UNIVERSITY OF BELGRADE
FACULTY OF ORGANIZATIONAL SCIENCES

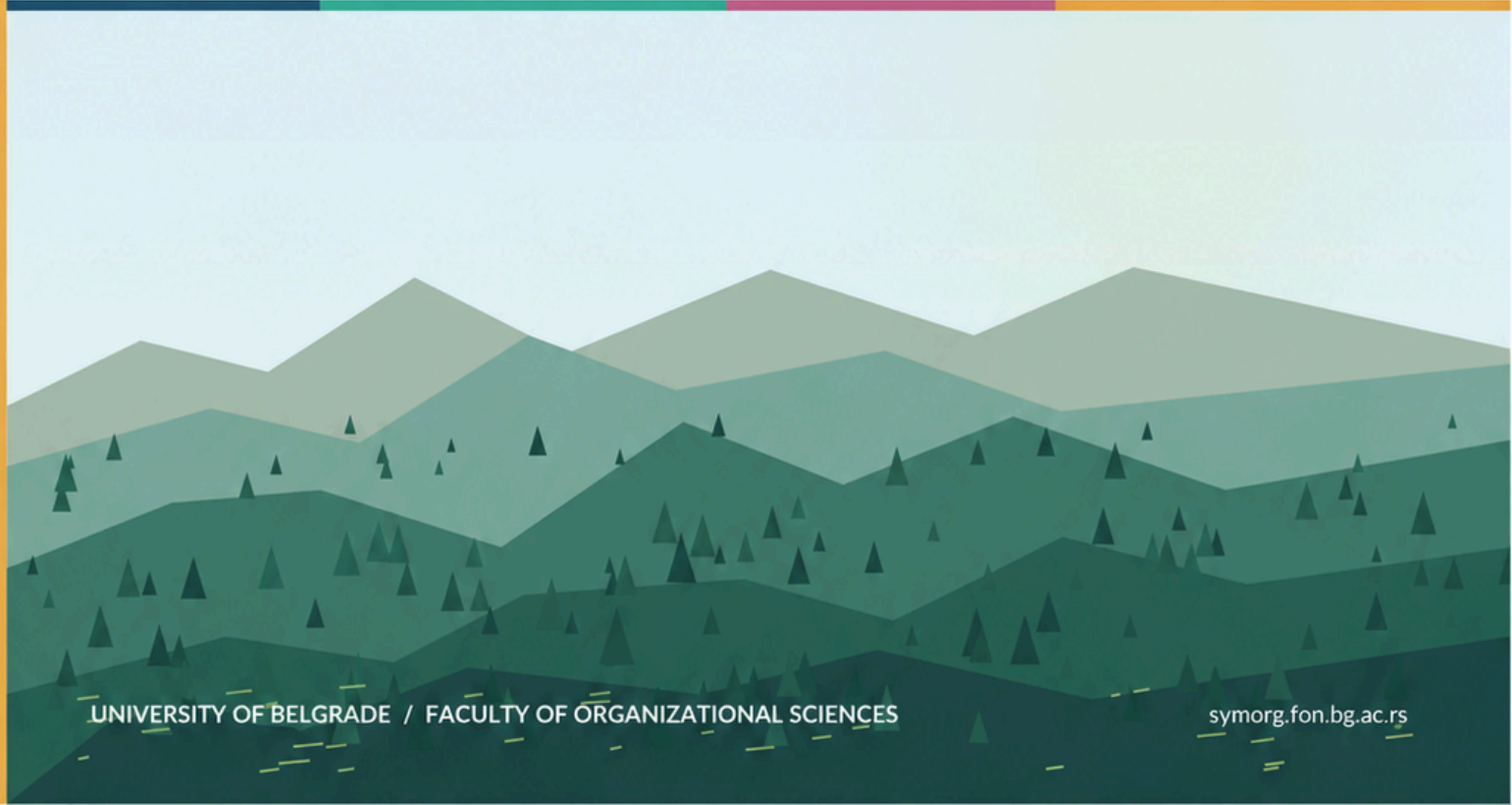
2026 symorg

20TH INTERNATIONAL SYMPOSIUM ON ORGANIZATIONAL SCIENCES

CONFERENCE PROCEEDINGS

Resilient and Intelligent Organizations
in Times of Permanent Disruption

10-12 JUNE 2026 | KOPAONIK, SERBIA



UNIVERSITY OF BELGRADE / FACULTY OF ORGANIZATIONAL SCIENCES

symorg.fon.bg.ac.rs

UNIVERSITY OF BELGRADE
FACULTY OF ORGANIZATIONAL SCIENCES
Belgrade, Serbia

20th INTERNATIONAL SYMPOSIUM ON ORGANIZATIONAL SCIENCES

SymOrg 2026

CONFERENCE PROCEEDINGS

Resilient and Intelligent Organizations
in Times of Permanent Disruption

June 10–12, 2026 • Kopaonik, Serbia

EDITORS
Boris Delibašić · Maja Glogovac
Gordana Savić · Manuel Mazzara

University of Belgrade, Faculty of Organizational Sciences
Belgrade • 2026

20th International Symposium on Organizational Sciences (SymOrg 2026)

SymOrg 2026 – Conference Procedings

Resilient and Intelligent Organizations in Times of Permanent Disruption
June 10–12, 2026 | Kopaonik, Serbia

Editors

Boris Delibašić — University of Belgrade, Faculty of Organizational Sciences, Serbia
Maja Glogovac — University of Belgrade, Faculty of Organizational Sciences, Serbia
Gordana Savić — University of Belgrade, Faculty of Organizational Sciences, Serbia
Manuel Mazzara — Innopolis University, Faculty of Computer Science and Engineering, Russia

Publisher

University of Belgrade, Faculty of Organizational Sciences

Jove Ilića 154, 11000 Belgrade, Serbia

Year

2026

Printed by

Grafika Galeb doo, Niš

Print run

20

ISBN

ISBN-978-86-7680-532-7

Designed by

Sandro Radovanović, Minja Marinović

Proceedings available at

https://symorg2026.fon.bg.ac.rs/papers/proceedings

2026 Faculty of Organizational Sciences, University of Belgrade. This proceedings and all included papers are licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0).

Publisher's note: The publisher makes no representations, either explicit or implied, regarding the accuracy of the information contained in this collection and cannot accept any legal responsibility for any errors or omissions that may have been made.

CIP - Каталогизација у публикацији, Народна библиотека Србије, Београд

004:005(082)(0.034.2)
004.8:005(082)(0.034.2)
658(082)(0.034.2)

INTERNATIONAL symposium on organizational sciences SymOrg 2026 (22 ; 2026 ; Kopaonik)
Resilient and intelligent organizations in times of permanent disruption [Elektronski izvor] : Conference proceedings / 22th International symposium on organizational sciences SymOrg 2026, June 10–12, 2026 • Kopaonik ; [conference organizer University of Belgrade, Faculty of Organizational Sciences] ; editors Boris Delibašić ... [et al.]. - Belgrade : University, Faculty of Organizational Sciences, 2026 (Niš : Grafika Galeb). - 1 elektronski optički disk (CD-ROM) : tekst, slika ; 12 cm

Sistemski zahtevi: Nisu navedeni. - Nasl. sa naslovnog ekrana. - Tiraž 20. - Preface / Boris Delibašić ... [et al.]. - Bibliografija uz svaki rad.

ISBN 978-86-7680-532-7
1. Delibašić, Boris, 1978- [уредник] [аутор додатног текста]
а) Менаџмент -- Зборници б) Предузећа -- Организација -- Зборници в) Информационе технологије -- Примена -- Зборници г) Вештачка интелигенција -- Примена -- Зборници

COBISS.SR-ID 205297161

PREFACE

This book contains papers presented at the 20th International Symposium on Organizational Sciences (SymOrg 2026), an international scientific conference held from June 10 to 12, 2026, in Kopaonik, Serbia. The volume is published in open access on Zenodo by the University of Belgrade, Faculty of Organizational Sciences, which organized the symposium, and it appears alongside two companion volumes of selected papers published in Springer's Communications in Computer and Information Science series as CCIS 3111 and CCIS 3112. Established in 1989, SymOrg has developed into a long-running international meeting place for researchers, educators, business leaders, practitioners, entrepreneurs, and public-sector representatives whose work advances organizational sciences.

The theme of SymOrg 2026, Resilient and Intelligent Organizations in Times of Permanent Disruption, reflected the conditions in which contemporary organizations operate. Disruption can no longer be treated only as an occasional shock followed by a return to stability. Technological acceleration, geopolitical and economic uncertainty, environmental pressures, changing workforce expectations, and the rapid diffusion of artificial intelligence have made adaptation a continuous organizational responsibility. The conference therefore approached resilience and intelligence as mutually reinforcing capabilities: resilience enables organizations to absorb shocks, learn, and renew themselves, while intelligence enables them to recognize change, make evidence-informed decisions, and translate knowledge into timely action.

The scientific program connected these questions across 15 tracks covering strategic communication, software engineering, blockchain, green and digital transformation, innovation and sustainable management, data and business analytics, organizational design, operations and logistics, e-business ecosystems, human resource management, lean management, strategy and project management, quality and standardization, finance, and generative artificial intelligence. This breadth is characteristic of organizational sciences: complex organizational problems rarely belong to a single discipline, and meaningful solutions require the integration of managerial, technological, economic, behavioral, and institutional perspectives.

SymOrg 2026 brought together contributions from 22 countries, involving 352 authors and 146 papers in the conference program. The review process began with 161 submissions. After withdrawals and retractions, 154 valid submissions passed through technical and editorial prescreening for formatting, double-blind requirements, scope, and publication ethics. Each valid paper received at least two independent reviews by members of the Program Committee, together with an additional technical review. Track Chairs considered these reviews, wrote a meta-review, and provided final recommendations. On this basis, 61 papers were selected for publication in the two Springer volumes, while the present volume collects a further 78 accepted papers, written by 204 authors, which were presented at the symposium and are published here in open access.

The papers gathered in this volume span all 15 conference tracks. They examine data and business analytics, decision support, and machine learning in organizational and clinical settings; e-business ecosystems, digital public services, and information system integration; blockchain, cryptographic protocols, and enterprise cybersecurity; generative artificial intelligence and autonomous agents in business process automation; green and digital transformation, sustainability reporting, and environmental performance; organizational design, resilience, and consulting practice; strategy, hybrid project management, and corporate entrepreneurship; quality management, standardization, and process improvement; human resource management and workplace learning; lean management; adaptive marketing and strategic communication; applied software engineering; and finance under conditions of uncertainty. Taken together, the contributions show that organizational intelligence is not produced by technology alone. It depends on the quality of data and models, but also on institutional capacity, responsible leadership, competent people, transparent processes, and the ability to coordinate knowledge across organizational boundaries.

SymOrg has always been more than a sequence of paper sessions. The 2026 program was designed to create several complementary spaces in which academic knowledge could meet organizational practice. The International Forum of Organizational Sciences, under the theme Unlocking Regional Potential Through Academia–Industry Collaboration, connected academic leaders from regional faculties with companies operating across the region. Its two-part format

addressed both the future development of organizational and management sciences and the transfer of organizational practices, experience, and knowledge between universities and industry.

The Business Forum brought together long-term strategic partners of the Faculty to discuss how resilient and intelligent organizations are built in conditions of uncertainty and technological transformation. The panel AI That Delivers: Turning Disruption into Advantage focused on practical experiences of implementing artificial intelligence and converting technological change into sustainable business value. The roundtable Competency-Based Education: The Voice of the Market addressed the alignment of higher education with evolving labor-market needs and the shared responsibility of universities and employers for developing relevant competencies.

Entrepreneurial and early-career development formed another important part of the program. Student teams from the fifth edition of the Route2Launch program presented startup ideas and demonstrated how academic knowledge and creativity can be developed into viable ventures. The Students Startup Achievement Panel explored entrepreneurial ecosystems in higher education, including living labs, open innovation, knowledge and technology transfer, and academia–industry collaboration. The Organizational Sciences Doctoral Colloquium gave doctoral students an interdisciplinary setting in which to present developing research, receive feedback from peers and senior scholars, and establish future research collaborations. It also included a workshop on frameworks for writing systematic literature reviews. Finally, the Business Practice Showcase created an interactive poster-based meeting point where companies and researchers presented achievements, best practices, research findings, and ideas with potential for application and commercialization.

The plenary program further enriched this exchange through contributions by Justin Paul, Alexander D. Stajković, Ulrich Harmes-Liedtke, Vasco Amaral, and Andrea Ivanišević. Across plenary lectures, thematic and poster sessions, forums, panels, roundtables, the doctoral colloquium, startup presentations, and informal encounters, SymOrg 2026 sought to connect research with the decisions and capabilities required by organizations in practice.

We are grateful to all authors for entrusting their work to SymOrg, to the reviewers and Track Chairs for their careful and constructive evaluations, and to the Program, Organizing, and Technical Committees for sustaining the academic and operational quality of the symposium. We thank the speakers, session chairs, moderators, mentors, volunteers, and participants whose contributions made the program intellectually rich and professionally engaging. We also acknowledge the academic, institutional, business, and media partners whose support strengthened the conference and its connection with the wider community. The publication of this volume was co-financed by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia.

We hope that this volume, together with the companion Springer volumes published as CCIS 3111 and CCIS 3112, will serve both as a record of the research presented at SymOrg 2026 and as a source of ideas for scholars and practitioners working to build organizations that remain capable, responsible, and forward-looking in times of permanent disruption.

July 2026

Prof Boris Delibašić
Prof Maja Glogovac
Prof Gordana Savić
Prof Manuel Mazzara

ORGANIZATION

Editors

Boris Delibašić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Maja Glogovac	University of Belgrade, Faculty of Organizational Sciences, Serbia
Gordana Savić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Manuel Mazzara	Innopolis University, Faculty of Computer Science and Engineering, Russia

Program Committee Chair

Boris Delibašić	University of Belgrade, Faculty of Organizational Sciences, Serbia

Organizing Committee

Maja Glogovac (Chair)	University of Belgrade, Faculty of Organizational Sciences, Serbia
Biljana Tošić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Dejana Nikolić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Milica Latinović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Nemanja Milanović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Sandro Radovanović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Željko Spasenić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Zoran Rakićević	University of Belgrade, Faculty of Organizational Sciences, Serbia

Technical Committee

Tea Borozan (Co-Chair)	University of Belgrade, Faculty of Organizational Sciences, Serbia
Sara Dimitrijević (Co-Chair)	University of Belgrade, Faculty of Organizational Sciences, Serbia
Aleksandra Vojvodić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Bojana Ivanović Bovan	University of Belgrade, Faculty of Organizational Sciences, Serbia
Boris Vajagić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Danilo Veličkov	University of Belgrade, Faculty of Organizational Sciences, Serbia
Dejan Markovski	University of Belgrade, Faculty of Organizational Sciences, Serbia
Dragana Avlijaš	University of Belgrade, Faculty of Organizational Sciences, Serbia
Đorđe Cvetković	University of Belgrade, Faculty of Organizational Sciences, Serbia
Filip Kužet	University of Belgrade, Faculty of Organizational Sciences, Serbia
Isidora Gaćeša	University of Belgrade, Faculty of Organizational Sciences, Serbia
Ivana Kužet	University of Belgrade, Faculty of Organizational Sciences, Serbia
Jovana Jović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Magdalena Majstorović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Marko Pepić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Marko Savićević	University of Belgrade, Faculty of Organizational Sciences, Serbia
Minja Marinović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Nela Rakić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Nemanja Đukić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Ognjen Nikolić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Petar Stanimirović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Selena Stanojević	University of Belgrade, Faculty of Organizational Sciences, Serbia

PROGRAM COMMITTEE

Aleksandar Jovović	University of Belgrade, Faculty of Mechanical Engineering, Serbia
Ariadna Llorens	Universitat Politècnica de Catalunya, Spain
Beata Marta Zielosko	University of Silesia in Katowice, Faculty of Science and Technology, Poland
Biljana Stošić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Bojan Ilić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Bojana Olgić Draženović	University of Rijeka, Faculty of Economics and Business, Croatia
Chase Thiel	University of Wyoming, College of Business, USA
Chi Maher	Ravensbourne University London, United Kingdom
Danica Lečić-Cvetković	University of Belgrade, Faculty of Organizational Sciences, Serbia
Dejan Mirakovski	Goce Delčev University, Faculty of Natural and Technical Sciences, North Macedonia
Dejan Petrović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Dejan Simić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Denisa Abrudan	West University of Timișoara, Faculty of Economics and Business Administration, Romania
Dragan Vukmirović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Dragoslav Slović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Eric Hertzler	Université Paris-Est Créteil, IAE Paris-Est School of Management, France
Gabriele Pizzi	University of Bologna, Italy
Goran Kuljanin	DePaul University, Driehaus College of Business, USA
Gordana Savić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Ivan Luković	University of Belgrade, Faculty of Organizational Sciences, Serbia
Ivana Kovačević	University of Belgrade, Faculty of Organizational Sciences, Serbia
Iztok Podbregar	University of Maribor, Faculty of Organizational Sciences, Slovenia
Jasmina Ćetković	University of Montenegro, Faculty of Economics, Montenegro
Jasmina Dlačić	University of Rijeka, Faculty of Economics, Croatia
Jelena Jovanović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Jelena Pokimica	Boise State University, USA
Jovan Filipović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Kurt Sandkuhl	University of Rostock, Germany
Manuel Mazzara	Innopolis University, Faculty of Computer Science and Engineering, Russia
Marija Janković	Université Paris-Saclay, CentraleSupélec, France
Marija Kuzmanović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Marijana Despotović-Zrakić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Marina Klačmer Čalopa	University of Zagreb, Faculty of Organization and Informatics, Croatia
Marko Mihić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Miha Marič	University of Maribor, Faculty of Organizational Sciences, Slovenia
Milica Kostić-Stanković	University of Belgrade, Faculty of Organizational Sciences, Serbia

Miloš Knežević	University of Montenegro, Faculty of Civil Engineering, Montenegro
Mladen Čudanov	University of Belgrade, Faculty of Organizational Sciences, Serbia
Nataša Petrović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Nenad Brkić	University of Sarajevo, Faculty of Economics, Bosnia and Herzegovina
Nikolina Ljepava	University of Khorfakkan, College of Business Administration, United Arab Emirates
Nina Begičević Ređep	University of Zagreb, Faculty of Organization and Informatics, Croatia
Ondrej Jaško	University of Belgrade, Faculty of Organizational Sciences, Serbia
Peter Marton	University of Žilina, Faculty of Management Science and Informatics, Slovakia
Polona Šprajc	University of Maribor, Faculty of Organizational Sciences, Slovenia
Predrag Trpeski	Ss. Cyril and Methodius University in Skopje, Faculty of Economics, North Macedonia
Sandra Jednak	University of Belgrade, Faculty of Organizational Sciences, Serbia
Sanja Bijakšić	University of Mostar, Faculty of Economics, Bosnia and Herzegovina
Saša Drezgić	University of Rijeka, Faculty of Economics, Croatia
Saša Lazarević	University of Belgrade, Faculty of Organizational Sciences, Serbia
Slađana Barjaktarović Rakočević	University of Belgrade, Faculty of Organizational Sciences, Serbia
Valentino Vranić	Pan-European University, Faculty of Informatics, Slovakia
Velimir Štavljanin	University of Belgrade, Faculty of Organizational Sciences, Serbia
Vesna Damnjanović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Zorica Bogdanović	University of Belgrade, Faculty of Organizational Sciences, Serbia

REVIEWERS

Achillefs Anagnostopoulos	University of Thessaly, Greece
Alenka Baggia	University of Maribor, Faculty of Organizational Sciences, Slovenia
Ana Kovačević	University of Belgrade, Faculty of Security Studies, Serbia
Ana Miličević	University of Belgrade, Faculty of Organizational Sciences, Serbia
Ana Poledica	University of Belgrade, Faculty of Organizational Sciences, Serbia
Ana Rakić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Ana Uzelac	University of Belgrade, Faculty of Transport and Traffic Engineering, Serbia
Andrija Petrović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Anida Zahirović Suhonjić	IPI Academy Tuzla, Bosnia and Herzegovina
Biljana Panić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Biljana Tošić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Bojan Jovanović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Bojan Tomić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Branka Novčić Korać	University of Belgrade, Faculty of Organizational Sciences, Serbia
Branka Rodić	Academy of Applied Studies Belgrade, Serbia
Catalin Vrabie	National University of Political Studies and Public Administration, Romania
Danijela Stojanović	Institute of Economic Sciences, Serbia
Dejana Nikolić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Dragan Pamučar	University of Belgrade, Faculty of Organizational Sciences, Serbia
Dragana Makajić-Nikolić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Dragana Stojanović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Dusan Barać	University of Belgrade, Faculty of Organizational Sciences, Serbia
Dušan Savić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Farhad Lotfi	Medical University of Vienna, Austria
Ilija Antović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Ivan Milenković	University of Belgrade, Faculty of Organizational Sciences, Serbia
Ivan Tomašević	University of Belgrade, Faculty of Organizational Sciences, Serbia
Ivana Dragović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Ivana Mijatović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Jelena Ruso	University of Belgrade, Faculty of Organizational Sciences, Serbia
Josip Čičak	University of Rijeka, Faculty of Economics and Business, Croatia
Jovan Krivokapić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Jovanka Vukmirović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Katarina Pažur Aničić	University of Zagreb, Faculty of Organization and Informatics, Croatia
Lena Đorđević Milutinović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Marija Bogićević Sretenović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Marija Todorović	University of Belgrade, Faculty of Organizational Sciences, Serbia

Marjan Senegačnik	University of Maribor, Faculty of Organizational Sciences, Slovenia
Marko Putniković	University of Belgrade, Serbia
Marko Suvajdžić	University of Florida, USA
Martin Kiselički	Ss. Cyril and Methodius University in Skopje, Faculty of Economics, North Macedonia
Milan Okanović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Milan Radojičić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Milan Vukićević	University of Belgrade, Faculty of Organizational Sciences, Serbia
Milena Popović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Milica Janković	University of Belgrade, School of Electrical Engineering, Serbia
Milica Jovanović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Milica Maričić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Milica Muhadinović	University of Montenegro, Faculty of Economics, Montenegro
Milos Danilović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Milos Jovanović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Miloš Milić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Milos Milosavljević	University of Belgrade, Faculty of Organizational Sciences, Serbia
Milos Radenković	Union University, School of Computing, Serbia
Nemanja Backović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Nemanja Milanović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Nikola Zornić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Nina Turajlić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Ognjen Pantelić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Pavle Milosević	University of Belgrade, Faculty of Organizational Sciences, Serbia
Pinaki Chakraborty	Netaji Subhas University of Technology, India
Radmila Janičić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Radul Milutinović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Sanja Marinković	University of Belgrade, Faculty of Organizational Sciences, Serbia
Saša Čegar	University of Rijeka, Faculty of Economics and Business, Croatia
Sava Cavoški	Sistemi, Serbia
Sergei Prokhorov	Institute for the History of Science and Technology, Russia
Slađana Janković	University of Belgrade, Faculty of Transport and Traffic Engineering, Serbia
Slavica Cicvarić Kostić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Slobodan Antić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Snežana Popović	Union University, School of Computing, Serbia
Sonja Radenković	Belgrade Banking Academy, Faculty of Banking, Insurance and Finance, Serbia
Stefan Komazec	University of Belgrade, Faculty of Organizational Sciences, Serbia
Tamara Vlastelica	University of Belgrade, Faculty of Organizational Sciences, Serbia
Tatjana Ivanović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Tatjana Stojadinović	Medicines and Medical Devices Agency of Serbia, Serbia
Uroš Šošević	University of Belgrade, Faculty of Organizational Sciences, Serbia

Veljko Dmitrović	University of Belgrade, Faculty of Organizational Sciences, Serbia
Veljko Jeremić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Violeta Cvetkoska	Ss. Cyril and Methodius University in Skopje, North Macedonia
Željko Spasenić	University of Belgrade, Faculty of Organizational Sciences, Serbia
Zoran Rakićević	University of Belgrade, Faculty of Organizational Sciences, Serbia
Zorica Mitrović	University of Belgrade, Serbia

PARTNERS

Program-engaged sponsors

Banca Intesa; Egzakta; Nelt Group; Raiffeisen Bank; Rivian and Volkswagen Group; AD IMLEK; Milsped Group; Elixir Zorka; EY; Oracle; Oktacron; MK Group; and Kampster.

Academic partners

University of Zagreb, Faculty of Organization and Informatics; University of Maribor, Faculty of Organizational Sciences; University of Montenegro, Faculty of Economics; and Academy of Management.

Institutional partners

Ministry of Science, Technological Development and Innovation of the Republic of Serbia; Innovation Fund of the Republic of Serbia; Regulatory Authority for Electronic Communications and Postal Services (RATEL); and Institute for Artificial Intelligence Research and Development of Serbia.

Sponsors

Belgrade Airport; Coca-Cola HBC Serbia; Aparteko; Štark; Vinarija Jeremić; Bambi; Nestlé; and Nutrino Lab.

Media partners

NIN; Bloomberg Adria; Tanjug; Una TV; BizLife; and Newsmax Balkans.

Conference website

https://symorg2026.fon.bg.ac

The publication of this volume was co-financed by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia.

TABLE OF CONTENTS

ADAPTIVE MARKETING AND STRATEGIC COMMUNICATIONS IN DISRUPTIVE ENVIRONMENTS.....1

GAMIFICATION IN HIGHER EDUCATION: HOW GAME ELEMENTS INFLUENCE USER MOTIVATION, ENGAGEMENT, AND EXPERIENCE.....2
Mina Živanović, Velimir Štavljanin, and Dejana Nikolić

DIGITAL COMMUNICATION STRATEGIES OF IVF CLINICS IN INFORMING ABOUT GAMETE DONATION: A COMPARATIVE ANALYSIS OF INTERNATIONAL AND SERBIAN WEBSITES.....13
Biljana Djordjevic

GENDER DIFFERENCES IN CONSUMERS' PERCEPTION OF ENVIRONMENTAL AND SOCIAL CSR ACTIVITIES.....22
Anja Vučić

CATCH IT IF YOU CAN: MEASURING VISUAL ATTENTION OF PRODUCT PLACEMENT IN MOVIES USING EYE TRACKING.....33
Kristina Hristova, Nikolaos Dimitriadis, Marko Ćircović, Matteo Venerucci, and Omid Asgari

APPLIED SOFTWARE ENGINEERING.....46

DESIGNING A DOMAIN-SPECIFIC CONVERSATIONAL AI FOR A BUILDING MATERIALS RETAILER.....47
Askar Akhmetkhanov and Marko Pezer

TRANSFORMATION OF UNSUSTAINABLE STRUCTURES INTO SUSTAINABLE STRUCTURES USING THE BDPSS FORMALISM.....64
Ana Trujić, Dragica Ljubisavljević, Milica Bačić, Miloš Milić, Dušan Savić, Ilija Antović and Siniša Vlajić

APPLICATION OF FEATURE MODELS IN MODULAR SOFTWARE DEVELOPMENT.....77
Isidora Vidojević, Vladimir Belča and Đorđe Nestorović

BLOCKCHAIN TECHNOLOGY IN BUSINESS AND INFORMATION SYSTEMS.....86

CRYPTOGRAPHIC MODEL CHAINING FOR PRIVACY-PRESERVING AI ENSEMBLES AND CROSS-CHAIN COLLABORATION.....87
Miloš Živadinović, Ivan Milenković, and Dejan Simić

INTENTION TO ADOPT BLOCKCHAIN TECHNOLOGY FOR IMPROVING ASSESSMENT QUALITY IN HIGHER VOCATIONAL EDUCATION IN SERBIA.....103
Miloš Mijić and Branko Ćebić

DEVELOPMENT OF DAPP FOR FASHION E-COMMERCE SUPPLY CHAINS.....119
Danica Sovtić, Aleksandra Trpkov, Marta Magovčević, Božidar Radenković, Sergei Prokhorov and Aleksandra Labus

HYBRID RISK MODELING FRAMEWORK FOR COMPARING POW AND POS CONSENSUS MECHANISMS....152
Jovan Petrović, Dušan Mitrović, Bojan Marčeta

MAPPING AI AGENTS IN ENTERPRISE CYBERSECURITY: A TAXONOMY OF ARCHITECTURES, AUTONOMY AND ADOPTION PATTERNS.....164
Petar Ilić, Uroš Šošević and Dušan Mitrović

A COMPARISON OF FLECTRA, DOLIBARR AND ODOO ERP SYSTEMS BASED ON HR MODULE FUNCTIONALITIES.....178
Srđa Bjeladinović and Marko Asanović

CHALLENGES OF GREEN AND DIGITAL TRANSFORMATION.....195

THE INFLUENCE OF ESG PERFORMANCE ON ATTRACTING FOREIGN DIRECT INVESTMENTS.....196
Luka Petrović and Marko Ćirović

IMPACT OF EXTREME TEMPERATURES ON CARDIOVASCULAR MORBIDITY AND MORTALITY.....209
Marjan Senegačnik and Davorin Žnidarič

ASSESSING SUSTAINABILITY IN SURFACE WATER TREATMENT: A CIRCULAR ECONOMY PERSPECTIVE WITH DATA-DRIVEN DECISION SUPPORT 222
Nevena Purić, Jelena Andreja Radaković, Vladimir Mandić, Maja Đolić

MEASURING EFFICIENCY IN GREEN AND SUSTAINABLE ENERGY ACROSS EUROPE USING DATA ENVELOPMENT ANALYSIS 237
José Rodrigo Juárez Brenner

CREATIVITY, INNOVATION AND SUSTAINABLE MANAGEMENT **251**

WHAT INFLUENCES SUSTAINABILITY REPORTING IN THE REPUBLIC OF SERBIA? AN EMPIRICAL ANALYSIS 252
Bojana Novićević Čečević, Jelena J. Stanković, Žarko Popovic

TO BE RESILIENT, COMMITTED OR BRANDED: EXPLORING PREDICTORS OF SUSTAINABLE CAREER 265
Ivana Kovačević, Denisa Abrudan, Mateja Manojlović, and Achilleas Anagnostopoulos

DATA AND BUSINESS ANALYTICS **283**

A TWO-LEVEL FRAMEWORK FOR JOINT OPTIMIZATION OF FEATURE ENGINEERING AND HYPERPARAMETER TUNING IN PREDICTIVE MAINTENANCE CLASSIFICATION 284
Martin Guyon

DIFFERENCES IN STATISTICS ANXIETY ACROSS ACADEMIC YEARS: THE ROLE OF PERCEIVED USEFULNESS OF STATISTICS 300
Luna Radević, Milica Maričić and Tatjana Von Rosen

REPRODUCIBILITY IN IS RESEARCH AS A WEAK SIGNAL: A BIBLIOMETRIC FORESIGHT ANALYSIS 310
Sara Koprivica, Aleksandra Hornjak, and Darko Stefanovic

ANALYSIS OF THE LEVEL OF ACCEPTANCE AND APPLICATION OF ARTIFICIAL INTELLIGENCE IN MEDICAL DIAGNOSTICS 325
Tijana Ponjarac

DESIGNING RESILIENT AND INTELLIGENT ORGANIZATIONS IN TIMES OF PERMANENT DISRUPTION **347**

RESILIENCE AS ORGANIZING WHEN STRUCTURE FAILS 348
Ana Aleksić Mirić and Maksim Vojvodić

SPECIFIC CHARACTERISTICS OF THE CONSULTING PROCESS IN SMES: CONCEPTUAL FRAMEWORK 361
Jovan Krivokapić and Milan Jovanović

RESEARCH-INTENSIVE TEAMS AND ORGANIZATIONAL RESILIENCE: THE ROLE OF COGNITIVE ROUTINES 371
Isidora Simeunović

DIGITAL OPERATIONS AND LOGISTICS MANAGEMENT **382**

DISTRIBUTION STRATEGIES IN E-COMMERCE: INSIGHTS FROM TEMU AND ALIEXPRESS 383
Anđela Zdravić, Miljana Luković, Teodora Đokić and Biljana Cvetić

E-BUSINESS ECOSYSTEMS AND TECHNOLOGIES **401**

CONTEMPORARY E-COMMERCE AND AI ECOSYSTEM – THEORETICAL AND PRACTICAL ISSUES 402
Sonja Vučenović, Nikola Macura, and Marija Pavličević

SERVICE CATALOGUES AND METAREGISTERS AS ENABLERS OF HEALTH SERVICE BUS INTEGRATION: THE ALIMS CASE STUDY 414
Tatjana Stojadinović, Ilija Antović, Svetlana Jovanović and Tijana Mladenović

FROM DIGITAL JUDICIARY TO AI-AUGMENTED JUSTICE: APPLYING THE BERKANA TWO LOOPS MODEL TO THE SERBIAN JUDICIARY 426
Dušan Kuzmanović, Ana Đilas, and Zorica Bogdanović

DEVELOPMENT OF AN AI-DRIVEN MOBILE SERVICE FOR SMART COOKING 444
Filip Stevanović and Dušan Kostić
A WEB APPLICATION FOR UNDERSTANDING CONFLICTING HISTORICAL NARRATIVES WITH USER CONTRIBUTIONS AND GENAI 457
Petar Tomić and Miloš Jolović
ZERO-KNOWLEDGE PROOFS IN E-COMMERCE: SURVEY OF APPLICATIONS AND A CASE STUDY IN PRIVACY-PRESERVING AGE VERIFICATION 474
Amar Mujezinović and Aleksandar Joksimović
A CCTS-BASED CONCEPTUAL FOUNDATION FOR EXPLAINABLE SCHEMA MATCHING DECISIONS 491
Ivan Jovanović, Milica Škembarević Sretenović, Ivan Luković, Nenad Ivezić, Anđela Karašićević
DIGITAL TRANSFORMATION OF A PRIVATE HIGHER EDUCATION INSTITUTION: EVIDENCE FROM A SERBIAN UNIVERSITY AND IMPLICATIONS FOR THE SECTOR 509
Miloš Vrbanac
APPLICATION OF 3D INDOOR NAVIGATION SYSTEMS IN ACADEMIC ENVIRONMENTS 520
Danica Živković, Dušan Kostić and Marijana Despotović-Zrakić
TOWARD A SECURE-BY-DESIGN HIGHER EDUCATION ECOSYSTEM: DECENTRALIZED IDENTITY, PRIVACY, AND TRUST 532
Tamara Naumović, Danica Živković, Marija Janković, Dušan Barać
ARTIFICIAL INTELLIGENCE FOR DIGITAL TRANSFORMATION IN SPORT: AN INTEGRATED CONCEPTUAL MODEL FOR ELITE SPORT, RECREATIONAL SPORT, AND PUBLIC HEALTH 547
Aleksandar Đurić

HUMAN RESOURCE MANAGEMENT FOR BUILDING RESILIENT EMPLOYEES **562**

A MIXED METHODS STUDY OF LEARNING ORGANIZATIONS: IMPLICATIONS FOR HEALTH CARE PRACTICE 563
Jelena Pokimica, Jelena Ruso, Indiana Laub, and Aleksandar Milovanovic
UNRAVELLING THE EMPLOYEE RESILIENCE CONCEPT: THE SUPREMACY OF EFFECTIVE LEARNING AT WORK 580
Vuk Mirčetić, Ivana Kovačević, and Marko Slavković
EXPLORING THE RELATIONSHIP OF SELF-DIRECTED LEARNING AND EMPLOYEE RESILIENCE IN THE AI ERA 595
Emilija Jeremić

ORGANIZATIONAL STRATEGY AND HYBRID PROJECT MANAGEMENT **606**

INTEGRATING DESIGN THINKING INTO THE INNOVATION FUNNEL: A FRAMEWORK FOR IMPROVING INNOVATION PERFORMANCE IN ENTREPRENEURSHIP 607
Aleksandar Ivković and Marko Mihić
THE INFLUENCE OF EXPERIENCE AND AGE ON PERCEPTIONS OF MANAGERIAL FACTORS IN THEATRE PROJECTS 620
Dragana Pantić and Dragan Bjelica
TOWARD INTEGRATED IT RESILIENCE: SYNERGIZING RISK MANAGEMENT, MATURITY ASSESSMENTS, AND IT AUDIT 636
Kristijan Lazić, Dragan Jovičić, Vladan Pantović, Zoran Keković

QUALITY MANAGEMENT AND STANDARDIZATION **647**

THE IMPACT OF THE APPLICATION OF QUALITY MANAGEMENT SYSTEM STANDARDS IN AUTOMOTIVE SUPPLY CHAINS 648
Simo Denić and Jelena Ruso
CLOSING THE GOVERNANCE GAP: WHY ISO/IEC 42001 COMPLETES THE ISO MANAGEMENT SYSTEM TRILOGY 672
Nikola Vojtek and Ivana Mijatović

CONTEXT AND IDENTITY IN HIGHER EDUCATION INSTITUTIONS IN INDUSTRY 5.0 686
Nedeljka Živković, Isidora Milošević, Ana Kićanović and Olesea Plotnic
INTEGRATING ARTIFICIAL INTELLIGENCE INTO QUALITY MANAGEMENT: OPPORTUNITIES, CHALLENGES, AND FUTURE DIRECTIONS 695
Mladen Đurić
IMPLEMENTATION OF THE DMAIC METHODOLOGY FOR QUALITY IMPROVEMENT IN AVIATION INDUSTRY: CASE STUDY OF "BELGRADE AIRPORT" 700
Mila Marković, Mladen Đurić, and Jovana Dimitrijević

REDESIGNING FINANCE FOR PERMANENT DISRUPTION **718**

WHEN BRAND VALUATION FAILS: GEOPOLITICAL RISK BLINDNESS AND PRIVATE EQUITY DUE DILIGENCE 719
Marko Jovanović, Maja Đurica, and Miha Marič
LINKING FINANCIAL LITERACY TO FINANCIAL WELL-BEING: EVIDENCE FROM SERBIA 731
Sladjana Barjaktarovic Rakocevic, Rade Rakocevic, Nela Rakic
THE WALKING DEAD: CAUSES AND CONSEQUENCES OF THE EMERGENCE OF THE ZOMBIE ECONOMY 743
Dario Maradin, Stella Suljić Nikolaj, Vesna Buterin, and Sara Sadrić

TRANSFORMATIONAL POTENTIAL OF GENERATIVE AI **758**

TRANSFORMATIVE POTENTIAL OF ARTIFICIAL INTELLIGENCE IN ECOLOGICAL AND ECOSYSTEM MANAGEMENT: AN ANALYSIS OF SUCCESSFUL SUSTAINABLE SOLUTIONS 759
Davorin Žnidarič and Marjan Senegačnik
MULTI-AGENT OBJECT DETECTION FRAMEWORK BASED ON RASPBERRY PI YOLO DETECTOR AND SLACK-OLLAMA NATURAL LANGUAGE INTERFACE 772
Vladimir Kalušev, Branko Brkljač, and Milan Brkljač
AI AGENTS IN BUSINESS PROCESS AUTOMATION: A SCENARIO-BASED COMPARATIVE ANALYSIS OF RPA AND HYBRID AUTOMATION IN FINANCIAL PROCESSES 792
Dragan Metikoš, Slobodan Antić, and Lena Đorđević Milutinović
TOWARDS AN ADAPTIVE HIERARCHICAL DEEP REINFORCEMENT LEARNING ARCHITECTURE FOR RESILIENT FINANCIAL SYSTEMS IN NON-STATIONARY MARKETS 806
Dragisa Stanojevic and Dragan Vukmirovic
EXPLORING DIFFERENCES IN GENERATIVE AI ADOPTION ACROSS EUROPEAN COUNTRIES USING STRUCTURAL EQUATION MODELING 824
Miljana Luković, Milica Maričić, and Miloš Danilović
CHATBOTS IN INTEGRATED MARKETING COMMUNICATIONS IN HIGHER EDUCATION: STRATEGIC IMPLICATIONS AND IMPLEMENTATION CHALLENGES 841
Dragana Todovic Milanovic, Velimir Stavljanin, Srdjan Masic and Milica Kostic-Stankovic

SHORT PAPERS **856**

SAP REINSURANCE MANAGEMENT MODULE CUSTOMIZATION FOR AUTOMATED INTEGRATION OF STRUCTURAL ELEMENTS 857
Katarina Mladenović, Ivan Jovanović, Srđa Bjeladinović
MOTIVATION FOR DIGITALIZATION IN AGRICULTURE: SIGNIFICANCE AND OPPORTUNITIES FOR ENHANCEMENT IN THE SOUTHERN GREAT PLAIN REGION OF HUNGARY 861
Patrik Urbán
QUO VADIS, HUMANITAS? MAPPING THE HUMAN COST OF ENVIRONMENTAL DEGRADATION THROUGH THE SDG FRAMEWORK 865
Nataša Petrović, Jelena Andreja Radaković, Aleksandar Jovović, Boris Najman, Polona Šprajc
ARE CHINESE UNIVERSITIES ENTREPRENEURIALLY ORIENTED? 869
Zelimir Todorovic

BARGAINING SOLUTIONS FOR FAIR CLASSIFICATION: A COMPARISON OF NASH AND KALAI SMORODINSKY SOLUTIONS.....873
Sandro Radovanović, Milan Vukićević, Boris Delibašić, Milija Suknović
FROM PREDICTION TO DECISION: A POSITION ON INTEGRATING MACHINE LEARNING AND DECISION SUPPORT SYSTEMS.....877
Boris Delibašić, Sandro Radovanović, Milija Suknović
REVENUE OPTIMIZATION OF NEW YORK CITY TAXI SERVICES USING LINEAR PROGRAMMING AND METAHEURISTICS.....881
Isidora Gaćeša, Gordana Savić, Violeta Cvetkoska
INTEGRATING NEURAL NETWORKS IN CARDIAC BIOPHYSICAL MODELING.....886
Bogdan Milićević, Miloš Ivanović, Boban Stojanović, Miljan Milošević, Miloš Kojić, and Nenad Filipović
INTEGRATING BIOMARKERS, VOICE ANALYSIS, AND SENSORS: STRATIFYHF DECISION SUPPORT PLATFORM.....891
Tijana Geroski, Anđela Blagojević, Ognjen Pavić, Lazar Dašić, Nemanja Marković, Nduka Okwose, Đorđe Jakovljević, Nenad Filipović
APPLICATION OF FUZZY LOGIC AND RISK MATRIX IN THE PRELIMINARY ASSESSMENT OF INDIVIDUALS' CREDITWORTHINESS.....896
Ilija Ivezić, Dragan Pamučar, Bisera Andrić Gušavac
INNER GOVERNANCE: UTILIZING TRAIT EMOTIONAL INTELLIGENCE TO ANCHOR RESILIENCE IN ORGANIZATIONS.....900
Teklehaimanot Embaye
LEAN STARTUP METHODOLOGY FOR INNOVATION DEVELOPMENT AND PRODUCTION.....906
Jovana Adžić
INTEGRATING STRATEGIC MANAGEMENT AND CORPORATE ENTREPRENEURSHIP IN PUBLIC ENTERPRISES: EVIDENCE FROM SERBIA.....911
Sara Dimitrijević, Dejan Petrović
FROM QUARTERLY ACHIEVEMENT TO RENEWED ENGAGEMENT: BEHAVIORAL AND MOTIVATIONAL PATTERNS DURING STRATEGY IMPLEMENTATION IN A LARGE PUBLIC UTILITY.....916
Goran Ličanin 1, Mladen Čudanov, Zorica Mitrović
STUDENTS' ENVIRONMENTAL AWARENESS: LINKING POLLUTANTS, EDUCATION, SCIENCE, AND ISO 14001 STANDARD.....920
Dušan V. Trajković, Filip Gaus, Teodora Križanić, Jelena Ruso
APPLIED AI FOR CARDIOVASCULAR DISEASE.....925
Nenad Filipovic, Djordje Jakovljevic, Zoran Bosnic and Dimitrios Boucharas
CAN AI REPLACE HUMAN EXPERTS IN CROWD VOTING? A COMPARATIVE STUDY USING BARGAINING-BASED AGGREGATION.....929
Ana Vukicevic, Milan Vukicevic, Sandro Radovanovic, Boris Delibasic, Milija Suknovic

THEMATIC SECTION

Adaptive Marketing and Strategic Communications in Disruptive Environments

Resilient and Intelligent Organizations in Times of Permanent Disruption

SYMORG 2026 • CONFERENCE PROCEEDINGS

Original Research Article *https://doi.org/10.5281/zenodo.23472128*

Gamification in Higher Education: How Game Elements Influence User Motivation, Engagement, and Experience

Mina Živanović[1][0009-0003-3098-9753], Velimir Štavljanin[2][0000-0001-5450-806X], and Dejana Nikolić[3][0000-0002-4772-7464]

[1] Academy of Technical and Art Applied Studies Belgrade, Belgrade, Republic of Serbia, zivanovic654@gmail.com
[2] University of Belgrade, Faculty of Organizational Sciences, Belgrade, Republic of Serbia, velimir.stavljanin@fon.bg.ac.rs
[3] University of Belgrade, Faculty of Organizational Sciences, Belgrade, Republic of Serbia, dejana.nikolic@fon.bg.ac.rs

Abstract. In contemporary higher education, improving student experience and increasing student engagement and motivation are key challenges for educational institutions. One of the approaches that is gaining more and more attention is gamification, which implies the application of game elements in contexts which are not game related, with the aim of improving the learning process. The aim of this paper is to analyze the impact of gamification on the motivation, engagement and experience of students in higher education, based on a narrative review of the relevant literature. The results of literature analysis indicate that gamification can contribute to increasing student motivation, which further encourages their engagement and leads to a more positive overall learning experience. The contribution of the paper is reflected in the integration of existing theoretical knowledge and the development of a theoretical basis for future empirical research, as well as for the improvement of practice in higher education.

Keywords: Gamification, Higher education, Motivation, Engagement, Learning experience.

1 Introduction

In today's conditions of digital transformation and intense competition in the higher education sector, universities face the need to redefine the way they create and deliver value to students. Students are increasingly seen not only as participants in the educational process, but also as users of services who form expectations, evaluate quality and make decisions analogous to consumers in the market.

In this context, improving student experience has become a central objective of higher education institutions. Accordingly, universities are increasingly focused on optimizing the teaching–learning process and fostering higher levels of student engagement, participation and motivation [1, 2, 3]. One approach that has attracted considerable attention in this regard is gamification, commonly defined as the application of

game elements in non-game contexts, such as education, with represents a practical tool for the improvement of the learning process, particularly by strengthening motivation and active involvement of students [4, 5].

Gamification has gained significant traction among educators due to its potential to make learning more dynamic and engaging, which has consequently led to growing body of research with this regard [6]. Its development has been further supported by advances in digital technologies and the increasing popularity of video games, which have become more immersive and interactive over time [7, 8]. As a result, the transfer of game design principles into educational settings has enabled the emergence and expansion of gamification as an innovative pedagogical approach [9].

Building on the growing body of research on gamification in higher education, this paper aims to provide a comprehensive conceptual understanding of how game elements influence key behavioral outcomes, namely student motivation, engagement, and overall experience. As a conceptual framework, the paper proposes that gamification has a direct effect on increasing student motivation and more active engagement of students in teaching activities, while contributes to a more positive perception of the learning process and a better overall student experience.

The paper is based on a narrative literature review encompassing scholarly publications from the fields of education, educational technology and consumer behavior. Relevant sources were identified through search of Google Scholar database, using keywords (such as “gamification”, “higher education”, “student motivation”, “student engagement”, and “student experience”) and the priority was given to peer-reviewed journal articles and conference proceedings tightly related to the topic. Following a narrative review approach, the selected literature was analyzed and synthesized to identify key findings and emerging trends regarding the application of gamification in higher education. Rather than providing a systematic assessment of empirical evidence, this review paper seeks to integrate fragmented insights into a coherent framework that explains the role of gamification as an effective mechanism within educational contexts which enhances the perceived value of the educational service. By doing so, the paper contributes to bridging the gap between theory and practice and provides a foundation for future empirical research in this domain. Furthermore, the study contributes to regional research by addressing a topic which has, so far, received limited attention in the context of Southeast European higher education systems, thereby providing a theoretical basis for future empirical investigations.

2 The concept of gamification in higher education

Two dominant perspectives on the use of games in higher education can be identified in literature: game-based learning, which involves the use of actual games to support teaching and learning processes, and gamification, which refers to the incorporation of game-design elements into non-game contexts [10]. While game-based learning relies on fully developed games as instructional tools, gamification focuses on selectively integrating specific elements of game design to enhance user engagement and motivation.

Gamification is mostly defined following the definition provided by Deterding, Dixon, Khaled and Nacke [11], who described it simply as: “the use of game design elements in non-game contexts”. Those authors emphasize several important distinctions: gamification involves the use rather than extension of games, focuses on design elements rather than full-fledged games or technologies, and applies these elements across various contexts irrespective of their purpose or medium. Similarly, gamification is also described as “the application of game features, mainly video game elements, into non-game context for the purpose of promoting motivation and engagement in learning” [12]. According to other definition, gamification is explained as “the art of changing human activity to a game like environment to gain user engagement” [13]. In educational settings, gamification typically entails the integration of elements such as point systems, peer competition, teamwork, and leaderboards to encourage participation, facilitate knowledge acquisition, and support learning assessment [14].

Within higher education, gamification is increasingly employed as a means of transforming traditional teaching approaches into more interactive, engaging and dynamic student-centered experience. Its growing adoption has been facilitated by the immediacy of feedback mechanisms and the implementation of reward systems, which enhance user interaction and sustained participation [7]. Common gamification elements, such as points, badges, leaderboards, levels, and challenges, affect student behavior through a variety of motivational mechanisms. Empirical evidence suggests that combinations of these elements can positively influence students' perceptions, academic performance, attendance, goal orientation and attitude towards learning [6, 15, 16, 17].

Among the wide range of game elements, four are frequently highlighted as particularly effective in the higher education context: narrative, challenge, progression and feedback. Narrative involves the use of storytelling or real-world scenarios to contextualize learning; challenge refers to tasks that are both stimulating and achievable; progression denotes the structured flow of activities that maintains learners’ motivation; and feedback encompasses timely and targeted responses that support learning and improvement [18].

Finally, recent developments in the field point to an increasing emphasis on customization and personalization within gamified learning environments, enabling the adaptation of game elements to individual learner preferences and needs, thereby further enhancing their effectiveness [19].

3 Psychological and behavioral outcomes of gamified learning

The integration of game-based approaches into teaching and learning has been widely recognized as a practical and effective strategy for enhancing educational outcomes [9]. In particular, gamification has emerged as a valuable pedagogical tool in higher education, with numerous studies highlighting its potential to improve key behavioral and psychological outcomes, including student engagement, motivation, satisfaction, and overall learning experience [3, 20, 21].

The existing literature suggests a sequential and interrelated relationship among these constructs. Specifically, gamification can enhance students’ motivation, which in

turn fosters higher levels of engagement; increased engagement then contributes to greater satisfaction, ultimately leading to a more positive and holistic educational experience. In this context, gamification, when embedded within the design of educational services, can act as a critical tool enabling these interconnected outcomes.

3.1 Gamification and student motivation

A substantial body of literature highlights the positive impact of gamification on student motivation in higher education. Importantly, gamification operates through both extrinsic and intrinsic motivational pathways [22, 23]. While extrinsic elements, such as rewards and rankings, encourage participation, intrinsic elements, such as a sense of progress, autonomy, and achievement, contribute to deeper engagement and sustained involvement in the learning process.

Research findings indicate that gamified learning environments can significantly enhance students' motivation and learning achievements compared to traditional approaches. For example, experimental evidence shows that students exposed to gamified classroom settings demonstrate higher levels of motivation and more positive attitudes towards learning activities [24]. Gamification is often perceived as an innovative and enjoyable method of delivering educational content, which further reinforce students' willingness to participate and engage [14]. In their study, Zourmpakis, Kalogiannakis, and Papadakis [25] demonstrate that adaptive gamification environments increase students' motivation to learn, as learners respond positively to personalized game elements embedded within instructional activities. Similarly, incorporating gamification into post-secondary education has been shown to stimulate motivation by increasing class attendance and participation, which are closely related to improved student performance [26].

The motivational potential of gamification can be further explained through its ability to address fundamental psychological needs, such as competence, autonomy, and relatedness. When these needs are satisfied, students are more likely to adopt deeper cognitive learning strategies and achieve better academic outcomes [27]. In this sense, gamification does not merely make learning more enjoyable but also creates a dynamic and supportive environment that fosters sustained motivation.

However, the effects of gamification on motivation are not uniform across all students. Research indicates that its effectiveness may depend on whether they are motivated intrinsically or extrinsically [28]. Well-designed gamified systems, which allow for autonomy in task selection, progression through levels and social interaction, can enhance both students' motivation and performance, particularly in technically demanding courses [13].

Finally, broader evidence confirms the central role of motivation as a key mechanism through which gamification influences learning outcomes. A systematic literature review by Jaramillo-Mediavilla et al. [29] concludes that gamification significantly contributes to the development of knowledge, skills and academic competencies of students, by fostering motivation through interactive and engaging learning experiences.

Studies conducted in the region also confirm that gamification can have a significant positive impact on the motivation of students in higher education. Analyzing the possi-

bilities of implementing gamification in higher education institutions in North Macedonia, Kiselicki and Josimovski [30] conclude that game elements can contribute to strengthening both intrinsic and extrinsic motivation of students, especially through a sense of achievement, advancement and recognition. Similar findings have been recorded in recent research in Serbia, which indicates that the application of gamified activities in an online environment increases students' interest in teaching content, encourages more active participation and contributes to maintaining motivation during the learning process [31]. Furthermore, the results of the study conducted in Bosnia indicate a significant influence of gamification on students' motivation to learn foreign language, as well as on their overall academic achievements [32].

Taken together, these findings underline that motivation represents a core pathway through which gamification enhances student engagement and academic achievement.

3.2 Student engagement in gamified settings

From a consumer behavior perspective, engagement reflects the degree of users' active involvement in interactions with products or services. Student engagement is commonly defined as the degree to which students are actively involved in learning activities that are empirically linked to desired educational outcomes. This multidimensional construct encompasses academic, emotional, and behavioral components, reflecting not only students' participation but also their cognitive investment and affective commitment to learning. In educational settings, engagement manifests through students' attention, participation, and emotional commitment to learning activities. In this context, gamification has emerged as an innovative pedagogical strategy, as the integration of game elements can effectively stimulate student involvement and enhance engagement across these dimensions [33].

A growing body of empirical research supports the positive impact of gamification on student engagement (e.g. [34, 35]). Increased engagement, in turn, has been shown to positively influence academic performance and knowledge acquisition [36].

Experimental studies using pre- and post-test measures have demonstrated significant improvements in students' engagement levels, academic performance, and attitudes towards learning when gamified approaches are applied, particularly in subjects such as mathematics [37]. Gamification elements, such as challenge, rewards, and interactive tasks encourage proactive engagement and enhance enjoyment in the learning process [18]. Similarly, evidence from technology-enhanced learning environments indicate that gamified systems, incorporating elements such as ranking, points, and badges can significantly influence students' behavioral engagement, including participation and persistence in learning activities [38].

Beyond behavioral aspects, recent studies highlight the broader impact of gamification on multiple dimensions of engagement. For instance, findings indicate that specific game elements are positively associated with students' cognitive, emotional, and performance-related engagement, particularly in online learning contexts [39]. In addition, gamified tools have been recognized as an interesting way to motivate students, and

have been shown to increase attendance, participation and overall academic performance by fostering a sense of competition and achievement among students [40]. Positive perceptions of gamified learning environments are further reflected in high levels of students' satisfaction and sustained engagement in e-learning settings [41]. The gamification model developed by Urh et al. [42] shows that the integration of elements such as points, rankings and prizes in e-learning contributes to greater involvement of students in teaching activities and encourages their continuous participation in the learning process. From the perspective of higher education, gamification is seen as a mechanism that increases interactivity and active participation of students, which is of particular importance in environments where it is challenging to maintain attention and interest over a long period of time. The improvement of student engagement by implementing a learning platform with gamification elements was found in the study conducted by Milenković et al. [43] in the Serbian context.

However, despite the generally positive findings, the literature also points to certain inconsistencies. While gamification is widely recognized for enhancing behavioural engagement, its influence on deeper cognitive and emotional dimensions remains less conclusive. Moreover, the mechanisms through which gamification translates into meaningful learning outcomes are still not fully understood, indicating a notable gap between theoretical conceptualizations and practical applications [44]. Despite such inconsistencies, overall, it can be concluded that previous empirical findings dominantly suggest that gamification holds significant potential to enhance student engagement by creating more interactive, motivating, and immersive learning environments.

3.3 Student experiences related to gamified learning environments

In addition to motivation and engagement, gamification plays an important role in shaping the overall student experience [45]. Gamification enhances the learning experience by enabling interactive, feedback-rich, and low-risk environments, where students can experiment, make decisions, and receive immediate responses [14]. Studies suggest that factors such as positive attitude, good experience and perceived usefulness and performance expectancy of gamified systems contribute to improved student performance and stronger intentions to adopt such approaches [4, 7].

Although gamification may not necessarily reduce teaching workload or increase efficiency, it contributes to more effective teaching practices and enriches the overall educational experience [18]. Empirical findings also indicate that gamified learning environments can improve students' understanding of complex concepts, increase active participation and create more stimulating and interactive classroom settings [46].

More advanced implementations, such as Gamified Flipped Learning (GFL), further demonstrate the potential of gamification to support the development of higher-order skills, suggesting a positive impact on students' innovation skills and overall perception of the learning process [47]. Urh et al. [42] pointed out that the introduction of gamification in e-learning leads to a higher level of user satisfaction and a more positive perception of the learning process. One of the recently conducted meta-analysis confirms

a statistically significant and positive overall effects of gamification on learning outcomes but further highlight the importance of applying context-specific and theory-driven gamification strategies tailored to students' needs, which possess greater potential to stimulate positive experiences [48]. Similarly, another meta-analysis points out the higher potential of gamification to increase students' enjoyment compared to traditional lecture-based teaching methods, although its effect on academic performances appeared to be moderate [49]. On the basis of these findings, it can be concluded that gamification contributes to the improvement of the overall student experience by creating an interactive, motivating and personalized learning environment that leads to a positive perception of the educational process.

Although literature predominantly points to the positive effects of gamification in higher education, it is necessary also to look at its limitations and potential negative consequences. Namely, some research shows that the effects of gamification are not universal and depend on students' characteristics, type of students' motivation, design of gamification elements, teaching context etc. [50, 51]. Overreliance on extrinsic rewards can lead to the dominance of extrinsic motivation and a decrease in intrinsic interest in learning, particularly if students begin to perceive learning activities solely in the function of obtaining rewards [52, 53, 54]. In addition, authors point out that gamification may lead to the surface-level engagement which does not guarantee better learning outcomes [55]. Also, some findings indicate that students may experience both positive and negative emotions when exposed to gamified learning context, meaning that their experiences might significantly differ [56]. Therefore, future research should pay more attention to identifying contextual factors that determine the success of gamification, including individual characteristics of students, cultural context, study discipline and degree of personalization of gamification systems.

4 Conclusion

Based on the review of the relevant literature, it can be concluded that gamification represents an innovative approach to improving the educational process in higher education, especially in the domain of increasing motivation, engagement and the overall experience of students. By integrating game elements into teaching activities, it is possible to create a more dynamic and interactive learning environment. Observed from the perspective of the value proposition, gamification contributes to the perception of the quality of the educational service, which further emphasizes its importance in modern educational systems.

In terms of recommendations for future research, it is necessary to further empirically test the proposed conceptual relationships between motivation, engagement and student experience, taking into account different contexts, disciplines and types of gamification elements. It is especially important to investigate the long-term effects of gamification, as well as the differences between intrinsic and extrinsic motivational mechanisms. Also, the application of more complex methodological approaches, such as

longitudinal studies and experimental designs, is recommended in order to better understand cause-effect relationships.

From a practical aspect, higher education institutions should approach the implementation of gamification strategically, with careful design of elements that encourage not only external motivation, but also deeper, intrinsic involvement of students. Special focus should be placed on the personalization of teaching activities, the balance between challenges and rewards, as well as the provision of timely and quality feedback. In this way, gamification can become an effective tool for improving the quality of education and creating a positive student experience.

Disclosure of Interests. Authors declare no competing interests.

Acknowledgement:

During the preparation of this work the authors used Open AI Chat GPT 4o in order to improve language and readability.

References

1. Alsubhi, M.A., Sahari, N., Wook, T.S.M.T.: A conceptual engagement framework for gamified E-Learning platform activities. International Journal of Emerging Technologies in Learning (iJET). 15(22), 4-23 (2020). https://doi.org/10.3991/ijet.v15i22.15443.
2. Huang, R., Ritzhaupt, A.D., Sommer, M., Zhu, J., Stephen, A., Valle, N., Hampton, J., Li, J.: The impact of gamification in educational settings on student learning outcomes: a meta-analysis. Educational Technology Research and Development. 68, 1875–1901 (2020). https://doi.org/10.1007/s11423-020-09807-z.
3. Alonso-Sánchez, J.A., Alonso, J.L.N., Santana-Monagas, E.: Gamification in Higher Education: A case study in Educational Sciences. TechTrends. 69, 507–518 (2025). https://doi.org/10.1007/s11528-025-01056-2.
4. Chung, C.-H., Shen, C., Qiu, Y.-Z.: Students' acceptance of gamification in higher education. Int. J. Game-Based Learn. 9(2), 1–19 (2019). https://doi.org/10.4018/IJGBL.2019040101
5. Sanchez, D.R., Langer, M., Kaur, R.: Gamification in the classroom: Examining the impact of gamified quizzes on student learning. Computers & Education. 144, 103666 (2019). https://doi.org/10.1016/j.compedu.2019.103666.
6. Ortiz, M., Chiluiza, K., Valcke, M.: Gamification in higher education and STEM: a systematic review of literature. In: EDULEARN16 Proceedings, Barcelona, Spain, pp. 6548–6558 (2016)
7. Varannai, I., Sasvari, P., Urbanovics, A.: The use of gamification in higher education: an empirical study. Int. J. Adv. Comput. Sci. Appl. 8(10), 1-6 (2017). https://doi.org/10.14569/IJACSA.2017.081001
8. Štavljanin, V., Cvijović, J., Kostić Stanković, M.: Research of indirect advertising in video game industry. Industrija 45(4), 113–132 (2017). https://doi.org/10.5937/industrija45-16065
9. Metwally, A.H.S., Nacke, L.E., Chang, M., Wang, Y., Yousef, A.M.F.: Revealing the hotspots of educational gamification: An umbrella review. International Journal of Educational Research. 109, 101832 (2021). https://doi.org/10.1016/j.ijer.2021.101832.

10. Wiggins, B.E.: An overview and study on the use of games, simulations, and gamification in higher education. International Journal of Game-Based Learning. 6, 18–29 (2016). https://doi.org/10.4018/ijgbl.2016010102.
11. Deterding, S., Dixon, D., Khaled, R., Nacke, L.: From game design elements to gamefulness: defining gamification. In: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, pp. 9–15. ACM (2011). https://doi.org/10.1145/2181037.2181040
12. Alsawaier, R.S.: The effect of gamification on motivation and engagement. Int. J. Inf. Learn. Technol. 35(1), 56–79 (2018). https://doi.org/10.1108/IJILT-02-2017-0009
13. Pratama, F.A., Silitonga, R.M., Jou, Y.-T.: Rimigs: the impact of gamification on students' motivation and performance in programming class. Indones. J. Electr. Eng. Comput. Sci. 24(3), 1789–1795 (2021). https://doi.org/10.11591/ijeecs.v24.i3.pp1789-1795
14. Alsadaji, A.J., Al-Fayyadh, S., Al-Saadi, S.F., Mohammed, Z.J.: Gamification in higher education: an overview. Med. Educ. Bull. 3(4), 555–562 (2022). https://doi.org/10.22034/meb.2022.348675.1062
15. An, Y.: The impact of gamification on doctoral students' perceptions, emotions, and learning in an online environment. TechTrends 67, 706–717 (2023). https://doi.org/10.1007/s11528-022-00833-7
16. Zeng, J., Sun, D., Looi, C.-K., Fan, A.C.W.: Exploring the impact of gamification on students' academic performance: a comprehensive meta-analysis. Br. J. Educ. Technol. 55, 2478–2502 (2024). https://doi.org/10.1111/bjet.13471
17. Lampropoulos, G., Sidiropoulos, A.: Impact of gamification on students' learning outcomes and academic performance: a longitudinal study. Educ. Sci. 14(4), 367 (2024). https://doi.org/10.3390/educsci14040367
18. Langendahl, P.-A., Cook, M., Mark-Herbert, C.: Gamification in higher education: toward a pedagogy to engage and motivate. Working Paper, Uppsala: Swedish University of Agricultural Sciences (2016). https://doi.org/10.13140/RG.2.1.4594.9042
19. Khaldi, A., Bouzidi, R., Nader, F.: Gamification of e-learning in higher education: a systematic literature review. Smart Learn. Environ. 10, 10 (2023). https://doi.org/10.1186/s40561-023-00227-z
20. Urh, M., Vukovic, G., Jereb, E., Pintar, R.: The model for introduction of gamification into e-learning in higher education. Procedia Soc. Behav. Sci. 197, 388–397 (2015). https://doi.org/10.1016/j.sbspro.2015.07.154
21. Deif, A.: Insights on lean gamification for higher education. Int. J. Lean Six Sigma 8(3), 359–376 (2017). https://doi.org/10.1108/IJLSS-04-2016-0017
22. Mitchell, R., Schuster, L., Jin, H.S.: Gamification and the impact of extrinsic motivation on needs satisfaction: making work fun? J. Bus. Res. 106, 323–330 (2020). https://doi.org/10.1016/j.jbusres.2018.11.022
23. Mohanty, S., Christopher, B.P.: A study on role of gamification elements in training outcomes: comparing the mediating effect of intrinsic and extrinsic motivation. Learn. Organ. 30(4), 480–500 (2023). https://doi.org/10.1108/TLO-08-2022-0098
24. Aşıksoy, G.: The effects of the gamified flipped classroom environment on students' motivation and learning achievements. Qual. Quant. 52, 129–145 (2018). https://doi.org/10.1007/s11135-017-0597-1
25. Zourmpakis, A.-I., Kalogiannakis, M., Papadakis, S.: Adaptive gamification in science education: impact on students' motivation. Computers 12(7), 143 (2023). https://doi.org/10.3390/computers12070143

26. Lister, M.: Gamification: the effect on student motivation and performance at the post-secondary level. Issues Trends Educ. Technol. 3(2), (2015). https://doi.org/10.2458/azu_itet_v3i2_lister
27. García-López, I.M., Acosta-Gonzaga, E., Ruiz-Ledesma, E.F.: Investigating the impact of gamification on student motivation, engagement, and performance. Educ. Sci. 13(8), 813 (2023). https://doi.org/10.3390/educsci13080813
28. Buckley, P., Doyle, E.: Gamification and student motivation. Interact. Learn. Environ. 24(6), 1162–1175 (2016). https://doi.org/10.1080/10494820.2014.964263
29. Jaramillo-Mediavilla, L., Basantes-Andrade, A., Cabezas-González, M., Casillas-Martín, S.: Impact of gamification on motivation and academic performance: a systematic review. Educ. Sci. 14(6), 639 (2024). https://doi.org/10.3390/educsci14060639
30. Kiselicki, M., Josimovski, S.: Implementation of gamification in HEIs in the Republic of Macedonia. Econ. Anal. 51(3–4), 57–66 (2018). https://doi.org/10.28934/ea.18.51.34.pp57-66
31. Grubor, M., Stojanović, M., Mitić, P., Veličković, A.: The impact of gamification on the motivation of online students. In: 16th International Conference on eLearning (ELEARNING2025) Proceedings, Belgrade, Serbia, pp. 253-265 (2025). https://doi.org/https://doi.org/10.46793/eLearning2025.253G
32. Huseinović L.: The Effects of gamification on student motivation and achievement in learning English as a foreign language in higher education. MAP Education and Humanities. 4, 10-36 (2023). https://doi.org/https://https://doi.org/10.53880/2744-2373.2023.4.10
33. Ruiz, J.J.R., Sanchez, A.D.V., Figueredo, O.R.B.: Impact of gamification on school engagement: a systematic review. Front. Educ. 9, 1466926 (2024). https://doi.org/10.3389/feduc.2024.1466926
34. Adams, S.P., Du Preez, R.: Supporting student engagement through gamification of learning activities. Technol. Knowl. Learn. 27, 119–138 (2022). https://doi.org/10.1007/s10758-021-09500-x
35. Murillo-Zamorano, L.R., López-Sánchez, J.Á., López-Rey, M.J., Bueno-Muñoz, C.: Gamification in higher education: the ECOn+ star battles. Comput. Educ. 194, 104699 (2023). https://doi.org/10.1016/j.compedu.2022.104699
36. Fortuna, J.M., de la Fuente, G., Velasco, P.: Does gamification mediate the relationship between digital social capital and student performance? Int. J. Manag. Educ. 21(3), 100846 (2023). https://doi.org/10.1016/j.ijme.2023.100846
37. Maryana, M., Halim, C., Rahmi, H.: The impact of gamification on student engagement and learning outcomes in mathematics education. Int. J. Bus. Law Educ. 5(2), 1697–1708 (2024). https://doi.org/10.56442/ijble.v5i2.682
38. Smiderle, R., Rigo, S.J., Marques, L.B., Coelho, J.A.P.M., Jaques, P.A.: The impact of gamification on students' learning, engagement and behavior. Smart Learn. Environ. 7, 3 (2020). https://doi.org/10.1186/s40561-019-0098-x
39. Meng, C., Zhao, M., Pan, Z., Pan, Q., Bonk, C.J.: Investigating the impact of gamification components on online learners' engagement. Smart Learn. Environ. 11, 47 (2024). https://doi.org/10.1186/s40561-024-00336-3
40. Moliner-Heredia, R., Abellán-Nebot, J.V.: Evaluation of the impact of gamification on students' performance and engagement. Int. J. Mech. Eng. Educ. 51(3), 174–193 (2023). https://doi.org/10.1177/03064190231160357
41. Poondej, C., Lerdpornkulrat, T.: Gamification in e-learning: a Moodle implementation and its effect on student engagement and performance. Interact. Technol. Smart Educ. 17(1), 56–66 (2020). https://doi.org/10.1108/ITSE-06-2019-0030

42. Urh, M., Vukovic, G., Jereb, E., Pintar, R.: The model for introduction of gamification into e-learning in higher education. Procedia Soc. Behav. Sci. 197, 388–397 (2015). https://doi.org/10.1016/j.sbspro.2015.07.154
43. Milenković, I., Šošević, U., Simić, D., Minović, M., Milovanović, M.: Improving student engagement in a biometric classroom: the contribution of gamification. Univ Access Inf Soc 18, 523–532 (2019). https://doi.org/10.1007/s10209-019-00676-9
44. Rivera, E.S., Garden, C.L.P.: Gamification for student engagement: a framework. J. Further High. Educ. 45(7), 999–1012 (2021). https://doi.org/10.1080/0309877X.2021.1875201
45. Oliveira, W., Hamari, J., Joaquim, S., Toda, A.M., Palomino, P.T., Vassileva, J., Isotani, S.: The effects of personalized gamification on students' flow experience, motivation, and enjoyment. Smart Learn. Environ. 9, 16 (2022). https://doi.org/10.1186/s40561-022-00194-x
46. Campillo-Ferrer, J.-M., Miralles-Martínez, P., Sánchez-Ibáñez, R.: Gamification in higher education: impact on student motivation and competencies. Sustainability 12(12), 4822 (2020). https://doi.org/10.3390/su12124822
47. Ahmed, H.D., Asiksoy, G.: The effects of gamified flipped learning on students' innovation skills and self-efficacy. Sustainability 13(18), 10163 (2021). https://doi.org/10.3390/su131810163
48. Gyedu, F.O., Partey, P.A., Boafo, F.A., Agbo, D.D.: Gamification Effect: Enhancing Student Learning Outcomes in Higher Education: A Meta-Analysis and Policy Implications. SAGE Open. 16, (2026). https://doi.org/10.1177/21582440261421375.
49. Zeng, J., Sun, D., Looi, C., Fan, A.C.W.: Exploring the impact of gamification on students' academic performance: A comprehensive meta-analysis of studies from the year 2008 to 2023. British Journal of Educational Technology. 55, 2478–2502 (2024). https://doi.org/10.1111/bjet.13471.
50. Buckley, P., Doyle, E.: Gamification and student motivation. Interact. Learn. Environ. 24(6), 1162–1175 (2016). https://doi.org/10.1080/10494820.2014.964263
51. Rivera, E.S., Garden, C.L.P.: Gamification for student engagement: a framework. J. Further High. Educ. 45(7), 999–1012 (2021). https://doi.org/10.1080/0309877X.2021.1875201
52. Mitchell, R., Schuster, L., Jin, H.S.: Gamification and the impact of extrinsic motivation on needs satisfaction: making work fun? J. Bus. Res. 106, 323–330 (2020). https://doi.org/10.1016/j.jbusres.2018.11.022
53. Mohanty, S., Christopher, B.P.: A study on role of gamification elements in training outcomes: comparing the mediating effect of intrinsic and extrinsic motivation. Learn. Organ. 30(4), 480–500 (2023). https://doi.org/10.1108/TLO-08-2022-0098
54. Fuchs, K.: Challenges with gamification in higher education: A narrative review with implications for educators and policymakers. Int. J. Changes Educ. 1(1), 51–56 (2024). https://doi.org/10.47852/bonviewIJCE32021604
55. Jose, B., Cherian, J., Jaya, P.J., Kuriakose, L., Leema, PWR.: The ghost effect: how gamification can hinder genuine learning. Front. Educ. 9, 1474733 (2024). https://doi.org/10.3389/feduc.2024.1474733
56. An, Y.: The impact of gamification on doctoral students' perceptions, emotions, and learning in an online environment. TechTrends 67, 706–717 (2023). https://doi.org/10.1007/s11528-022-00833-7

Original Research Article *https://doi.org/10.5281/zenodo.23472230*

Digital Communication Strategies of IVF Clinics in Informing About Gamete Donation: A Comparative Analysis of International and Serbian Websites

Biljana Djordjevic[1][0009-0000-6223-5102]

[1] University of Belgrade, Faculty of Organizational Sciences, Belgrade, Republic of Serbia
biljana.dimitrijevic@yahoo.com

Abstract. Gamete donation represents a particularly sensitive area within reproductive healthcare, as it is shaped by complex ethical, legal, and emotional considerations. In contemporary healthcare settings, IVF clinic websites often serve as the first and most accessible point of contact for individuals seeking information about reproductive treatment options. Consequently, the way donation-related information is presented online plays an important role in shaping users' understanding, expectations, and decision-making. This study investigates how IVF clinics communicate information about gamete donation on their websites through a comparative analysis of selected international clinics and clinics in Serbia. The research is based on a structured content analysis, with particular attention given to the visibility, accessibility, and clarity of information. The analysis also examines how clinics communicate procedural details, legal frameworks, and potential risks, as well as the degree of transparency and user orientation in their digital communication. The findings reveal clear cross-country differences, particularly regarding information availability, transparency practices, and donor-oriented communication strategies. The study contributes to the field of digital health communication by emphasising the importance of clear, structured, and accessible online information in sensitive areas of reproductive medicine.

Keywords: Digital communication, Gamete donation, IVF.

1 Introduction

Gamete donation is a key component of assisted reproductive technologies; however, its availability and implementation vary significantly across countries. These differences are primarily shaped by regulatory frameworks, ethical considerations, and broader healthcare system structures [1]. Alongside regulatory variation, a lack of clear, structured, and easily accessible information remains a persistent challenge for both potential donors and recipients.

Within digital healthcare, IVF clinic websites serve as a primary interface between healthcare providers and patients. For many individuals, these websites are the first point of contact when exploring treatment options. Beyond their informational role, they also function as communication platforms that shape users' perceptions of medical

procedures, risks, and expected outcomes. Clinics typically provide information on eligibility criteria, treatment processes, and legal conditions. However, the level of detail, clarity, and transparency varies considerably across contexts [2].

From a theoretical standpoint, this study is grounded in key concepts from digital health communication, particularly patient-centred communication, information asymmetry, and transparency. Incomplete or poorly structured information can heighten uncertainty and reduce users' ability to make informed decisions, especially in ethically and emotionally sensitive areas such as reproductive medicine.

Although previous research has examined the general transparency and marketing practices of IVF clinic websites, less attention has been devoted specifically to gamete donation communication. This gap is particularly evident in Europe, where diverse regulatory and clinical environments lead to heterogeneous communication practices. Prior studies also highlight deficiencies in communicating the medical, psychological, and ethical dimensions of donation [3].

Accordingly, this study analyses digital communication practices through a comparative content analysis of IVF clinic websites, focusing on:

- availability and clarity of information
- transparency (e.g., success rates)
- communication of risks and legal aspects
- use of emotional and persuasive communication elements
- targeting of donors and recipients
- user engagement mechanisms

1.1 Research Question

- How is gamete donation information presented on IVF clinic websites?
- What differences exist between Serbian and international clinics?
- To what extent do these websites demonstrate transparent and user-oriented communication practices?

2 Methodology

The study included 54 IVF clinics across six countries: Serbia, the Czech Republic, Greece, Denmark, Sweden, and Italy, with nine clinics selected from each country. The selection of countries was designed to cover those with different regulatory frameworks, ethical approaches and market structures in reproductive medicine in the European Union. The clinics were selected in the order they were listed in the Google search engine after typing the name of the country and the IVF clinic.

The coding procedure was performed in an Excel spreadsheet where the search categories (Variables) were defined. And for each piece of content on the website, the author added to the spreadsheet whether the category (Variables) searched for was on the website or not. These predefined categories are used across all sites to provide consistency and systematicity of the content.

The unit of analysis was the official website of each clinic. Only publicly available content related to gamete donation was included in the analysis. A structured coding

framework was developed to ensure systematic evaluation of key dimensions, including information availability, transparency indicators, communication strategies, and user engagement features.

Data were analysed using descriptive statistics (frequencies and percentages). Fisher's exact test was applied to assess differences between countries, with statistical significance set at $p < 0.05$.

3 Results

3.1 General Findings

The analysis shows that most IVF clinic websites include basic information about gamete donation. Table 1 presents transparency information on success rates, which were available at 46.3% of sites. Then, elements of emotional marketing were moderately represented, with 48.1% of sites using emotional images, while storytelling was present on 25.9% of sites. Persuasive promotional language was rarely used (7.4%). Analysis of target groups revealed that only 14.8% of websites have a special section for donors, while 40.7% had content intended for recipients. A call to action for donors (Donor CTA) was present on 18.5% of sites, while a contact form existed on most sites (85.2%). In terms of visual and functional features, 38.9% of sites were rated as average functional, while the same percentage had a modern design. In the majority of cases, navigation was easy or medium easy (85.2%), which indicates a relatively good accessibility of the content.

Table 1. Descriptive statistics of website characteristics

Variable	Category	n	%
Country	Serbia	9	16.7
	Czech Republic	9	16.7
	Greece	9	16.7
	Denmark	9	16.7
	Sweden	9	16.7
	Italy	9	16.7
Donation explained	No	16	29.6
	Yes	38	70.4
Donor criteria	No	22	40.7
	Yes	32	59.3
Risk mentioned	No	42	77.8
	Yes	12	22.2
Legal information provided	No	35	64.8
	Yes	19	35.2

Variable	Category	n	%
Transparency (success rate shown)	No	29	53.7
	Yes	25	46.3
Emotional images	No	28	51.9
	Yes	26	48.1
Storytelling	No	40	74.1
	Yes	14	25.9
Persuasive language	No	50	92.6
	Yes	4	7.4
The donor section exists	No	46	85.2
	Yes	8	14.8
Recipient section exists	No	32	59.3
	Yes	22	40.7
Donor call to action	No	44	81.5
	Yes	10	18.5
Contact form	No	8	14.8
	Yes	46	85.2
Website design	Outdated design	12	22.2
	Average, functional but not modern	21	38.9
	Modern design	21	38.9
Website navigation	Hard	8	14.8
	Average	19	35.2
	Easy	27	50.0

While 70.4% of websites provide a general explanation of donation, more detailed information is less frequently available. Donor criteria are presented on 59.3% of websites, while risks are mentioned on only 22.2%. Legal information is available on 35.2% of sites.

Transparency, measured through the presence of success rates, is observed on 46.3% of websites. Donor-oriented communication is limited, with only 14.8% of websites including dedicated donor sections.

3.2 Comparative Analysis

The results of the statistical analysis showed statistically significant differences between countries across several dimensions (Table 2).

Table 2. Comparison of website characteristics by country (Fisher's exact test)

Variable	Category	Serbia n (%)	Czech Rep. n (%)	Greece n (%)	Denmark n (%)	Sweden n (%)	Italy n (%)	Fisher (p)
Donation explained	No	7 (13.0)	0 (0.0)	2 (3.7)	1 (1.9)	5 (9.3)	1 (1.9)	18.02 (.001)
	Yes	2 (3.7)	9 (16.7)	7 (13.0)	8 (14.8)	4 (7.4)	8 (14.8)	
Donor criteria	No	9 (16.7)	1 (1.9)	2 (3.7)	1 (1.9)	7 (13.0)	2 (3.7)	27.01 (<.001)
	Yes	0 (0.0)	8 (14.8)	7 (13.0)	8 (14.8)	2 (3.7)	7 (13.0)	
Risk mentioned	No	8 (14.8)	8 (14.8)	7 (13.0)	6 (11.1)	5 (9.3)	8 (14.8)	4.63 (.529)
	Yes	1 (1.9)	1 (1.9)	2 (3.7)	3 (5.6)	4 (7.4)	1 (1.9)	
Legal information	No	7 (13.0)	6 (11.1)	5 (9.3)	5 (9.3)	6 (11.1)	6 (11.1)	1.60 (.985)
	Yes	2 (3.7)	3 (5.6)	4 (7.4)	4 (7.4)	3 (5.6)	3 (5.6)	
Transparency (success rate)	No	9 (16.7)	5 (9.3)	3 (5.6)	2 (3.7)	6 (11.1)	4 (7.4)	14.46 (.011)
	Yes	0 (0.0)	4 (7.4)	6 (11.1)	7 (13.0)	3 (5.6)	5 (9.3)	
Emotional images	No	5 (9.3)	3 (5.6)	5 (9.3)	3 (5.6)	5 (9.3)	7 (13.0)	4.98 (.475)
	Yes	4 (7.4)	6 (11.1)	4 (7.4)	6 (11.1)	4 (7.4)	2 (3.7)	
Storytelling	No	8 (14.8)	6 (11.1)	4 (7.4)	7 (13.0)	7 (13.0)	8 (14.8)	5.83 (.336)
	Yes	1 (1.9)	3 (5.6)	5 (9.3)	2 (3.7)	2 (3.7)	1 (1.9)	
Persuasive language	No	9 (16.7)	7 (13.0)	8 (14.8)	8 (14.8)	9 (16.7)	9 (16.7)	4.39 (.690)
	Yes	0 (0.0)	2 (3.7)	1 (1.9)	1 (1.9)	0 (0.0)	0 (0.0)	
The donor section exists	No	9 (16.7)	9 (16.7)	9 (16.7)	5 (9.3)	5 (9.3)	9 (16.7)	14.15 (.001)
	Yes	0 (0.0)	0 (0.0)	0 (0.0)	4 (7.4)	4 (7.4)	0 (0.0)	

Variable	Category	Serbia n (%)	Czech Rep. n (%)	Greece n (%)	Denmark n (%)	Sweden n (%)	Italy n (%)	Fisher (p)
Recipient section exists	No	8 (14.8)	8 (14.8)	8 (14.8)	3 (5.6)	3 (5.6)	2 (3.7)	19.23 (.001)
	Yes	1 (1.9)	1 (1.9)	1 (1.9)	6 (11.1)	6 (11.1)	7 (13.0)	
Donor call to action	No	9 (16.7)	9 (16.7)	7 (13.0)	5 (9.3)	5 (9.3)	9 (16.7)	12.36 (.008)
	Yes	0 (0.0)	0 (0.0)	2 (3.7)	4 (7.4)	4 (7.4)	0 (0.0)	
Contact form	No	4 (7.4)	0 (0.0)	0 (0.0)	0 (0.0)	3 (5.6)	1 (1.9)	10.56 (.023)
	Yes	5 (9.3)	9 (16.7)	9 (16.7)	9 (16.7)	6 (11.1)	8 (14.8)	
Website design	Outdated	5 (9.3)	1 (1.9)	1 (1.9)	1 (1.9)	2 (3.7)	2 (3.7)	10.39 (.401)
	Average	3 (5.6)	2 (3.7)	5 (9.3)	3 (5.6)	4 (7.4)	4 (7.4)	
	Modern	1 (1.9)	6 (11.1)	3 (5.6)	5 (9.3)	3 (5.6)	3 (5.6)	
Website navigation	Hard	1 (1.9)	1 (1.9)	1 (1.9)	1 (1.9)	2 (3.7)	2 (3.7)	10.09 (.387)
	Average	7 (13.0)	2 (3.7)	3 (5.6)	3 (5.6)	2 (3.7)	2 (3.7)	
	Easy	1 (1.9)	6 (11.1)	5 (9.3)	5 (9.3)	5 (9.3)	5 (9.3)	

The analysis reveals statistically significant differences between countries in several dimensions, including donation explanation, donor criteria, and transparency. There is a statistically significant correlation between countries and the content of the explanation about reproductive cell donations on the websites ($p<0.01$). Such information is significantly less present on the websites of clinics from Serbia and Sweden compared to other countries. Interestingly, clinics in Sweden more often had a call to action directed at donors, which may indicate a more organized and transparent approach to donor recruitment, and not necessarily a more pronounced marketing approach. Significant differences were also observed in relation to the presence of donor criteria ($p<0.001$); this data was significantly less available on the websites of clinics from Serbia and Sweden. There is a uniformly low level of information about risks and regarding the availability of legal information among all analyzed countries, and no statistically significant association was observed between the country and the presence of information about risks ($p=0.529$), nor regarding the availability of legal information ($p=0.985$). Information on transparency when displaying success rates shows significant differences ($p=0.011$); this information was less often present on websites from

Serbia, the Czech Republic, and Sweden. In the field of emotional marketing (images, storytelling, and promotional language), no statistically significant difference was found between countries ($p>0.05$), which indicates a relatively uniform approach in this segment. Also, a statistically significant difference was observed in the presence of a call to action for donors ($p=0.008$), which was less common on websites from Serbia, the Czech Republic, Greece, and Italy. The presence of a contact form also showed a statistically significant difference between countries ($p=0.023$), with contact forms being present on clinic websites from Serbia and Sweden.

Finally, no statistically significant association was found between country and website design ($p=0.401$), nor between country and ease of navigation ($p=0.387$), indicating a relatively uniform level of technical and visual quality among the analysed websites.

4 Discussion

The results of this research show differences in the way IVF clinics communicate with reproductive cell donation target groups through their websites. ASRM guidelines emphasise the importance of transparency in communication [4], and the results of our research indicate a limited presence of transparent information on the analysed websites, which is confirmed by previous studies that analysed the quality of information on IVF clinic websites, where insufficient compliance with recommended communication standards was also observed [2]. In particular, the difference can be seen in terms of information availability, transparency, and access to potential donors and recipients. Although most clinics provide basic information about donation, more detailed aspects, such as donor criteria, procedural risks, and the legal framework, remain underrepresented, which may limit user awareness.

Fertility clinic websites are not only sources of information but also key platforms for shaping the narrative of donation, which may vary across cultural and regulatory contexts [5]. The findings from this research confirm that IVF clinic websites are not only a source of information, but also participate in understanding reproductive cell donations, and thus it can be seen that different national contexts lead to different communication narratives (UK-responsibility, Belgium-medical model, Spain-sisterhood/solidarity).

Compared to international clinics, clinics in Serbia show a lower level of structure and availability of information, as well as a lower representation of content directed towards donors. In particular, the limited use of calls to action and the lack of clearly defined sections for donors were observed, which indicates the underutilised potential of digital platforms for communication and engagement of this target group. Such findings may be related to the regulatory and organisational specificities of a system in which reproductive cell donation relies heavily on imports, which reduces the need for active local donor recruitment.

Clinics in countries known as destinations for cross-border reproductive medicine, such as Greece and the Czech Republic, are more likely to use digital marketing elements, including more clearly structured content and greater availability of information, which is in line with their market orientation and international competition. The results showed that clinics in Sweden were more likely to include calls to action directed at donors, despite the expectation that systems with a stronger ethical framework would

have less digital marketing. This finding may indicate that the presence of a call to action is not necessarily associated with a commercial approach. Based on the results, we see the existence of different models of digital communication in the field of reproductive cell donation: these differences are particularly evident in the way of communicating with potential donors, as well as in the degree of transparency and availability of information.

Finally, the design and functionality of websites do not show significant differences between countries; the results clearly indicate that the key differences lie in the content and the way of communication. Digital communication strategies need to be improved, especially in the context of providing clear, structured, and transparent information about reproductive cell donation. The reason for this is that IVF clinic digital communication can be seen not only as a means of information, but as a key social marketing instrument that actively participates in shaping perceptions, social norms, and potential behaviour regarding reproductive cell donation.

5 Conclusion

The results of this research indicate the existence of significant differences in the way reproductive cell donation is communicated on the websites of IVF clinics in different European countries. Most of the analysed clinics provide basic information about the donation, and shortcomings appear in terms of detail, transparency, and structured content, especially when it comes to the criteria for donors, risks, and legal aspects of the procedure. Compared to international clinics, clinics in Serbia show a lower level of digital communication development in this area, with a limited representation of donor-oriented content and a weaker use of interactive elements, such as calls to action. Such findings point to the underutilised potential of websites as a key channel for information and communication with target groups.

On the other hand, the results suggest the existence of different approaches in the international context, from market-oriented models characteristic of countries with developed IVF tourism, to systematically organised approaches with an emphasised ethical framework. It is particularly important that the presence of elements such as calls to action does not necessarily reflect a commercial approach, but can be part of a transparent and structured donor recruitment system.

It is necessary to improve digital communication in the field of reproductive cell donation, with a special focus on providing clear, accessible, and reliable information for potential donors and recipients. Further research can contribute to a deeper understanding of how digital communication affects user awareness and engagement in this area.

References

1. Calhaz-Jorge, C., Smeenk, J., Wyns, C., De Neubourg, D., Baldani, D.P., Bergh, C., Cuevas-Saiz, I., De Geyter, C., Kupka, M.S., Rezabek, K.: Survey on ART and IUI: legislation, regulation, funding and registries in European countries- an update. Human Reproduction 39(9), 1909-1924 (2024). https://doi.org/10.1093/humrep/deae163

2. Carneiro, M.M., Koga, C.N., Mussi, M.C.L., Fradico, P.F., Ferreira, M.C.F.: Quality of information provided by Brazilian fertility clinic websites: Compliance with Brazilian Medical Council (CFM) and American Society for Reproductive Medicine (ASRM) guidelines. JBRA Assisted Reproduction 27(2), 169-173 (2023). https://pmc.ncbi.nlm.nih.gov/articles/PMC10279427/
3. Iranifar, E., Ebrahimzadeh Zagami, S., Amirian, M., Ebrahimpour, H., Latifnejad Roudsari, R.: A systematic review of assisted and third-party reproduction guidelines regarding management and care of donors. BMC Medical Ethics 24, 84 (2023). https://doi.org/10.1186/s12978-024-01804-2
4. American Society for Reproductive Medicine: Guidelines for advertising by ART programs: Fertility and Sterility 82 (Suppl.1), 527-528 (2004). https://www.fertstert.org/article/S0015-0282(04)00911-2/fulltext
5. Coveney, C., Hudson, N., Lafuente-Funes,., Jacxsens, L., Provoost, V.: From scarcity to sisterhood: The Framing of egg donation on fertility clinic websites in the UK, Belgium, and Spain. Social Science & Medicine 277, 296 (2021). https://www.sciencedirect.com/science/article/pii/S0277953622000910?via%3Dihub#cited-by

There is no external funding for research and no conflict of interest.

The authors verified the scientific content of the manuscript, and Artificial Intelligence tools were used to improve the text and proofread the language.

Original Research Article *https://doi.org/10.5281/zenodo.23472271*

Gender Differences in Consumers' Perception of Environmental and Social CSR Activities

Anja Vučić[1] [0009-0007-6205-3369]

[1] Faculty of Organizational Sciences, University of Belgrade
Email: anjaavucic@gmail.com

Abstract. This paper examines gender differences in consumers' perception of corporate social responsibility activities, focusing on the environmental and social dimensions of Corporate Social Responsibility (CSR). The study is based on primary data collected through an online survey conducted among 225 respondents in Serbia. Respondents evaluated selected CSR activities using a five-point Likert scale, with an additional "I don't know" option, which was included in the descriptive analysis as a separate category, while mean values were calculated only for valid Likert scale responses. The data were analyzed using descriptive statistics and the Chi-square test of independence in the SPSS software. The results indicate that gender is statistically associated with consumers' perception of several environmental and social CSR activities. Overall, female respondents are more likely to perceive the importance of all analyzed CSR activities compared to male respondents. The findings suggest that gender may represent an important demographic factor in understanding consumer responses to CSR activities and provide practical implications for companies when designing and communicating CSR initiatives. However, these findings should be interpreted with caution because the sample was not representative, the gender distribution was unequal, and the analysis was focused on individual CSR items.

Keywords: CSR, consumers' perception, demographic characteristics, gender

1 Introduction

In today's business environment, companies are focusing more on their CSR activities due to their increasing importance. CSR activities influence the company's reputation, and therefore its business overall. Even though CSR activities are becoming more important, not all consumers perceive them in the same way. Some of the consumer's characteristics that are influencing perception of CSR are various demographic, social, cultural, and personal characteristics (Maričić, 2005). There are different studies conducted while CSR was becoming more popular topic, but not all of them gave the same results (Banerjee & McKeage, 1994; Zelezny, Chua, & Aldrich, 2000; Xiao & McCright, 2015; Calabrese, Costa & Rosati, 2016; Jones, Reilly, Cox & Cole, 2017). The consumer's behavior is frequently changing as well as the market terms, which might be the reason why previous studies have not given the common conclusion on how demographic characteristics influence the perception of CSR (Azamussan &

Shanmugam, 2022). This inconsistency in results was an indicator that further research is needed.

The aim of this paper is to examine whether gender is associated with differences in consumers' perception of selected CSR activities, specifically in the domain of environmental and social activities. The empirical analysis is primarily conducted at the item level, which means that 12 selected environmental and social CSR items, adopted from Alvarado-Herrera, Bigné, Aldás-Manzano, and Currás-Pérez (2017), are analyzed separately. The research conducted is based on survey data collected from 225 respondents in Serbia and analyzed using descriptive statistics and the Chi-square test of independence. By identifying differences between genders towards CSR perception, the study contributes to a better understanding of consumer attitudes and provides practical implications for companies when designing and communicating CSR initiatives. This paper is based on research previously presented in the author's master's thesis, and therefore part of the results presented in the paper were previously included in the corresponding thesis (Vučić, 2025).

2 Theoretical background

CSR can typically be defined through four key levels. Those are economic, legal, ethical, and philanthropic responsibilities (Carroll, 1996). As a result, CSR can be understood in different ways, depending on the perspective and context. CSR can be defined as company's effort to minimize negative effects of their business, maintain their relationships with all key stakeholders and integration of ecological, social and philanthropic aspects in their business decisions (Marco-Lajara, Zaragoza-Sáez, Falcó & Millan-Tudela, 2022). From ethical perspective, CSR is closely connected to company's ethical behavior meaning its decisions are based on moral principles, while contributing to solving environmental and social issues (Castelo Branco & Delgado, 2011). CSR can be defined as operating economically responsible and voluntarily acting socially responsible above legal minimums forward company's stakeholders. This includes integrating responsible behavior in every activity which contributes to global sustainability (Sarkar & Searcy, 2016). Hence, CSR is any business activity that in any way contributes to increasing quality of life (Cazeri, et al., 2018). During the years, CSR has changed its form, and it has matured to the concept of ESG (Environmental, Social, Governance). ESG is a strategic approach which distinguishes three main dimensions of CSR, environmental, social, and governance. The main difference between CSR and ESG is that CSR is based on voluntary initiatives, while ESG is integrated in corporate strategy, business decisions and every activity (Passas, 2024). The environmental dimension of ESG is focused on business activities that impact the environment, aiming to minimize negative effects and ensure sustainable operations in order to improve resource efficiency, and influence their positive public image (De Souza Barbosa, et al., 2023). Social dimension is based on relationships with stakeholders, ensuring that their rights and interests are respected. The company's engagement in this area improves

company performance, reputation, and long-term competitiveness (Yoo, 2025). The governance dimension is based on the principles that are guiding corporate decision-making practices. Effective governance ensures ethical behavior and enables the integration of environmental and social practices into everyday business activities (Martiny, et al., 2024). Since this paper focuses on consumers' perception of CSR activities, the empirical analysis includes the environmental and social dimensions, as these are more visible and understandable to consumers.

Perception is an individual consumer's interpretation of the environment around them. Since every individual might have a different reaction to the same stimuli and influences from the environment, this shows that stimuli can be interpreted in different ways. This interpretation depends not only on objective stimuli, but also on subjective expectations, knowledge, beliefs, attitudes, experiences, and other factors. Perception is also influenced by emotional and irrational elements, but also consumer's characteristics which influence different interpretations of the same stimuli (Husić-Mehmedović et al., 2012). In the context of CSR, this means that the same CSR activities may be evaluated differently by different consumer groups. One's perception of CSR is based on their individual characteristics and perspectives on individual level. Those perceptions might be influenced by company's efforts, but also the way that company is communicating their CSR activities. Consumers must trust the company in order to believe in their motives of doing socially responsible business (Aguinis, Rupp & Glavas, 2024).

Consumer's characteristics that have the highest influence on their perception are social, cultural and personal (Solomon, 2016). Social factors include the need of individuals' to be a part of a group which might influence their attitudes and perception. Social factors also might include demographic characteristics, including gender on which this paper focuses (Jisana, 2014; Maričić, 2015). Cultural values have a great impact on consumer perceptions, since the norms and beliefs that are influenced by culture shape individuals' behaviors and their choices (Gajdzik et al., 2024). Personal factors include the unique psychological characteristics of an individual which is shown as the totality of individual's behavior in different situations (Singh, 2016).

Social identity theory is widely used to show the influence of gender on the CSR perspective. This theory suggests that gender is the major influence that shapes value perceptions (Tajfel & Turner, 1986). Banerjee & McKeage (1994) identified that females are more conscious about environmental issues. Zelezny, Chua & Aldrich (2000) concluded that women generally report stronger environmental attitudes and behaviors than men, while Xiao & McCright (2015) emphasize that these gender differences are modest. Therefore, Xiao & McCright (2015) conclude that gender is relevant but not the dominant factor in shaping environmental perceptions. However, other researchers suggest that gender has a major role in shaping attitudes towards CSR, since female consumers are more responsive to CSR initiatives (Jones, Reilly, Cox & Cole, 2017). The research conducted in India in 2022 hasn't shown major differences in CSR's perception based on gender (Bhrambhatt, 2023), while on the other hand, research con-

ducted in South Corea in 2014 concluded that female consumers have a higher perception toward CSR compared to males (Hur, Kim & Jang, 2015). Authors Calabrese, Costa & Rosati (2016) found that female respondents expressed slightly higher expectations regarding companies' commitment to CSR efforts than male respondents. Authors Vlastelica, Jović & Pavković (2021) found that in Serbia women have higher awareness of importance of ecological and social practices. Therefore, there is more space to conduct the research in this field to identify differences in perception of CSR activities between genders.

3 Methodology

This study was conducted in order to examine the relationship between the demographic variable of gender and consumers' perceptions of the specific environmental and social activities of CSR. The quantitative research approach was applied, based on primary data, collected through an online survey. Data were gathered between October and December 2024 using a structured questionnaire distributed via email and social media platforms. The sample consisted of 225 respondents from Serbia, selected by a non-probability convenience sampling method, due to online distribution of questionnaire. Participation in the study was voluntary and anonymous.

The questionnaire included demographic questions and 12 items measuring consumers' perceived importance of selected CSR activities. Six items referred to the environmental dimension of CSR, while other six items referred to the social dimension of CSR. Those items were adopted from the scale developed by Alvarado-Herrera, Bigné, Aldás-Manzano & Currás-Pérez (2017). Consumers' attitudes toward the selected company's CSR activities were measured using a five-point Likert scale, with an additional option "I don't know". The descriptive analysis calculating mean values, standard deviation and variance excluded "I don't know answers" since they do not represent a level of agreement as rest of the five-point Likert scale. However, they were included in the descriptive analysis as a separate category, in order to show the proportion of respondents who stated that they did not have enough knowledge or awareness to evaluate a specific CSR activity. This might result in some bias in the interpretation of mean values.

The collected data were analyzed using IBM SPSS Statistics. Descriptive statistics were used to summarize the data, followed by the Chi-square test of independence to examine the relationship between respondents' attitudes toward CSR and their demographic characteristics. The chi-square test was selected because the analysis examined the association between two categorical variables, gender and response categories for each CSR item. Although the items were mostly measured using a Likert scale, the responses were treated as ordinal categorical data and not as continuous variables, since the "I don't know" answers were also included in the Chi-square test of independence. The null hypothesis was defined as: H0: "There is no statistically significant association between gender and consumers' perceived importance of the selected environmental

and social CSR activities", with a significance level set at 0.05. The hypothesis was tested through 12 item level Chi-square tests, with corresponding Cramer's V tests.

Table 1. – Demographics of sample

Demographic	Category	N	%
Gender	Male	73	32.44%
	Female	152	67.56%
	Prefer not to say	0	0%
Age group	18-24	148	65.78%
	25-34	39	17.33%
	35-49	18	8.00%
	50-65	20	8.89%
	65+	0	0%
Education level	Finished primary school	11	4.89%
	Finished high school	49	21.78%
	Graduated	110	48.89%
	Master's degree	45	20.00%
	PhD	10	4.44%
Profession	High school student	11	4.89%
	University student	100	44.44%
	Employed	94	41.78%
	Unemployed	18	8.00%
	Retired	0	0%
	Other	2	0.89%

4 Results and discussion

In order to measure the perception of CSR activities, respondents were asked to evaluate 12 questions about company's activities in environmental and social aspect of CSR. The descriptive statistics of total attitudes towards those activities is shown in Table [2].

Table 2. Descriptive statistics

Statement: It is important to me that the company:	Variable	Total N	Valid N	"I don't know" (N)	"I don't know" (%)	M	SD	S^2
Protects the environment	EKO1	225	198	27	12%	3.419	1.247	1.554
Financially supports environmental protection programs	EKO2	225	196	29	12,9%	3.245	1.305	1.704
Implements programs to reduce pollution	EKO3	225	196	29	12,9%	3.301	1.255	1.576
Recycles its waste	EKO4	225	199	26	11,6%	3.493	1.251	1.564

Uses only the natural resources necessary for its operations	EKO5	225	194	31	13,8%	3.242	1.307	1.708
Allocates resources into the development of environmentally friendly products	EKO6	225	193	32	14,2%	3.228	1.263	1.594
Financially supports programs for education	SOC1	225	195	30	13%	3.344	1.256	1.577
Financially supports programs for health	SOC2	225	188	37	16%	3.404	1.265	1.600
Financially supports programs for culture	SOC3	225	194	31	14%	3.232	1.260	1.588
Donates funds in humanitarian purposes	SOC4	225	193	32	14%	3.617	1.237	1.529
Contributes to improving the local community's quality of life	SOC5	225	195	30	13%	3.487	1.253	1.571
Operates in accordance with moral principles	SOC6	225	180	45	20%	3.933	1.190	1.415

The proportion of "I don't know" responses ranged from 11.6% to 20.0% across the analyzed items, which suggest that a notable proportion of respondents may not have enough knowledge or awareness to evaluate some CSR activities. The highest share was for the SOC6, where 20.0% of respondents chose this option.

For analyzing the influence of demographic characteristics of respondents on environmental and social variables related to the perception of corporate social responsibility, the Chi-square test of independence was conducted. This approach enabled the identification of potential relationships between demographic characteristics and respondents' attitudes toward corporate social responsibility. In order to assure data relevance, the categories analyzed were "1-5" categories on Likert scale and "I don't know". Since the expected cell counts were adequate, as no cells had expected counts below 5, there was no need to collapse the categories.

The decision to accept or reject the hypothesis is based on the value of the p-value. If the obtained p-value is less than or equal to 0.05, it indicates a higher probability that H0 is not correct, and the hypothesis is rejected. In that case, it is concluded that the variables are statistically dependent. If the p-value is greater than 0.05, H0 cannot be rejected, meaning that the variables are independent (Vuković & Bulajić, 2014).

Table 3. Results of the Chi-square test of independence

	EKO1	EKO2	EKO3	EKO4	EKO5	EKO6	SOC1	SOC2	SOC3	SOC4	SOC5	SOC6
Gender	**.010**	**.016**	.082	**.044**	**.042**	**.001**	**.023**	**.027**	**.027**	**.002**	**.001**	**.035**
Cramer's V	.258	.249	.209	.225	.226	.298	.241	.237	.241	.292	.321	.230
Interpretation	weak to moderate	weak to moderate	not significant	weak to moderate	weak to moderate	moderate	weak to moderate	weak to moderate	weak to moderate	moderate	moderate	weak to moderate

Table [3] shows that there is a statistical association between gender and several CSR activities. However, the significant associations were generally weak to moderate in strength. The highest associations were between gender and EKO6, SOC4 and SOC5. Also, results may not be fully accurate since 12 separate Chi-square tests were conducted. Therefore, the hypothesis is only partially supported.

Table [4] presents mean values by gender for the analyzed CSR items. The "I don't know" responses were excluded from the calculation of mean values because they cannot be shown as a point on the Likert scale. The aggregate EKO score was calculated as the average of the six environmental items, while the aggregate SOC score was calculated as the average of the six social items. These aggregate scores are presented only as descriptive summaries and were not used as separate hypothesis tests.

Table 4. Differences between male and female mean answers

Variable	M (Male)	M (Female)	Difference
EKO	**2.96**	**3.51**	**0.55**
EKO1	3.06	3.61	0.55
EKO2	2.85	3.45	0.60
EKO3	2.97	3.47	0.50
EKO4	3.15	3.67	0.52
EKO5	3.00	3.37	0.37
EKO6	2.76	3.48	0.72
SOC	**3.11**	**3.70**	**0.59**
SOC1	2.92	3.55	0.63
SOC2	3.00	3.61	0.61
SOC3	2.89	3.40	0.51
SOC4	3.13	3.86	0.73
SOC5	3.11	3.69	0.58
SOC6	3.67	4.11	0.44

Table [4] shows that in this study all the CSR activities are perceived as more important by female than by male respondents. However, there are also some similarities. Both groups identify SOC6 as the most important CSR activity, while in environmental aspects both groups identify EKO4 as the most important. The biggest difference in opinions can be seen in EKO6 where female respondents rated this activity 0.72 points higher than men. A similar pattern happened for SOC4, where the difference in average scores is 0.73. While the Chi-square test was used to determine the existence of statistically significant associations, mean values were compared to interpret the direction of differences based on gender.

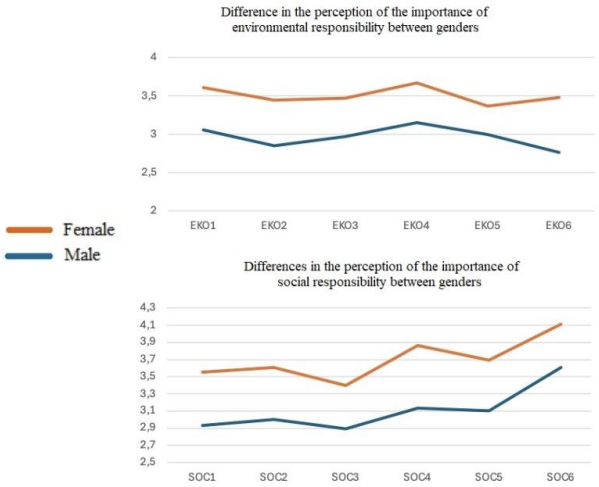


Fig. 1. Average ratings of CSR activities by gender

Based on the figure and average values, it can be concluded that the ranking of importances of activities within the same group of CSR activities is mostly similar, except for EKO6. Both men and women see the social responsibility as more important than the environmental responsibility. Overall, female respondents reported higher perceived importance of CSR activities than male respondents. The average difference in ratings for environmental activities is 0.53, while for social activities it is slightly higher, at 0.59.

These findings suggest that female respondents in this sample are more sensitive toward both environmental and social CSR activities. This is consistent with many previous studies (Zelezny, Chua, & Aldrich, 2000; Hur, Kim & Jang, 2015; Calabrese, Costa & Rosati, 2016; Vlastelica, Jović & Pavković, 2021) suggesting that women may express higher concern for environmental and social issues. For practical implication, the findings indicate that companies should consider demographic differences when designing CSR communication strategies.

However, the study has several limitations. The sample was collected through an online survey and is dominated by younger and female respondents, which limits the ability to generalize the findings. In addition, the study focused only on environmental and social CSR activities, while the governance dimension of ESG was not included in the empirical analysis. Another limitation is that the mean values were calculated only for respondents who provided Likert scale answers, while "I don't know" responses were excluded from mean calculations. This may influence the interpretation of average scores. Also, the respondents' answers are based on self-assessment, which means they may be giving answers influenced by socially desirable responses. There is also a limitation related to the fact that there were 12 separate item level tests, which means that some statistically significant results may have appeared by chance, so they should be

interpreted with caution. As the conducted research also points to the complexity of the relationship between consumer characteristics and their attitudes and perception regarding the importance of CSR, it can serve as a foundation for further studies on this topic, conducted with a larger sample.

5 Conclusion

The results of this paper indicate that gender is associated with differences in consumers' perception of CSR activities. Based on the conducted Chi-square tests, gender showed statistically significant relationships with most CSR perception variables, while the comparison of mean values showed that female respondents recognized all analyzed CSR activities as more important than male respondents. The most noticeable differences were observed in variables that are related to the development of environmentally friendly products and donations for humanitarian purposes.

However, since the associations were not statistically significant for all items and gender was not statistically associated with all analyzed items, hypothesis is only partially supported. Therefore, the findings can only indicate item level associations.

These findings suggest that gender may be an important demographic factor and has an influence on how consumers evaluate CSR activities. From a managerial perspective, companies should be aware that CSR communication may not be perceived equally by different gender groups. Female consumers in this sample were more aware of importance of both environmental and social aspects of CSR, which indicates that companies might benefit from gender sensitive CSR communication strategies. Anyhow, companies should be aware of many limitations of this paper.

References

1. Maričić, B. Ponašanje potrošača. CID, Ekonomski fakultet. Beograd. (2005)
2. Azamussan, S., & Shanmugam, M. (2022). An empirical examination of the effects of demographic groups on socially responsible consumption behaviour. Social Responsibility Journal, 18(5), 951-967.
3. Carroll A. Business and Society -Ethics and Stockholder Management, South-Western College Publishing, ITP,Ohio, (1996).
4. Marco-Lajara, B., Zaragoza-Sáez, P., Falcó, J. M., Millan-Tudela, L. A. Corporate social responsibility: A narrative literature review. Frameworks for sustainable development goals to manage economic, social, and environmental shocks and disasters, 16-3. (2022)
5. Castelo Branco, M., Delgado, C. Research on corporate social responsibility and disclosure in Portugal. Social Responsibility Journal, 7(2), 202-217. (2011)
6. Sarkar, S., Searcy, C. Zeitgeist or chameleon? A quantitative analysis of CSR definitions. Journal of Cleaner Production, 135, 1423-1435. (2016)
7. Cazeri, G. T., Anholon, R., da Silva, D., Ordoñez, R. E. C., Quelhas, O. L. G., Leal Filho, W., de Santa-Eulalia, L. A. An assessment of the integration between corporate social

responsibility practices and management systems in Brazil aiming at sustainability in enterprises. Journal of cleaner production, 182, 746-754. (2018)
8. Passas, I. The evolution of ESG: From CSR to ESG 2.0. Encyclopedia, 4(4), 1711-1720. (2024)
9. De Souza Barbosa, A., da Silva, M. C. B. C., da Silva, L. B., Morioka, S. N., de Souza, V. F. Integration of Environmental, Social, and Governance (ESG) criteria: their impacts on corporate sustainability performance. Humanities and Social Sciences Communications, 10(1), 1-18. (2023)
10. Yoo, B. C. Exploring Stakeholder and Organizational Influences on ESG Management in the Logistics Sector. Sustainability, 17(9), 4243. (2025)
11. Martiny, A., Taglialatela, J., Testa, F., Iraldo, F. Determinants of environmental social and governance (ESG) performance: A systematic literature review. Journal of Cleaner Production, 456, 142213. (2024)
12. Husić-Mehmedović, M., Kukić, S. Čičić, M. Consumer behaviour. School of Economics and Business in Sarajevo (2012)
13. Aguinis, H., Rupp, D. E., Glavas, A. Corporate social responsibility and individual behaviour. Nature Human Behaviour, 8(2), 219-227. (2024)
14. Solomon, M. R. Consumer behaviour: A European perspective. Pearson education. (2016)
15. Jisana, T. K. Consumer behaviour models: an overview. Sai Om Journal of Commerce & Management, 1(5), 34-43. (2014)
16. Gajdzik, B., Bartuś, K., Jaciow, M., Wolniak, R., Wolny, R., Grebski, W. W. Evolution of Polish e-consumers' environmental awareness and purchasing behavior over ten years. Sustainability, 16(11), 4686. (2024)
17. Singh, A. Consumer Behaviour, First edition. Himalaya Publishing House Pvt. Ltd., Mumbai. (2016)
18. Banerjee, B., McKeage, K. "How green is my value: exploring the relationship between environmentalism and materialism", in Allen C.T. and John, D.R. (Eds) Advances in Consumer Research, Association for Consumer Research, Provo, UT, Vol. 21, pp. 147 – 52. (1994)
19. Bhrambhatt, V. Personality Assessment of Green Consumers using the Big Five Model (OCEAN). International Journal of Innovative Science and Research Technology ISSN No:-2456-2165. Volume 8, Issue 11. (2023)
20. Vlastelica, T., Jović, M., Pavković, V. Preferencije i stavovi potrošača u odnosu na socijalne i ekološke dimenzije korporativne društvene odgovornosti. Marketing (0354-3471), 52(3). (2021)
21. Alvarado-Herrera, A., Bigne, E., Aldas-Manzano, J., Curras-Perez, R. A scale for measuring consumer perceptions of corporate social responsibility following the 3. sustainable development paradigm. Journal of Business Ethics, 140, 243-262. (2017).
22. Vuković, N., & Bulajić, M. Osnove statistike. Beograd: Fakultet organizacionih nauka. (2014).
23. Vučić, A. Uticaj karakteristika potrošača na percepciju značaja ekoloških i društveno odgovornih aspekata poslovanja. Unpublished Master's Thesis, Faculty of Organizational Sciences, University of Belgrade. (2025).
24. Tajfel, H., Turner, J. C. The social identity theory of intergroup behaviour. Psychology of intergroup relations, 2, 7-24. (1986)
25. Zelezny, L. C., Chua, P. P., Aldrich, C. New ways of thinking about environmentalism: Elaborating on gender differences in environmentalism. Journal of Social Issues, 56(3), 443–457. (2000).

26. Xiao, C., McCright, A. M. Gender differences in environmental concern: Revisiting the institutional trust hypothesis in the USA. Environment and Behavior, 47(1), 17–37. (2015).
27. Jones, R. J., Reilly, T. M., Cox, M. Z., Cole, B. M. Gender makes a difference: Investigating consumer purchasing behavior and attitudes toward corporate social responsibility policies. Corporate Social Responsibility and Environmental Management, 24(2), 133–144. (2017)
28. Hur, W. M., Kim, H., Jang, J. H. The role of gender differences in the impact of CSR perceptions on corporate marketing outcomes. Corporate Social Responsibility and Environmental Management, 23(6), 345-357. (2015).
29. Calabrese, A., Costa, R., Rosati, F. Gender differences in customer expectations and perceptions of corporate social responsibility. Journal of cleaner production, 116, 135-149. (2016).

Original Research Article *https://doi.org/10.5281/zenodo.23472301*

Catch It if You Can: Measuring Visual Attention of Product Placement in Movies Using Eye Tracking

Kristina Hristova[1][0009-0006-6264-9026], Nikolaos Dimitriadis[2][0009-0000-1080-6612]
Marko Ćirović[3][0000-0002-5073-3902], Matteo Venerucci[4][0000-0002-7523-4691]
and Omid Asgari[5][0000-0001-9178-1196]

[1] CITY College, University of York Europe Campus, 546 26, Thessaloniki, Greece
[2] CITY College, University of York Europe Campus, 546 26 Thessaloniki, Greece
[3] Faculty of Organizational Sciences, University of Belgrade, 11000, Belgrade, Serbia
[4] Research Scientist, Tobii AB, 182 53 Danderyd, Sweden, Customer Success Education Services
[5] Institute for International Strategy (IIS), Tokyo International University (TIU) 170-0013; Higashi Ikebukuro 4 – 42 – 31, Toshima, Tokyo

Abstract. Product placement has been a popular marketing strategy for over a century, gaining even more traction with the rise of mass media. However, its effectiveness remains uncertain due to hidden risks and significant financial investment, making it crucial to understand viewer attention and attitudes toward product placements. This study aimed to measure implicit visual attention to product placements in movies using eye tracking, while also addressing the limitations of self-reported data often used in prior research. Conducted at CITY College, University of York Europe Campus in Thessaloniki, Greece, the study involved 44 participants who viewed movie scenes containing product placements while their eye movements were recorded, followed by a questionnaire to compare implicit and explicit attention and assess brand retention. Key findings revealed that while well-known brands are more easily noticed, the success of product placement depends more on its integration into the storyline, size, and position within scenes. Recommendations for marketers include aligning brand ideology with movie themes, creatively incorporating placements into storylines, and employing neuromarketing techniques to pre-test and evaluate effectiveness. These insights and strategies offer valuable guidance for optimizing product placements in movies and enhancing their success.

Keywords: neuromarketing, eye tracking, product placement, visual attention, movies

1 Introduction

Product placement has been a popular advertising strategy since 1896, gaining immense popularity among marketers. As latest available measurements, research agency, PQ Media [1], estimated a global product placement spending of $20.57 billion in 2019, with a forecast of tremendous growth in the years to follow. Most recently, the agency

estimated that product placement spending had increased by 12.3% in 2023, reaching an impressive $29.63 billion. With rising mass media consumption, and the substantial financial investment required, understanding the effectiveness of product placement is increasingly crucial. Existing studies on the effectiveness of product placement, focus mainly on explicit findings using self-reported data, while studies that employ neuromarketing tools, like eye tracking and EEG, concentrate on disclosure of product placement, and not as much on its effectiveness [2]. Neuromarketing tools could offer valuable insights into implicit consumer perceptions and decision-making processes, helping companies optimize product placements and reduce financial risks [3]. By combining implicit data (eye tracking) with traditional explicit (self-reported questionnaires) methods, this study aims to provide an enhanced understanding of consumer attentiveness to product placement in movies [4].

Examining what level of implicit attention is paid to product placements in movies, and if that is affected by brand familiarity, product location on screen (foreground vs. background), and personal relationships with the brand? Further, do these variables affect brand placement recall?

2 Theory and Literature Review

2.1 Attention Theory

The human brain is known to have limited capabilities of processing the enormous number of sensory stimuli which surround us at all times. This is why the brain uses the cognitive efforts of attention to seed out where neural resources should best be distributed, specifically utilizing bottom-up (involuntary, stimuli driven) and top-down attention (intention driven) [5]. In relation to product placement, bottom-up attention theory could help determine its effectiveness, by suggesting viewers' attention is automatically drawn toward visually salient stimuli, such as brands with a prominent presence on screen – central, large in size, with frequent on-screen appearances and/or integration into the storyline [6–8]. In turn, the likelihood of noticeability, memory retention and future recall increases.

Looking further into the limits of attention, human beings have a limited cognitive capacity for visual and central attention; our cognitive system slows down when asked to multi-task or focus on too many individual visual stimuli simultaneously [9–11]. Taking this into consideration, this visual and central attention limit could also influence the effectiveness of product placement, since scenes with attention-grabbing visuals and auditory stimuli could easily divert attention from the product.

Furthermore, Gross et al. [12], in a study using MEG to examine the synchronization of different brain segments during visual tasks, found that visual attention capacity is regulated by long-phase synchronizations within the frontal, temporal, and parietal lobes, specifically in the beta brain-wave frequency. When overwhelmed by too much visual information and distractions, the synchronization between the regions decreases, leading to a drop in attention capacity. Stronger synchronization allowed for more ef-

fective target detection, while weaker synchronization corresponded with missed detections, demonstrating how neural coordination enables selective attention by enhancing relevant stimuli while suppressing distractions.

2.2 Product Placement Research and Findings

Throughout the years, product placement has captivated the attention of researchers, partly due to the effectiveness of the concept needed to be measured owing to the high expenses related to it, as mention previously. Some of the more significant studies in relation to our topic of interest are the following.

An experimental study conducted by Gürses and Okan [13], found proven that more obvious brand placements formed better recall in viewers over subtle placements and a combination of audio and visual appearance of a brand yielded better recall. Evidence was derived using an implicit questionnaire; a situational case where participants were asked to make decisions that would indicate brand preference. Building on these findings, Valenzuela-Fernández, Martínez-Troncoso and Yáñez-Wieland [14] found that brand recall and awareness level is enhanced the higher the level of integration of the product into the main plot of the movie/series, but more importantly, it discarded the findings of a preceding study, by proving that unconscious memory is not as affected by background placement, as it was speculated. To derive the findings, regarding the unconscious effects of product placement, the colleagues used a purchase intention test.

Another study, conducted by Cholinski [15], further supports both findings, however he was not able to prove that positive effect on attitudes or brand preference, since no tests were conducted after the viewing of the movies and data was collected only through self-reporting by viewers, during small interviews once participants had finished watching the movies.

Furthermore, a study by Ülker-Demirel and Yıldız [16], examined the connection between the level of likeability or lack thereof of participants towards the characters, actors, product placement and movies, and their attitude towards the presented brands. The colleagues concluded that the viewpoint of viewers towards the listed elements, do have an effect over the perception of the featured brand.

The most recent study closely related to the topic we are examining was conducted by Kanimozhi and Dhas [17] and obtained a couple of findings. What the colleagues found was that product placement does prove to be an effective marketing method, leaving a long-lasting effect in the memories of viewers, as they were able to notice and later on recall placements they had observed. Moreover, they were able to conclude that the placement of established brands within media, the use of celebrity endorsement, and creative placements were able to enhance credibility, capture attention, and leave a long-lasting impression in consumers. The paper also outlines the importance of the relevance the placement holds to a movie's storyline, as well as the seamless integration of said placement into the plot, similar to the findings made by Valenzuela-Fernández, Martínez-Troncoso and Yáñez-Wieland [14].

As seen majority of the available studies examining product placement in mass media have mainly relied on explicit research methods, with some exceptions as Boerman,

van Reijmersdal and Neijens [18] examining the effects of brand placement in TV programs. Despite, the use of implicit research methods in the area remains relatively limited, hence why the current study incorporates eye-tracking with the aim to complement existing approaches and offer additional perspective into the discourse [19–22].

3 Methods

The study was conducted in the form of an eye tracking experiment, aiming to measure the implicit visual attention of participants towards product placement, while viewing selected movie scenes. Data was collected using "Tobii Pro X2-30", allowing us to capture and analyze the main metrics of interest, namely Time to First Fixation (TFF), Total Fixation Duration (TFD) and Fixation Count. To be able to compare performance across scenes and brand categories, we used secondary metrics, such as Percentage Fixated, Average Fixation Count (AFC), Average Time to First Fixation (ATFF), Average Total Fixation Duration (ATFD) and Average Total Fixation Points (ATFP) per second of exposure.

The stimulus material consisted of nine movie scenes featuring product placements, grouped according to brand familiarity (well-known vs. not well-known) and placement position (foreground vs. background) (Table 1). However, only 8 are included in the current analysis, as the excluded scene contained multiple brands with the aim of separate analysis.

Table 1. Complete list and detailed description of the movie used in the study.

Movie Title	Brand in Scene	Type of brand	Type of Scene and Description	Position of Brand	Scene Duration (minutes)
The Shinning	Heinz	Well-known	Calm conversation: characters discussing available food, while on a tour of a hotel (no scary moment)	Background	00:03:03
E.T.	Reese's Pieces	Well-known	Key plot scene: boy riding a bike and placing Reese's Pieces (no scary moment)	Foreground	00:01:23
Ready Player One	Starbucks	Well-known	Fight scene: fight between fictional characters from famous games and movies, in a virtual reality setting	Background	00:04:37
Jaws 2	Coca Cola	Well-known	Romantic: couple on a boat (no scary moment)	Foreground	00:01:05
Man of Steel	IHOP	Not Well-known	Fight scene: fight between superman and villain ending in a crash into an IHOP	Background	00:01:57

			Catch it if you can		
Little Nicky	Popeye's Chicken	Not Well-known	Calm conversation: a dog teaching a person how to eat and behave	Fore-ground	00:00:56
Incep-tion	Famima!!	Not Well-known	Slightly tense conversation: characters going over their plan of action	Back-ground	00:02:40
Free Guy	Alienware	Not Well-known	Slightly tense conversation: characters discussing the fate of a game they developed	Fore-ground	00:04:00

Note. **Depends on individuals' interest and familiarity with such technologies (e.g., known mainly by people who are interested in gaming).*

The study was conducted at CITY College, University of York Europe Campus in Thessaloniki, Greece and involved 44 participants, all of legal age (above 18 years old), willing and able to consent to the study upon receiving all relevant information.

Participants completed the study individually in a university lab setting. For the collection of authentic and unbiased data, participants were informed that we were measuring their visual attention while watching a selection of movie scenes. After calibration of the eye-tracking device, they watched the selected movie scenes while their eye movements were recorded. Once the viewing was completed, participants filled in a questionnaire containing closed and Likert-scale questions about the movie scenes and the featured brands. This allowed comparison between implicit eye tracking data and self-reported responses. After completing both stages, participants were debriefed on the topic of the study and given the option to withdraw their data.

4 Results

The conducted research led to the following findings and answers to the research questions imposed at the beginning.

4.1 Implicit Visual Attention and Retention of Product Placement

To begin with, the study found that the majority of participants implicitly noticed product placements in the movie scenes, with higher fixation levels influenced by both brand familiarity and brand placement. The results suggest that shape, size, color, frequency of appearances and on-screen position greatly influence the amount of implicit attention a product placement receives. More specifically, product placements positioned in the foreground attracted greater and more sustained levels of visual attention in comparison to those featured in the background.

This effect is amplified when the product placement was integrated into the film's storyline, as in the case of Reese's Pieces in E.T. Fixation count indicated that participants' visual attention repeatedly returned to such placements throughout the scenes, suggesting that they were successful in retaining implicit attention. Average fixation

duration was also relatively high, often exceeding typical adult fixation levels. Overall, placements closely tied to the storyline or characters attracted greater unconscious visual attention [14].

Despite the observations that participants paid high levels of unconscious attention to product placements, the retention they had of the brands was not as significant. When asked to select the brands they remember noticing, participants could pinpoint only a few of the ones they had noticed. Several participants unconsciously directed visual attention towards brands within the movies, despite, when asked which of the brands they remember noticing, they report no recollection of seeing a particular product placement.

We even detected false recollections, where participants testified noticing brands, they had not established fixations upon while viewing the scenes, or even brands that were not present in the viewed material. This suggests the retention of the unconsciously noticed product placements is relatively low. This could be justified by the previous findings of Guo et al. [23] and Boerman, van Reijmersdal and Neijens [18], concluding that pre-viewing product placement disclosure, yielded higher retention and recognition in viewers. Nevertheless, these findings do not necessarily reduce the value of product placement, as its influence may extend to unconscious consumer responses beyond conscious memory.

4.2 Brand Familiarity's Effect on Implicit Visual Attention towards Product Placement

In regards to brand familiarity's influence on product placement, both well-known and less well-known brands attracted unconscious visual attention, however, well-known brands performed better across the eye-tracking metrics [16].

There were, however, some exceptions, such as Starbucks, which despite being a globally recognized brand, performed substantially worse than both the other well-known brands and the less well-known brands. Clearly the cause is not low brand familiarity, but the brand's scene selection, position on screen, with unsuitable size, supporting the view that placement position and features can influence the level of unconscious visual attention paid to a product placement. Results of the comparison of Well-known vs Not Well-known Brands is shown in Table 2.

Table 2. Comparison of Well-known vs Not Well-known Brands showing AFC, ATFF, ATFD, and ATFP Results.

Brand & Film	Type of Brand	Fixation (%)	AFC (pp)	ATFF (s)	ATFD (s)
Reeses' Pieces (E.T.)	Well-known	93,18	6,14	17,67	1,30
Coca Cola (Jaws 2)	Well-known	81,82	2,73	44,95	0,66
Starbucks (Ready Player One)	Well-known	11,36	0,27	250,64	0,02
Heinz (The Shining)	Well-known	88,64	3,36	122,86	0,80

				Catch it if you can	
Alienware (Free Guy)	Not well-known	47,73	0,75	179,60	0,13
Popeye's Chicken (Little Nicky)	Not well-known	56,82	1,57	7,66	0,36
Famima!! (Inception)	Not well-known	50,00	0,84	70,49	0,25
IHOP (Man of Steel)	Not well-known	34,09	0,57	107,58	0,09

4.3 Well-known Brands

Looking at the well-known brands, Reese's Pieces was the highest-scoring well-known brand, likely because it was closely integrated into the scene and central to the storyline (see Fig. 2, A). Unlike Coca-Cola, it was actively used throughout the scene rather than merely shown, which made the placement more contextually relevant and attention-grabbing. This supports Williams et al. [24], who argue that placements more closely tied to the storyline or main characters produce stronger unconscious recognition and retention.

4.4 Not Well-known Brands

The eye-tracking data for the less well-known brands did not show clear differences based on placement position. Their overall performance was relatively modest, although Popeyes Chicken performed better than the other brands. This may be because food-related stimuli tend to attract attention more effectively than non-food stimuli [25]. By contrast, Alienware may have attracted less attention because it was both less familiar and non-food related. However, our data did not provide sufficient evidence to confirm these explanations or support the related findings of Ülker-Demirel and Yıldız [16].

4.5 On-screen Positioning's Effect on Implicit Visual Attention towards Product Placement

Position-wise, both foreground and background placements attracted participants' unconscious attention; however, background placements generally produced weaker eye-tracking results. In most cases, they had a longer Average Time to First Fixation (ATFF), indicating that participants took longer to notice them. They also received slightly shorter fixation durations and fewer fixations than foreground placements.

There were also some exceptions. For example, although Heinz was positioned in the background, it produced results comparable to those of well-known brands placed in the foreground, likely due to its repeated and more prominent presence increasing the chance of notice. In contrast, Starbucks performed far less successfully, likely because its logo was smaller and appeared in the background of a visually intense fight scene.

4.6 Brand and Movie Likability and Familiarity Effect

Further, the study examined whether familiarity with and likeability of the featured brands and films influenced unconscious visual attention to product placements. Although previous research suggests that brand likeability may affect unconscious perception, our data did not show such a relationship [16]. The same study also indicates that familiarity with and likeability of the film may shape unconscious responses to product placements.

5 Limitations of the study

This study is subject to several limitations. The analysis was based on a limited number of movie scenes and product placements, which may not represent the full diversity of placement strategies used in contemporary media. In addition, the study measured visual attention through eye tracking and compared it with self-reported responses, but it did not examine subsequent purchase behavior or long-term memory effects. The experiment was also conducted under controlled laboratory conditions, which may differ from natural viewing environments where distractions and multitasking are common. Future studies should examine larger and more diverse samples, include additional neuromarketing measures, and investigate the long-term behavioral impact of product placements.

6 Conclusion

In conclusion, combining explicit methods (self-reported questionnaires) with implicit methods (eye tracking) provided us with a richer understanding of viewers' responses to product placements in movies. The results showed that all placements were noticed unconsciously, although their ability to attract and sustain attention varied depending on factors such as visual prominence, frequency of appearance, screen position, and integration into the storyline. Placements that were more prominent or more closely tied to the plot, such as Reese's Pieces in E.T., attracted greater attention, while well-known brands performed better only when these additional factors were also favorable.

The self-reported data further showed that participants often lacked conscious recollection of the placements and some even falsely recalled brands that were not present. However, this does not rule out the possibility of unconscious influence on follow-up consumer behavior. In contrast to some previous studies, no relationship was found between participants' likeability or familiarity with the brands or movies and their unconscious attention to product placements. Overall, the study demonstrates the value of combining explicit and implicit methods in product placement research and highlights the contribution of eye tracking and other neuromarketing approaches to understanding consumer behavior.

Appendices

Fig. 1. Product Placement in Foreground with AOI/s Indicated in Red

Fig. 2. Product Placement in Background with AOI/s Indicated in Red

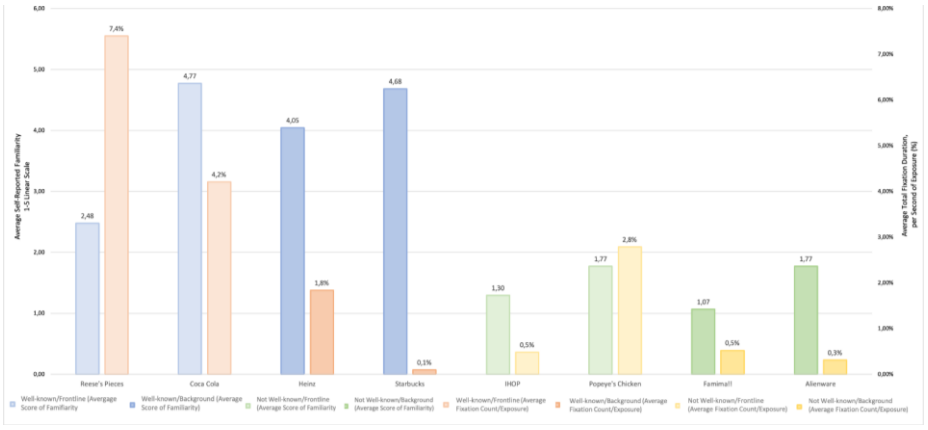


Fig. 3. ASF vs ATFP per Seconds of Exposure, per Brand

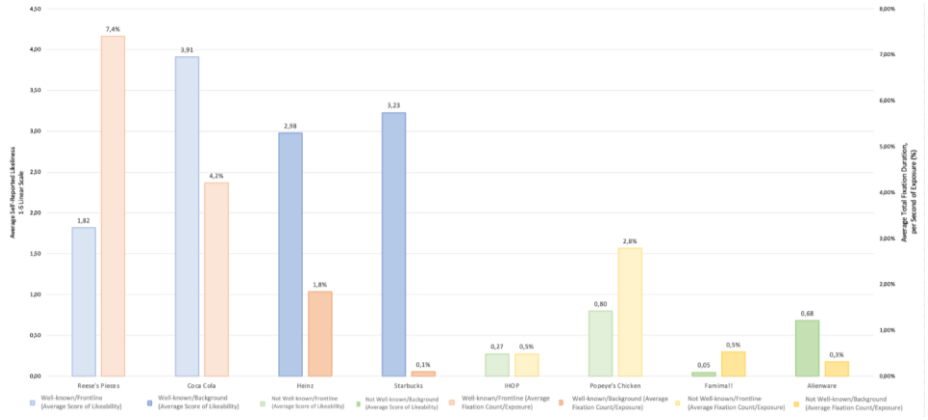


Fig. 4. ASL vs ATFP per Second of Exposure, per Brand

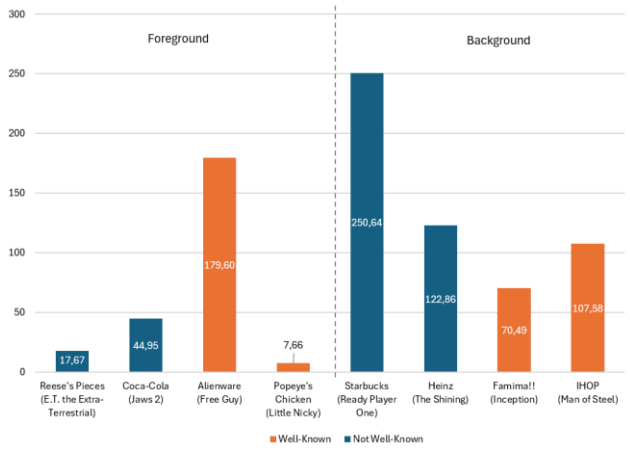


Fig. 5. ATFF Comparison Between Product Placement in Foreground vs Background

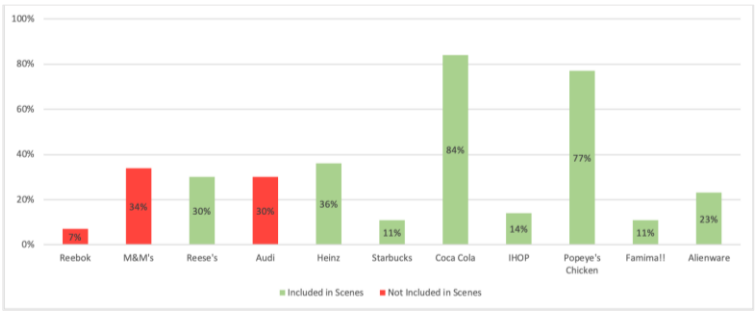


Fig. 6. How Often Brands were Noticed, Based on Self-Reported Data by All Participants Q2

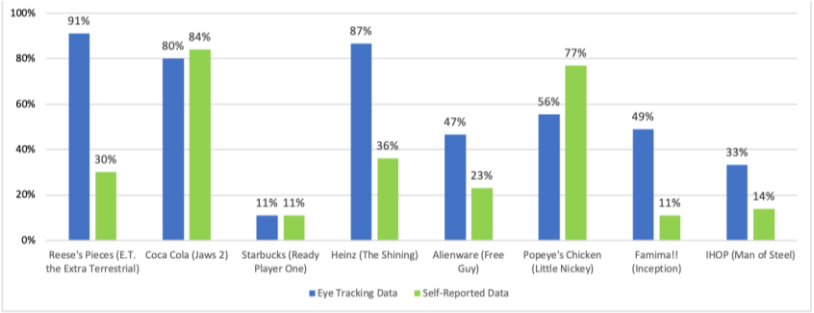


Fig. 7. Percentage of Participants Who Noticed Brands: Eye Tracking Data vs Self-Reported Data

Disclosure of Interests. Author Matteo Venerucci has been an employee of Tobii AB.

References

1. PQ Media: Global Product Placement Spending Grew 12.3% in 2023 to $29.6B, Slowing from 14.3% in 2022 Due to Impact of Strikes; Spending to Post Strong, But Slower, Growth in 2024. Prweb.com (2024)
2. Janić, M., Ćirović, M., Dimitriadis, N., Jovanović Dimitriadis, N., Alevizou, P.: Neuroscience and CSR: Using EEG for assessing the effectiveness of branded videos related to environmental issues. Sustainability 14(3), 1347 (2022). https://doi.org/10.3390/su14031347
3. Ćirović, M., Dimitriadis, N., Janić, M., Alevizou, P., Dimitriadis, N.J.: More than words: Rethinking sustainability communications through neuroscientific methods. Journal of Consumer Behaviour 23(1), 15–30 (2022). https://doi.org/10.1002/cb.2125
4. Symeonidis, S., Dimitriadis, N., Ćirović, M., Venerucci, M.: The eyes are the windows to the brain: Pupillometry and eye fixation patterns in smartphone product communications. Zbornik radova Ekonomskog fakulteta u Rijeci 43(2), 485–510 (2025) https://doi.org/10.18045/zbefri.2025.2.3
5. Katsuki, F., Constantinidis, C.: Bottom-Up and Top-Down Attention: Different Processes and Overlapping Neural Systems. The Neuroscientist 20(5), 509–521 (2014). https://doi.org/10.1177/1073858413514136
6. Corbetta, M., Shulman, G.L.: Control of Goal-Directed and Stimulus-Driven Attention in the Brain. Nature Reviews Neuroscience 3(3), 201–215 (2002). https://doi.org/10.1038/nrn755
7. Babin, B.J., Herrmann, J.-L., Kacha, M., Babin, L.A.: The Effectiveness of Brand Placements: A Meta-Analytic Synthesis. International Journal of Research in Marketing 38(4), 1017–1033 (2021). https://doi.org/10.1016/j.ijresmar.2021.01.003
8. Bressoud, E., Lehu, J.-M., Russell, C.A.: The Product Well Placed: The Relative Impact of Placement and Audience Characteristics on Placement Recall. Journal of Advertising Research 50(4), 374–385 (2010). https://doi.org/10.2501/S0021849910091622
9. Carrasco, M.: Visual Attention: The Past 25 Years. Vision Research 51(13), 1484–1525 (2011). https://doi.org/10.1016/j.visres.2011.04.012
10. Reimer, C.B., Schubert, T.: Visual and Central Attention Share a Capacity Limitation When the Demands for Serial Item Selection in Visual Search Are High. Attention, Perception, & Psychophysics 82(2), 715–728 (2020). https://doi.org/ 10.3758/s13414-019-01903-4
11. Musslick, S., Cohen, J.D.: Rationalizing Constraints on the Capacity for Cognitive Control. Trends in Cognitive Sciences 25(9), 757–775 (2021). https://doi.org/10.1016/j.tics.2021.06.001
13. Gross, J., Schmitz, F., Schnitzler, I., Kessler, K., Shapiro, K., Hommel, B., Schnitzler, A.: Modulation of Long-Range Neural Synchrony Reflects Temporal Limitations of Visual Attention in Humans. Proceedings of the National Academy of Sciences 101(35), 13050–13055 (2004). https://doi.org/10.1073/pnas.0404944101
14. Gürses, İ.Ö.A., Okan, E.Y.: Effectiveness of Product Placement: An Experimental Study in Turkey. Canadian Social Science 10(1), 56–63 (2014). https://doi.org/10.3968/J.CSS.1923669720141001.4202
15. Valenzuela-Fernández, L., Martínez-Troncoso, C., Yáñez-Wieland, F.: Influence of Placement on Explicit and Implicit Memory of College Students. Comunicar 22(44), 169–176 (2015). https://doi.org/10.3916/C44-2015-18

16. Cholinski, A.: The Effectiveness of Product Placement: A Field Quasi-Experiment. International Journal of Marketing Studies 4(5) (2012). https://doi.org/10.5539/ijms.v4n5p14
17. Ülker-Demirel, E., Yıldız, E.: The Effects of Audience's Attitudes on Actor, Character, Movie and Product Placement on the Brand Attitude. Istanbul Business Research 49(2), 339–359 (2021). https://doi.org/10.26650/ibr.2020.49.0013
18. Kanimozhi, V., Dhas, D.A.: A Study of Factors Affecting Consumer Evaluations and Memory of Product Placement in Movies. International Journal of Advanced Research in Science Communication and Technology 3(3), 298–304 (2023). https://doi.org/10.48175/IJARSCT-11449
19. Boerman, S.C., van Reijmersdal, E.A., Neijens, P.C.: Using Eye Tracking to Understand the Effects of Brand Placement Disclosure Types in Television Programs. Journal of Advertising 44(3), 196–207 (2015). https://doi.org/10.1080/00913367.2014.967423
20. Rayner, K.: Eye Movements and Attention in Reading, Scene Perception, and Visual Search. Quarterly Journal of Experimental Psychology 62(8), 1457–1506 (2009). https://doi.org/10.1080/17470210902816461
21. Vulpe-Grigorasi, A., Kren, Z., Slijepčević, D., Schmied, R., Leung, V.: Attention Performance Classification Based on Eye Tracking and Machine Learning. In: 2024 IEEE 17th International Scientific Conference on Informatics (Informatics), pp. 431–435 (2025). https://doi.org/10.1109/Informatics62280.2024.10900810
22. Wedel, M.: Attention Research in Marketing: A Review of Eye Tracking Studies. Robert H. Smith School Research Paper No. RHS 2460289 (2013). https://doi.org/10.2139/ssrn.2460289
23. Białowąs, S., Szyszka, A.: Eye-Tracking in Marketing Research. In: Managing Economic Innovations – Methods and Instruments, pp. 91–104 (2019).
24. Guo, F., Ye, G., Hudders, L., Lv, W., Li, M., Duffy, V.G.: Product Placement in Mass Media: A Review and Bibliometric Analysis. Journal of Advertising 48(2), 215–231 (2019). https://doi.org/10.1080/00913367.2019.1567409
25. Williams, K., Petrosky, A., Hernandez, E., Page Jr., R.: Product Placement Effectiveness: Revisited and Renewed. Journal of Management and Marketing Research 7, 1–24 (2011)
26. Werthmann, J., Roefs, A., Nederkoorn, C., Mogg, K., Bradley, B.P., Jansen, A.: Attention Bias for Food Is Independent of Restraint in Healthy Weight Individuals—An Eye Tracking Study. Eating Behaviors 14(3), 397–400 (2013). https://doi.org/10.1016/j.eatbeh.2013.06.005

THEMATIC SECTION

Applied Software Engineering

Resilient and Intelligent Organizations
in Times of Permanent Disruption

SYMORG 2026 • CONFERENCE PROCEEDINGS

Original Research Article *https://doi.org/10.5281/zenodo.23472378*

Designing a Domain-Specific Conversational AI for a Building Materials Retailer

Askar Akhmetkhanov[1[0009-0006-3394-7470]] and Marko Pezer[1[0009-0003-9056-7712]]

1 Innopolis University, Innopolis, Russia
a.akhmetkhanov@innopolis.university
m.pezer@innopolis.university

Abstract. Offline retail still lags behind e-commerce in providing immediate customer assistance. The paper presents the design and evaluation of a domain-specific conversational AI assistant for a building-materials retailer. The system is designed to answer customer questions about products, prices, availability, and recommendations while relying solely on the store's local database. The final solution is built as an offline retrieval-augmented generation (RAG) system with multilingual dense retrieval, BM25-based lexical search, reciprocal rank fusion, cross-encoder reranking, and a local language model. A speech recognition layer is added in the final iteration, allowing users to speak their prompts instead of typing. The paper focuses on the final architecture and reports the main design decisions that made the RAG assistant possible under retail constraints, including mixed Russian and English product terminology, connection to the store's database, and safe refusal for irrelevant questions. Evaluation on in-domain and off-topic prompts suggests that hybrid retrieval and reranking improve the quality of evidence passed to generation, while the refusal mechanism prevents use of the language model on irrelevant or weakly supported queries and therefore reduces the opportunity for unsupported generation. The resulting prototype provides evidence that an offline AI assistant for in-store retail is technically feasible and potentially useful as a customer-support system.

Keywords: conversational AI · retrieval-augmented generation · retail assistant · offline AI

1 Introduction

Physical stores still face a basic service gap; for example, customers often need quick answers about product location, availability, or price, but store employees are not always immediately available since most of them are busy unloading goods. This problem is especially visible in building materials retail, where product choice is complex and questions are often domain-specific. Prior work suggests that understaffing can reduce store performance and that computer agents can improve customer service efficiency [1,2]. However, most conversational systems are designed for online commerce or depend on cloud services rather than local in-store deployment. This paper addresses that gap through the design of an offline conversational assistant for a building-

materials retailer. The assistant is connected to the retailer's store catalog and supports natural customer questions about products, features, prices, availability, and simple recommendations. The central design goal is a more controllable assistant within a narrow and operationally relevant retail domain. The work is motivated by three practical constraints. First, the system must function without reliance on cloud APIs, which makes local inference and local data access necessary. Second, it has to work with Russian-language customer queries that may contain English brand names and technical terms. Third, it needs to reduce opportunities for unsupported answers and correctly handle irrelevant questions. These constraints make retrieval quality and grounded generation more important than raw fluency. The contribution of the paper consists of three parts. First, it presents a concise architecture for an offline retail assistant built around a retrieval-augmented generation (RAG) system. Second, it shows why the final solution uses hybrid retrieval and reranking rather than dense retrieval alone. Third, it evaluates the prototype using representative customer prompts and argues, within the limits of the reported evaluation, that safe refusal is a desirable design feature for a domain-specific retail assistant. Section 2 reviews the literature on retail assistants, voice interfaces, and retrieval-augmented generation. Section 3 summarizes the methodology, system requirements, and final design logic. Section 4 describes the implementation of the final prototype with emphasis on the final retrieval stack and the voice entry point. Section 5 presents the evaluation and discusses both the benefits and the current limitations of the approach. Section 6 provides the conclusion for the paper.

2 Literature Review

2.1 Conversational AI for Retail and Voice Interfaces

Conversational assistants are gaining popularity in business since they reduce routine workload and improve response time [3,4]. In retail, the strongest results have been reported for product search, recommendation, and basic service guidance [5]. At the same time, prior research also highlights known limitations of voice-user interfaces, especially speech recognition errors and recovery from misunderstandings [6,7]. These limitations are significantly important in noisy physical environments such as stores. For the present problem, the most relevant implication from previous work is that retail assistants must be domain-constrained and operationally reliable. A store assistant is useful only if it produces answers based on information from the retailer's database rather than fluent but unsupported responses.

Industry examples also support this conclusion. Recent deployments in Russia and abroad show that in-store assistants are usually introduced as narrowly scoped service systems rather than as open conversational agents. Their value comes from handling repeated operational questions such as prices, returns, promotions, product search, and store guidance. This observation is necessary for the present paper since it justifies a domain-specific design objective, which is about maximizing trustworthiness and practical usefulness within a narrow retail scope, even if that means deliberately limiting open-ended conversational behavior. A second recurring lesson from earlier work is

that voice interactions are helpful only when there are clear recovery strategies. Studies about voice user interfaces (VUIs), which are interfaces in which users interact with a system through spoken input and outputs of the system, repeatedly emphasize confirmation, repetition, and error handling because users quickly lose trust when the system mishears their intent and still answers confidently. In the context of a building-materials store, this risk is amplified by noisy environments and product names containing brand-specific or technical vocabulary. As a result, the final prototype treats voice input as an interface layer that must remain subordinate to the same grounding and guardrail logic used for text input. The literature on deployed retail assistants also suggests that human acceptance depends on the clarity of the task. Users are more willing to engage with an assistant when they can quickly understand what kinds of questions it can answer. This further supports a narrow product-support scope for the present system. Rather than positioning the assistant as a general conversational agent, the design presented here frames it as a specialized in-store support tool for product discovery, availability checks, and basic recommendations.

2.2 RAG and Retrieval Quality

RAG is a fit for domain-specific assistants since it contains retrieval of external knowledge [8,9]. Recent work has shown that retrieval design strongly affects final answer quality, specifically in specialized areas. Hybrid approaches include dense retrieval with lexical retrieval to improve recall, while rerankers improve precision by selecting the most relevant evidence from the candidate set [10,11,12]. This is especially useful when queries contain both semantic intent and exact lexical constraints such as brand names, model numbers, or product types. The literature, therefore, provides two conclusions that guide this paper. First, an in-store assistant should provide responses based on domain data rather than rely on a general-purpose language model. Second, the retrieval stage should be treated as the main control point for answer quality. Another important implication from the literature is that offline deployment is not just a preference; it changes the whole system design. When cloud APIs are unavailable or undesirable, the embedding model choice, database, reranker, and speech recognizer must be driven by local execution constraints and predictable latency. This leads to a modular pipeline in which each component can be inspected, tuned, and replaced independently.

The literature also helps explain why the final solution does not rely only on a single retrieval strategy. Dense retrieval is strong for semantic matching, but retail queries often contain exact lexical constraints such as product series, brand names, material types, or pack sizes. Purely lexical retrieval has the opposite weakness: it can miss semantically appropriate results when users express intent through paraphrases or lay vocabulary. Foundational work on dense retrieval and later studies on hybrid retrieval pipelines support the view that dense and lexical signals should be treated as complementary rather than competing mechanisms [19,10]. The hybrid design adopted in this paper, therefore, follows directly from that tension. It is not an ordinary improvement, but a response to the mixed semantic and lexical nature of realistic store queries. Recent RAG research also points toward increasingly modular and adaptive architectures.

Approaches such as modular RAG, Self-RAG, and Adaptive-RAG emphasize that retrieval should be dynamically controlled rather than treated as a fixed pre-processing stage [20,21,22]. Although the present system is deliberately simpler than those architectures, these studies reinforce the same core idea: answer quality depends significantly on how evidence is selected, filtered, and passed to generation. Finally, prior studies of domain adaptation and on-device language modeling suggest that retrieval often offers a more controllable path to specialization than end-to-end fine-tuning alone, especially when systems must operate under local deployment constraints [23,24]. For this reason, the paper focuses on the retrieval and grounding pipeline as its main contribution. The resulting system is easier to inspect, easier to evaluate under operational constraints, and better aligned with the need for safe refusal when relevant catalog evidence is missing.

3 Methodology

3.1 Research Design

The project has an iterative and design-oriented methodology. Instead of rating a single monolithic system, the work develops the assistant in stages and validates the practical impact of each major design decision. The baseline stage establishes that a fully local RAG pipeline can answer store-related questions from catalog data. The second stage improves retrieval quality by having a hybrid search and reranking. The final stage adds speech recognition so that spoken queries can enter the same grounded pipeline. The evaluation logic follows the same structure. The system is assessed not only by answer fluency but by whether it retrieves relevant products, keeps answers from the retrieved catalog evidence, refuses unsupported questions, and remains practically useful for a customer-facing retail scenario. This design treats the assistant as a decision pipeline rather than as a standalone language model. In a retail setting, this distinction plays an important role. The main decisions are whether the query is within the supported store domain, which product records should be used as evidence, whether the evidence is strong enough to answer, and whether the system should refuse instead of invoking the language model. The system is considered successful only when it can justify an answer through retrieved evidence or refuse when no such evidence is available. Controllability is achieved through local data access, relevance thresholds, and prompt restrictions; transparency is achieved by inspecting retrieved product records before generation; and incremental validation is achieved by comparing dense retrieval, hybrid retrieval, reranking, and voice input as separate stages.

3.2 System Requirements

The final solution is shaped by several requirements derived from the target environment. These requirements are not abstract design preferences; they follow directly from the operational realities of an offline retail assistant that must work with local data, limited infrastructure, and customer-facing reliability constraints. In practice, the

system is expected to support short, goal-oriented interactions in which clients ask about products, prices, availability, and simple recommendations, while the assistant must remain connected to the catalog evidence and avoid unsupported answers. For this reason, the requirements below define not only the expected functionality of the final prototype, but also the constraints that guided the choice of models, retrieval strategy, and responsegeneration logic. First, the assistant must operate within the retailer's internal network, without dependence on external APIs. This avoids per-request cloud costs and external data transfer, but it shifts maintenance effort to the retailer: local hardware must be provisioned, models and indexes must be updated, and catalog synchronization must be monitored. The prototype cost is therefore mainly implementation and local compute cost rather than API usage cost; no full operational cost analysis was conducted, which remains a deployment limitation. Second, all answers must remain grounded in the retailer's own catalog, stock data, and product metadata rather than in general background knowledge. Third, the system must support Russian-language queries that may contain English brand names and technical terms, because this mixture naturally appears in real retail data. Fourth, unsupported or off-topic questions should trigger safe refusal instead of unsupported answers. Finally, the assistant should produce short and clear responses that are genuinely useful for customers making purchase decisions in-store. Short interactions are supported by product-centered retrieval, concise prompt instructions, and answer formatting that prioritizes direct product names, prices, availability, and one or two decision-relevant details instead of open-ended conversation.

3.3 Final Solution: Iteration 3

The final solution keeps the validated retrieval pipeline and adds a speechrecognition front end. Spoken customer input is transcribed locally and then

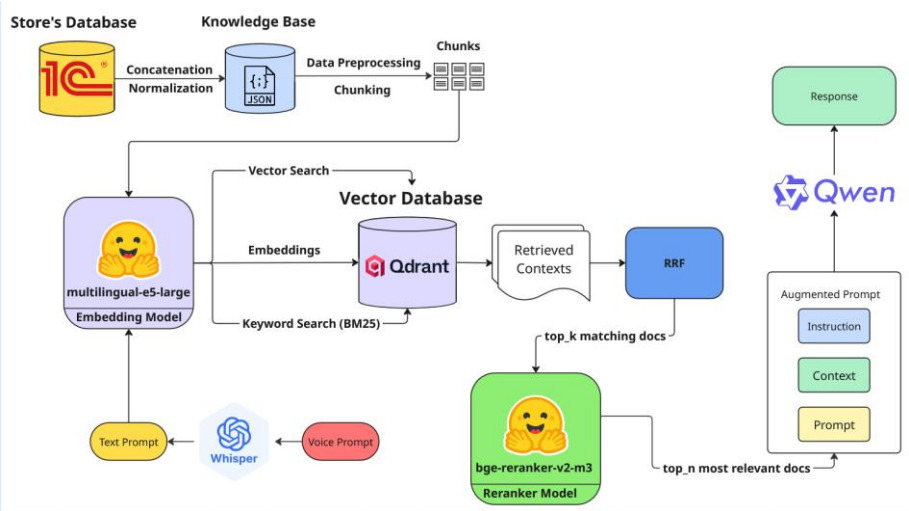


Fig. 1. Final solution architecture. Spoken input is transcribed locally and then passed through the same retrieval, reranking, and grounded answer-generation pipeline.

processed exactly like a text query, where the query is embedded, relevant catalog entries are retrieved, candidates are reranked, and an LLM produces the final answer. This design is intentionally conservative. Rather than changing several components at once, the final iteration adds only the interface layer while preserving the retrieval and generation logic already validated in earlier stages. The methodological importance of Iteration 3 is not that it introduces a new reasoning model, but that it turns a text-based prototype into a more realistic in-store assistant while keeping the system grounded in the store catalog.

3.4 Grounding and Safety Logic

Grounding is enforced at two levels. First, the retrieval stage limits the answerable space to products and attributes that exist in the retailer's catalog. Second, the generation stage is executed only when the retrieved evidence passes a fixed relevance-threshold check. In this prototype, the threshold represents the minimum reranker confidence needed for the top product evidence to be treated as answerable. The threshold was held constant during the 40-prompt evaluation, but it should be interpreted as a prototype-specific engineering parameter rather than as a universal semantic truth score. The current paper, therefore, reports its role in the decision logic, while more systematic threshold calibration is left for future work. The purpose of the cutoff is to prevent weak lexical overlap from automatically triggering generation. This design reduces the chance that the language model will answer based on prior world knowledge rather than from the store catalog. In the target application, such control is more valuable than broad generative flexibility, because a wrong answer about price, availability, or product suitability can directly mislead a customer.

The refusal mechanism is therefore part of the methodology rather than a fallback added after deployment. If no sufficiently relevant context is found, the system should abstain. This principle is preserved across all iterations and becomes especially important in the final voice-enabled version, where speech recognition errors could otherwise increase the risk of unsupported answers.

3.5 Iteration Logic and Scope Reduction

The system was developed across several iterations, but this paper focuses on the final solution and includes earlier iterations as they explain design rationale. Iteration 1 established that a fully local baseline could retrieve product-centered evidence and produce grounded answers for a subset of store questions. Iteration 2 improved the quality of retrieval by adding hybrid search and reranking. Iteration 3 then integrated spoken input while preserving the validated retrieval backbone. This progression matters because it shows that the final architecture is not a speculative design; it is the result of sequential refinement under stable deployment constraints.

Table 1. Role of the three development iterations in the final system.

Iteration	Main change	Contribution to the final paper
Iteration 1	Local baseline RAG	Demonstrated that grounded answers were feasible using only local catalog data.
Iteration 2	Hybrid retrieval + reranker	Improved evidence quality before generation and reduced weakly supported answers.
Iteration 3	Voice input layer	Turned the validated text pipeline into a more realistic in-store assistant interface.

This scope reduction is also methodologically useful for a paper. Instead of documenting every implementation branch, the paper highlights the decisions that most directly affect final-system quality, which are product-centered data normalization, multilingual embeddings, hybrid retrieval, reranking, and safe refusal. These are the parts of the project most likely to generalize to other offline, domain-specific retail assistants.

4 Implementation

4.1 Data and Core Components

The implementation uses approximately 10,000 products exported from the retailer's internal 1C-based system. 1C is a widely used enterprise customer relationship management (CRM) platform in Russian retail; in this project, it is the source system for product, property, and stock records. The raw data comes from three JSON sources, which are the main product catalog, product properties, and stock availability. These sources are normalized into a single structured knowledge base in which each product is represented as a text chunk containing fields such as name, category, brand, price, description, and availability. The normalization step is essential because the raw records are not directly suitable for retrieval. Product attributes are originally distributed across separate exports and use field naming conventions designed for retail systems rather than natural-language search. The implementation, therefore, merges records by product identifier, cleans noisy fields, and converts the resulting product profile into a structured text representation that can be embedded and retrieved. This transformation is what makes the catalog usable as a source for natural-language interaction. A product-centered chunking strategy is used instead of ordinary fixed-size text segmentation. This decision follows from the structure of the data, where each product already forms a meaningful unit for search and recommendation. Preserving that unit helps the retriever return self-contained evidence that the language model can reuse directly. It also reduces the likelihood that product attributes such as price, brand, stock, and description become separated across unrelated chunks. In a retail assistant, such separation would

harm both relevance and factual consistency. The final chunk format is designed to balance natural-language richness with retrieval efficiency. If the text is too sparse, semantically relevant products may not surface. If the text is too verbose, retrieval becomes noisier and the final prompt context becomes harder to control. The implementation, therefore, favors compact, structured product profiles that include the fields most likely to support customer-facing answers. This is achieved by selecting stable customer-facing fields from the merged export, removing unnecessary attributes, normalizing availability and price fields, and serializing each product in a consistent order so that retrieval sees comparable product descriptions. Dense embeddings are generated locally with multilingual E5 large, which was selected because the catalog is predominantly in Russian but also contains English brand names and mixed-language terminology [13]. Embeddings are stored in a local Qdrant index together with structured metadata for filtering and retrieval [14,15]. For answer generation, the prototype uses a local Qwen3-14B model served through Ollama, and unsupported cases are filtered before generation using a relevance threshold [16]. If no sufficiently relevant evidence is found, the system returns a refusal response instead of invoking the language model. The choice of multilingual-e5-large is also supported by internal comparison against ru-en-RoSBERTa in the prototype experiments. While both models are great for Russian-language retrieval, the final results favored multilinguale5-large because it provided a stronger balance between in-domain relevance, reliability to mixed-language product naming, and fewer observed unsupportedgeneration risk cases in downstream answers. This mattered in practice because many building-materials queries mix everyday Russian phrasing with English brand names or model identifiers.

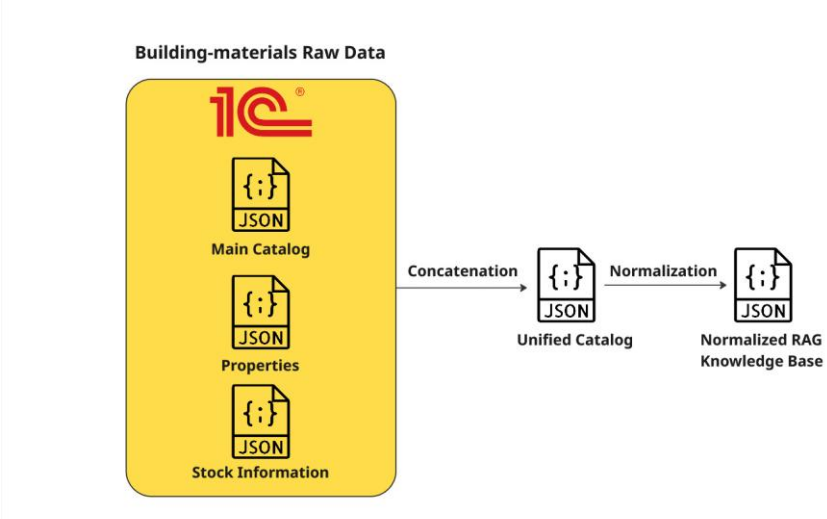


Fig. 2. Condensed data-ingestion workflow used to convert raw retailer exports into a retrieval-ready knowledge base.

Qdrant was selected not only as a vector index but also as a practical retrieval layer for structured retail data. In addition to dense retrieval, the stored payload enables brand-, category-, and availability-aware filtering, which creates a path for future query refinement. Even when such filters are not aggressively used in the current prototype, the implementation preserves this capability because retail assistants often need to evolve from general recommendation to constraintsensitive product search.

4.2 Retrieval and Grounded Generation

The implemented retrieval stack reflects the main findings from the iterative development process. Dense retrieval alone was sufficient for a baseline prototype, but the final system uses a hybrid configuration in which dense retrieval is combined with BM25 lexical retrieval and fused with reciprocal rank fusion (RRF). This improves recall for queries that mix semantic intent with exact terms such as brand names, product labels, or highly specific product categories. The fused candidate set is then reranked with the multilingual cross-encoder BAAI/bge-reranker-v2-m3, which improves precision before generation [17].

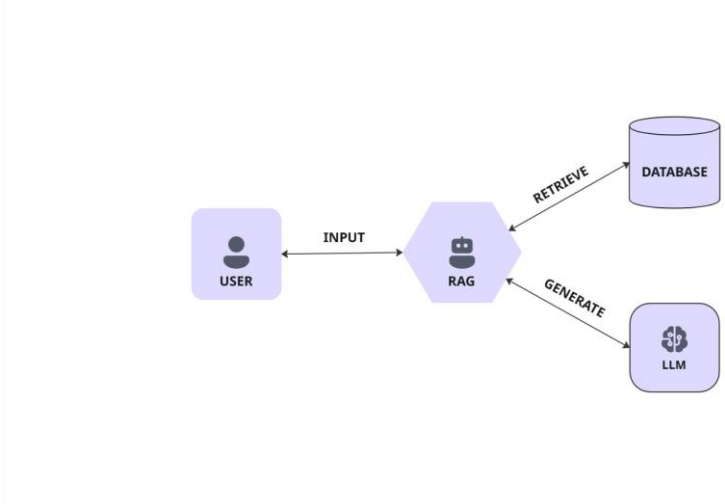


Fig. 3. High-level implementation architecture showing the main modules around retrieval, generation, and user interaction.

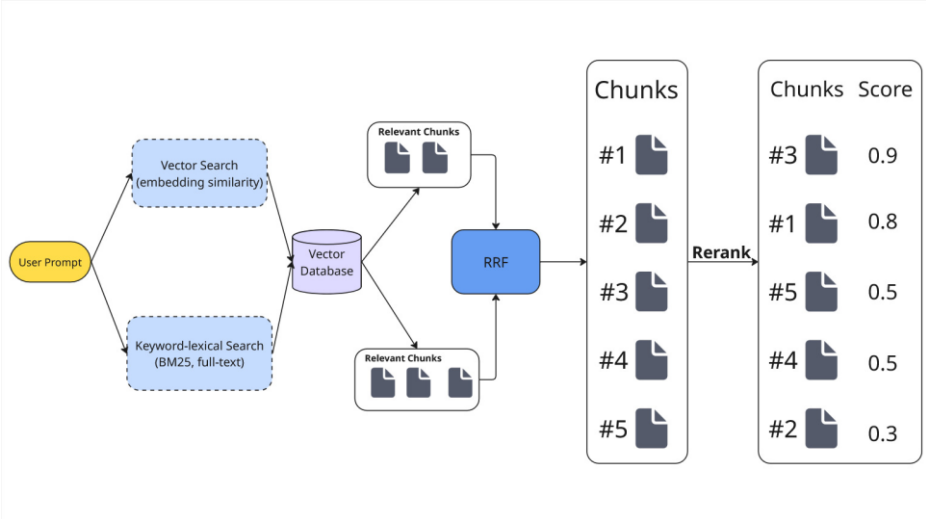


Fig. 4. Final retrieval stack used by the prototype: dense retrieval and BM25 are fused with RRF, then reranked before grounded generation.

This architecture is crucial for retail since many queries combine semantic and lexical constraints. For example, a customer could ask for a product by use case while also mentioning a brand, material, or technical specification. The hybrid stage broadens the candidate pool, while reranking keeps the final prompt context small and focused.

The generation stage is deliberately constrained by prompt design. The retrieved context is concatenated into a compact prompt together with a strict instruction to answer only from the provided catalog evidence. If relevant documents do not pass the threshold, the language model is not called. This guardrail is simple, but it is central to the implementation because it prevents the model from filling gaps with unsupported claims. Another implementation detail is that retrieval and reranking are treated as separable modules. This makes the prototype easier to debug and evaluate because candidate quality can be inspected before answer generation. In practical terms, this separation was useful during development: it made it possible to compare raw dense retrieval, fused hybrid retrieval, and reranked outputs on the same prompt set and to understand where answer failures actually came from. For a conference paper, this modularity is also useful because it clarifies that the performance gains of the final system are not attributed vaguely to “the model,” but to specific retrieval improvements.

4.3 Voice Input in the Final Prototype

The final prototype adds local speech recognition as the query-entry mechanism. The speech layer is based on the same design assumptions as the rest of the system: local operation, multilingual robustness, and compatibility with Russian speech containing English product names. Whisper is used as the baseline speech-recognition choice because it is open and multilingual [25]. Mixed Russian–English input is handled at the

interface level by transcribing the spoken query into text without translating brand names or technical terms; the resulting mixed-language text is then processed by the same multilingual embedding model and lexical retriever used for typed queries. After transcription, the recognized text is passed unchanged into the retrieval pipeline. An important implementation detail is that the voice layer does not replace grounded retrieval; it only changes how the user query enters the system. This keeps the behavior interpretable and ensures that spoken interaction does not bypass the same safety constraints used for text input. From a system perspective, this decision also reduces architectural risk. Speech recognition can introduce transcription errors, but those errors are then exposed to the same retrieval thresholding, reranking, and refusal logic as any typed query. In other words, the voice layer broadens accessibility without weakening the core grounding mechanism. This is a useful design principle for domainspecific assistants in noisy or operational environments.

5 Evaluation

5.1 Setup and Criteria

The evaluation is based on 40 representative customer prompts covering both in-domain and off-topic scenarios. The prompt set was constructed manually from the expected use cases of an in-store building-materials assistant: product lookup, product recommendation, price or availability-oriented questions, mixed semantic–lexical queries containing brands or technical terms, ambiguous storerelated prompts, and clearly off-topic prompts. The off-topic subset also includes prompts that may accidentally overlap with isolated catalog terms but do not correspond to genuine store information needs. This construction was chosen to test not only ordinary successful answers, but also boundary cases in which retrieval noise or lexical overlap could incorrectly trigger generation.

Four criteria are used throughout the evaluation: relevant retrieval, factuality of the final answer, guardrails for unsupported prompts, and practical usefulness for a customer. This combination is suitable for an offline retail assistant because linguistic fluency alone is not a meaningful success criterion.

Table 2. Evaluation criteria used for the prototype.

Criterion	Interpretation in the retail setting
Relevant retrieval	The retrieved products should match the customer's real intent rather than only isolated terms.
Factuality	The answer should stay within the information supported by the retrieved catalog entries.
Guardrails	The assistant should refuse when the query is outside the store domain or unsupported by evidence.

Criterion	Interpretation in the retail setting
Usefulness	The response should be actionable for a customer, for example by naming products, prices, or availability.

This style of evaluation is particularly relevant for retail assistants because operational usefulness cannot be captured by a single automatic score. A customerfacing answer should be relevant, grounded, concise, and actionable at the same time. For example, an answer that mentions the right product family but omits price or availability may still be less useful than a slightly shorter answer with stronger catalog grounding. The evaluation criteria, therefore, intentionally combine retrieval correctness with human-centered practical usefulness.

5.2 Main Results

The evaluation supports three main findings within the 40-prompt prototype benchmark. First, the local RAG assistant answered many store-related questions using only the retailer's own catalog data. Second, hybrid retrieval with reranking improved the inspected evidence reaching the language model compared with dense retrieval alone. Third, refusal behavior should be treated as a success criterion rather than a failure mode when the retrieved catalog evidence is too weak to support an answer.

Table 3. Main evaluation findings for the final prototype.

Aspect	Main finding
Embedding choice	multilingual-e5-large achieved a better balance of in-domain relevance and low unsupported-generation risk than ru-en-RoS-BERTa.
Retrieval quality	Hybrid retrieval improved recall for mixed semantic and lexical queries.
Reranking	Cross-encoder reranking improved precision by pushing the most relevant products to the top of the final context.
Guardrails	Off-topic prompts were correctly refused when valid catalog evidence was absent.
Practical trade-off	Better retrieval quality increased latency, but the trade-off remained acceptable for prototype validation.

Table 4. Summary statistics for the 40-prompt retrieval-variant evaluation.

Variant	Off-topic refusal	In-domain answer	Avg top-1	Avg R	Avg F/G/U	Avg total
ru-en-RoS-BERTa	0.875	0.500	0.542	0.625	0.650	2.575/4

Variant	Off-topic refusal	In-domain answer	Avg top-1	Avg R	Avg F/G/U	Avg total
multilingual-e5-large	1.000	0.875	0.807	0.600	0.925	3.375/4

Table 4 reports the main statistics used to select the final embedding configuration. The multilingual E5 configuration answered a larger share of valid in-domain prompts while preserving refusal behavior on off-topic prompts. Its average top-1 retrieval score was also higher, suggesting that the strongest retrieved candidate was usually more relevant before generation. Interpreting the binary criteria over 40 prompts, the final multilingual E5 configuration satisfied relevant retrieval in 24 prompts, factuality in 37 prompts, guardrails in 37 prompts, and usefulness in 37 prompts. The lower retrieval count compared with the factuality and guardrail counts indicates that retrieval was not perfect, but the downstream threshold and refusal logic often prevented weak evidence from becoming an unsupported answer. A representative in-domain example is a query asking for tile adhesive with good adhesion. In this case, hybrid retrieval returns several plausible candidates, while reranking moves the most relevant tile adhesives to the top. The resulting answer is more focused because the language model receives a cleaner and more relevant prompt context. In contrast, an off-topic prompt such as how to cook borscht still produces weak lexical matches at the retrieval stage, but the final system correctly falls back to the refusal response “I do not know based on the catalog data”. This is the desired behavior for a constrained retail assistant. A second lesson from the experiments is that retrieval improvements have a disproportionate effect on downstream answer quality. Once the candidate set contains mostly relevant products, the final answer becomes not only more accurate but also easier to phrase clearly, because the language model is no longer forced to synthesize across weakly related evidence. In the inspected prompt set, dense retrieval alone was more vulnerable to missing exact product terms, hybrid retrieval improved recall for mixed semantic and lexical queries, and reranking improved precision by moving the most relevant products into the final prompt context. This supports the central claim of the paper: for domain-specific retail assistants, the most important optimization target is often evidence selection rather than unconstrained generation. The embedding comparison also offers a practical methodological result. Even when a locally plausible Russian-focused model is available, broader multilingual retrieval can still be preferable in a mixed retail catalog. This matters because real product data rarely stays monolingual in a clean way; it often contains English brand names, manufacturer labels, and technical abbreviations embedded inside otherwise Russian descriptions.

5.3 Additional Practical Observations

Two practical observations emerged from the prototype. First, the system benefits from refusal not only in obviously off-topic cases but also in borderline cases where lexical overlap is present without true semantic support. This is important because retail catalogs often contain broad household vocabulary that can accidentally overlap with unrelated prompts. However, the present evaluation does not prove a measured reduction in

hallucination rates. Demonstrating hallucination reduction would require a controlled comparison in which the same borderline prompts are tested with and without refusal while keeping retrieval and generation settings fixed. This is therefore treated as a limitation and a future-work direction. Second, response usefulness depends on balancing completeness and brevity. Excessively long answers may be technically grounded yet still unsuitable for quick in-store interaction. The final system therefore benefits from a controlled answer style that prioritizes direct recommendations and concise supporting details. A further practical insight is that the architecture is well suited for incremental extension. Because retrieval, reranking, and generation are modular, future versions can add stronger metadata filters, location-aware reasoning, or richer store-specific constraints without redesigning the entire assistant. This modularity increases the long-term value of the prototype beyond the immediate conference-paper evaluation.

5.4 Discussion

The main lesson from the evaluation is that answer quality in this setting appears to depend strongly on retrieval quality. Once the candidate set is cleaner, the language model has a more relevant context from which to generate an answer. This finding aligns with the literature on RAG and supports the design choice to invest in hybrid retrieval and reranking instead of relying on prompt engineering alone. The current evaluation also has clear limitations. The benchmark remains prototype-scale, and several judgments are based on scenario inspection rather than large-scale automatic metrics. In addition, the final voice interface is integrated conceptually and architecturally, but it still requires a broader quantitative evaluation under real shop noise conditions. The visual user interface was not evaluated separately; the current results concern retrieval, grounding, refusal behavior, and answer usefulness rather than usability of the screen layout or interaction design. Latency is another important limitation. The prototype shows that better retrieval quality can come at the cost of longer end-to-end response time, especially when hybrid search, reranking, and generation are executed sequentially. This does not invalidate the design, but it means that deployment-oriented optimization remains an important part of future work. In practice, the current prototype should be understood as a proof of concept rather than a finalized production system.

5.5 Threats to Validity and Deployment Implications

Several validity threats should be kept in mind when interpreting the current results. The first is dataset specificity the prototype is grounded in one retailer's catalog structure, naming conventions, and product mix. Although the architectural lessons are likely transferable, absolute retrieval behavior may change in other retail domains with different data quality or product granularity. The second threat concerns evaluation scale. Because the benchmark is prototypesized and partially inspected manually, the reported findings should be read as strong design evidence rather than as a final statistical characterization of system performance. There are also deployment implications that strengthen the practical significance of the work. In a real store, the assistant would not be judged only by benchmark prompts; it would be judged by whether customers

receive timely, trustworthy answers under noisy and time-constrained conditions. From that perspective, the most important properties of the prototype are controllability, grounding, and graceful refusal. Even before full production optimization, these properties make the architecture a plausible foundation for future pilot deployment. A final implication concerns maintainability. Because the system is grounded in structured retailer data rather than handcrafted answer rules, updates to catalog content can propagate through the retrieval layer without redesigning the conversational logic from scratch. This creates a practical path from research prototype to operational assistant: improve data freshness, tune thresholds, optimize latency, and expand evaluation, while keeping the core architecture stable.

6 Conclusion

This paper presented an offline conversational assistant for building-materials retail. The final prototype combines local data access, multilingual dense retrieval, BM25 lexical retrieval, reciprocal rank fusion, cross-encoder reranking, and local language generation. A speech-recognition layer is added in the final iteration so that spoken input can be processed through the same grounded reasoning pipeline. The most important result is not simply that a language model can answer retail questions, but that a domain-specific assistant can behave more controllably when its generation is constrained by retrieval quality and safe refusal. Within the reported benchmark, hybrid retrieval and reranking strengthen the evidence selection stage, while refusal behavior prevents language-model invocation on low-evidence queries and reduces the opportunity for unsupported generation. Together, these findings suggest that offline, grounded conversational support is a viable direction for further customer-assistance research in physical retail environments. Future work should focus on reducing latency, expanding the benchmark with more realistic customer scenarios, and evaluating the voice layer under real in-store conditions. Another important direction is tighter use of metadata filters and store-location reasoning so that the assistant can provide even more actionable guidance during product search. A controlled with/without-refusal evaluation on borderline prompts is also needed before making stronger claims about measured hallucination reduction.

Acknowledgments. The authors received no specific funding for this work.

Disclosure of Interests. The authors have no competing interests directly or indirectly related to the work submitted for publication.

Ethical Approval. This work does not report a study involving human participants, animals, biological material, or personal data. The evaluation was conducted on manually constructed prompts and catalog-derived product data.

AI-Use Disclosure. AI-based tools were used only to help search for additional references relevant to the manuscript.

References

1. Mani, V., Kesavan, S., Swaminathan, J.: Estimating the impact of understaffing on sales and profitability in retail stores. Research Paper, Carolina Digital Repository (2015)
2. IBM Corporation: AI Virtual Agents: Business Impact. White paper (2018). Available at: https://www.ibm.com/downloads/documents/us-en/10a99803f8afda23. Accessed 2 June 2026
3. Blut, M., Wünderlich, N.V., Brock, C.: Facilitating retail customers' use of AIbased virtual assistants: A meta-analysis. Journal of Retailing 100(2), 293–315 (2024)
4. Khan, R.A.: Customer engagement through AI-based chatbots in e-commerce. SSRN Electronic Journal, 1–29 (2024)
5. Shanmugasundaram, G.: A review of AI-based chat-bot kiosks: Architecture, applications, and future directions. International Journal of Advance Scientific Research and Engineering Trends 9(1), 40–47 (2025)
6. Deshmukh, A.M., Chalmeta, R.: User experience and usability of voice user interfaces: A systematic literature review. Information 15(9), 579 (2024)
7. Alo, O., Wright, O., Rehman, M.A., et al.: The user experience of voice assistants in retailing: a qualitative comparative study. Qualitative Market Research: An International Journal, 1–29 (2025)
8. Lewis, P., Perez, E., Piktus, A., et al.: Retrieval-augmented generation for knowledge-intensive NLP tasks. arXiv preprint arXiv:2005.11401 (2021)
9. Gao, Y., Xiong, Y., Gao, X., et al.: Retrieval-augmented generation for large language models: A survey. arXiv preprint arXiv:2312.10997 (2024)
10. Yan, G., Zhang, Y., Cai, P., et al.: HetaRAG: Hybrid deep retrieval-augmented generation across heterogeneous data stores. arXiv preprint arXiv:2509.21336 (2025)
11. Nogueira, R., Cho, K.: Passage re-ranking with BERT. arXiv preprint arXiv:1901.04085 (2020)
12. Rumble, T., Gazd'ik, Z., Zarrin, J., Ahluwalia, J.: Comparative analysis of neural retriever-reranker pipelines for retrieval-augmented generation over knowledge graphs in e-commerce applications. arXiv preprint arXiv:2602.22219 (2025)
13. Wang, L., Yang, N., Huang, X., et al.: Multilingual E5 text embeddings: A technical report. arXiv preprint arXiv:2402.05672 (2024)
14. Qdrant: Qdrant vector search engine. Software documentation (2024). Available at: https://qdrant.tech/. Accessed 2 June 2026
15. Aquino, S., Myriel, D.: A complete guide to filtering in vector search. Qdrant technical article (2024). Available at: https://qdrant.tech/articles/ vector-search-filtering/. Accessed 2 June 2026
16. Yang, A., Li, A., Yang, B., et al.: Qwen3 technical report. arXiv preprint arXiv:2505.09388 (2025)
17. BAAI: BGE reranker documentation. Technical documentation (2024). Available at: https://bge-model.com/tutorial/5_Reranking/5.2.html. Accessed 2 June 2026
18. Kikani, D., Ramchandani, S.M.: The impact of AI chatbots on customer service: Efficiency vs. human touch. International Journal for Research Trends and Innovation 10(3), 265–269 (2025). Available at: https://www.ijrti.org/papers/ IJRTI2503134.pdf. Accessed 2 June 2026
19. Karpukhin, V., Oguz, B., Min, S., et al.: Dense passage retrieval for open-domain question answering. arXiv preprint arXiv:2004.04906 (2020)
20. Gao, Y., Xiong, Y., Wang, M., Wang, H.: Modular RAG: Transforming RAG systems into LEGO-like reconfigurable frameworks. arXiv preprint arXiv:2407.21059 (2024)

21. Asai, A., Wu, Z., Wang, Y., Sil, A., Hajishirzi, H.: Self-RAG: Learning to retrieve, generate, and critique through self-reflection. arXiv preprint arXiv:2310.11511 (2023)
22. Jeong, S., Baek, J., Cho, S., Hwang, S.J., Park, J.C.: Adaptive-RAG: Learning to adapt retrieval-augmented large language models through question complexity. arXiv preprint arXiv:2403.14403 (2024)
23. Ling, C., Zhao, X., Lu, J., et al.: Domain specialization as the key to make large language models disruptive: A comprehensive survey. arXiv preprint arXiv:2305.18703 (2024)
24. Xu, J., Li, Z., Chen, W., et al.: On-device language models: A comprehensive review. arXiv preprint arXiv:2409.00088 (2024)
25. OpenAI: Whisper: Robust speech recognition. Software / GitHub repository (2022). Available at: https://github.com/openai/whisper. Accessed 2 June 2026

Original Research Article *https://doi.org/10.5281/zenodo.23472423*

Transformation of Unsustainable Structures into Sustainable Structures Using the BDPSS Formalism

Ana Trujić[1[0009-0009-8874-6707]], Dragica Ljubisavljević[1[0009-0002-7647-933X]], Milica Bačić[1[0009-0009-6802-5080]], Miloš Milić[1[0000-0002-2521-7607]], Dušan Savić[1[0000-0003-1726-219X]], Ilija Antović[1[0000-0002-2673-6628]] and Siniša Vlajić[1[0000-0003-3830-2170]]

[1] Faculty of Organizational Sciences, University of Belgrade, 11000 Belgrade, Serbia
{ana.trujic, dragica.ljubisavljevic, milica.bacic, milos.milic, dusan.savic, ilija.antovic, sinisa.vlajic}@fon.bg.ac.rs

Abstract. This work focuses on the identification of unsustainable structures in software systems and their transformation into sustainable structures. The aim of this study is to define a conceptual transformation model that translates a problem structure into a solution structure of a software system using the Basic Design Pattern Structure of the Solution (BDPSS) formalism, which is an element of the Universal Structural–Semantic Model of Design Patterns (USSMDP) framework. This study describes how different roles (Client, Abstract Server, and Concrete Server) and class responsibilities are identified based on the relations between classes. It also explains how to identify the unsustainable parts of the structure in which classes directly depend on classes that implement concrete functionalities, which makes software maintenance and evolution more difficult. Based on this, a conceptual transformation model is proposed for refactoring the initial structure, which contains unsustainable parts, into a new structure that is easier to upgrade and maintain. A key contribution of this work is a conceptual transformation model that is not limited to specific software structures and can be used for the automated analysis and improvement of software solution structures.

Keywords: Model Transformation, Sustainable Software Structure, Software Design Patterns, USSMDP, BDPSS.

1 Introduction

At the beginning of software system development, it is important to define the structure properly, because software quality largely depends on design, and errors introduced in the early development phases directly affect later changes and improvements to the system [1, 2, 3]. The structure of a software system refers to the set of classes, interfaces, abstract classes, and the relations between them. It forms the basis for software solution development and should be analyzed before the implementation phase [4, 5, 6, 7]. The structure of a software system can be represented by program code, UML diagrams, ER diagrams, or other notations [8, 9]. Often, diagram-drawing tools like PlantUML are used for this purpose as they can use program code or their own language as input and draw these diagrams automatically [10]. In this work, a UML class diagram is used to present the classes and their relationships.

Although UML class diagrams are a standard way of modeling software systems, they do not provide rules for developing a high-quality model, nor do they guarantee that the defined structure will be easy to maintain [8]. In practice, classes often directly depend on concrete implementations, which reduces system flexibility [11]. Such structures, in which the addition or change of concrete classes affects existing system elements, can be considered unsustainable. In contrast, sustainable structures are based on abstraction and allow the system to be extended without changing the client logic [12, 13]. In this context, sustainable structures follow the principle of programming to an interface, not an implementation [13, 14].

Introducing abstraction is one way to reduce class dependencies, because it allows the client to access concrete functionality through abstract elements instead of using it directly [2, 13]. Accordingly, this study proposes a conceptual transformation model for translating unsustainable structures into sustainable software solution structures. The study uses the *Universal Structural–Semantic Model of Design Patterns* (USSMDP), which is a framework for describing the structure and behavior of design patterns. The structure of a design pattern is described by the *Basic Design Pattern Structure of the Solution* (BDPSS), while the *Basic Design Pattern Semantic Module* (BDPSM) is used to describe the semantics of design patterns [11]. As this study focused on unsustainable and sustainable structures, the BDPSS element of the USSMDP framework was used. Although the BDPSS was introduced to describe the structure of design patterns, it can also be used to describe other types of software system structures. In this study, BDPSS is used as an intermediate representation for identifying and describing different structures in UML class diagrams. In this way, any structure can be formally represented, analyzed, and transformed from a problem structure (unsustainable structure) into a solution structure (sustainable structure). The contribution of this study is a general conceptual model for transforming a problem structure into a solution structure, based on the BDPSS element, which provides a basis for developing tools for the automated analysis and transformation of software system structures.

After the introductory section, the paper presents an overview of our previous papers on which this research is based, as well as related research by other authors. Section 3 presents and explains the conceptual framework for transforming unsustainable structures into sustainable. Section 4 describes the Basic Design Pattern Structure of the Solution (BDPSS) element, defined within the Universal Structural–Semantic Model of Design Patterns framework (USSMDP) [11]. Section 5 provides an example of the application of the conceptual transformation model. Finally, Section 6 presents the conclusions.

2 Related Work

This study builds on our previous work in which the Universal Structural–Semantic Model of Design Patterns was introduced [11]. The USSMDP is a framework used to describe the structure and behavior of design patterns. The USSMDP comprises two separate elements: the Basic Design Pattern Structure of the Solution (BDPSS) and the Basic Design Pattern Semantic Module (BDPSM). The BDPSS describes the structure

of design patterns, while the semantics of design patterns are described through the BDPSM. One of the important results of the previous paper is that BDPSS was defined as a minimal and formal representation of design patterns, independent of implementation and modeling methods or tools. This established a basis for further research and the development of tools for automatic structure identification, analysis, and transformation within software solutions. In this direction, this paper presents a conceptual framework for transforming unsustainable structures into sustainable software solution structures.

A particularly relevant reference for this study is the work by Vlajić et al., *The General Form of GoF Design Patterns* [12], in which design patterns are viewed as a way to transform a *problem structure* (an unsustainable structure) into a *solution structure* (a sustainable structure). In that paper, the problem structure and solution structure of design patterns were described through the following elements: *Client, Abstract Server*, and *Concrete Server*, where the transformation from the problem structure to the solution structure was achieved by introducing an abstraction between the Client and Concrete Servers [12]. This approach was used in this study to identify the unsustainable *Basic Design Pattern Structure of the Problem* (BDPSSP) and transform it into a sustainable BDPSS structure.

Three relevant groups of papers have been identified in the scientific literature. In these studies, UML diagrams were used because UML is a widely accepted notation for modeling software systems. The first group consists of papers that consider UML as a standard language for modeling software systems and analyze how the quality of UML diagrams affects model maintainability, understandability, and modifiability [5, 6, 15].

The second group consists of papers dealing with UML model refactoring and transformation [16, 17]. Sunyé et al. [16] presented ways of refactoring UML models and showed that transformation can be performed at the model level, not only at the program code level. Marković and Baar [17] formalized the refactoring of UML class diagrams with Object Constraint Language (OCL) constraints, where refactoring rules are described by a formalism inspired by graph grammars and implemented as QVT (Query/View/Transformation) transformations. These studies define model refactoring approaches and transformation rules, but do not propose a general conceptual framework for identifying unsustainable structures in UML diagrams.

The third group of relevant papers focuses on identifying the unsustainable structures within UML models. Kim and Lu [18] propose an approach for identifying the presence of design patterns in UML models by applying logic programming. Similarly, Rahimi et al. [19] developed a tool to identify static design patterns in UML class diagrams. Haendler [20] considered the use of UML diagrams to identify software design smells, that is, indicators of poor software design, and showed that problematic software structures can be recognized from UML diagrams. These studies confirm that unsustainable structures can be identified within UML models, but they do not propose a model that would enable the improvement of critical parts of the system, whereas this study proposes such an approach.

3 Conceptual Model for Transforming Unsustainable Structures into Sustainable Structures

This study proposes a conceptual model for transforming unsustainable software system structures into sustainable structures using the BDPSS element of the USSMDP framework. The transformation is achieved by introducing abstraction to avoid direct class dependencies and obtain a sustainable structure [13].

The transformation process consists of the following steps (see Fig. 1):

1. Definition of the initial structure representing the input model
2. Translation of the input model into a BDPSS representation
3. Identification of unsustainable parts of the structure
4. Transformation of the unsustainable structure into a sustainable structure
5. Obtaining the transformed structure representing the output model

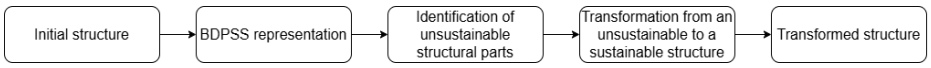


Fig. 1. Conceptual Model for Transforming Unsustainable Parts of the Software Structure

The following part of this section presents a detailed description of the steps of the proposed conceptual model for transforming software system structures.

3.1 Input Model

In general, the model input may be program code or different notations used to describe the system structure. In this study, the input to the transformation model is provided as a UML class diagram. A UML class diagram represents the static structure of a system through a set of classes, abstract classes, interfaces, and the relationships between them, such as inheritance, realization, dependency, aggregation, and composition [4]. These relations can be used to determine the relationships between classes in the structure and to identify the roles of their elements (Client, Abstract Server, and Concrete Server) [12, 13].

3.2 Translation of the Input Model into a BDPSS Representation

Based on the UML class diagram, the elements, their roles, and the relationships between them were identified. In this step, BDPSS structural forms are identified from the diagram, and the result is a BDPSS structure that describes the system participants and the relationships between them. The BDPSS allows parts of a UML class diagram to be represented by the element 3-tuple (Client, Abstract Server, Concrete Server) [11]. Using BDPSS structures, it is possible to clearly determine whether the structure contains an abstraction that separates the Client from concrete implementations or whether a direct dependency exists.

3.3 Identification of an Unsustainable Structure

In this step, a structure is identified in which the Client directly depends on one or more concrete classes. This structure is considered a problem structure because a change in a concrete implementation may require a change in the Client class. In such a structure, there is no abstraction to provide system flexibility and maintainability. Based on [12], a rule was identified that can be used as a detection rule within the proposed transformation model: *if changing or adding a concrete class requires a change in the Client, the structure is considered unsustainable*.

Within the BDPSS element of the USSMDP framework, the problem structure is viewed as a structure in which the Abstract Server element does not exist, that is, it has the value null (for example, BDPSS(Client, null, ConcreteServer)) [11].

3.4 Transformation of the Unsustainable Structure into a Sustainable Structure – Transformation Rule

The solution structure of a software system is a structure in which the Client does not directly depend on concrete classes, but on an abstraction that defines the required functionality. A sustainable structure is obtained from an unsustainable one by introducing an abstraction between the Client and concrete classes. Concrete classes implement abstract functionality and can be changed without affecting the Client. This idea is related to the software design principles of *inversion of control* and *dependency injection*, which are also based on reducing the direct dependencies between classes [13]. In the following formulas, *Cl* denotes the Client, *AS* denotes the Abstract Server, and *CS* denotes the Concrete Server.

In the BDPSS formalism, an unsustainable structure (problem structure) of a software system is defined as [11]:

BDPSS_© (Cl, null, CS) = (Cl © CS1) + ... + (Cl © CSn), © = (has, using), n – number of CS (1)

In the BDPSS formalism, a sustainable structure (solution structure) of a software system is defined as follows [11]:

BDPSS_© (Cl, AS, CS) = (Cl © AS) + (CS1 is AS) + ... + (CSn is AS), © = (has, using), n – number of CS (2)

The transformation rule involves introducing an Abstract Server between the Client and concrete classes. In this way, the problem structure is transformed into a solution structure, and the Client's direct dependency on concrete classes is replaced by a dependency on an abstraction. This transformation can be represented as follows:

BDPSS (Cl, null, CS) → BDPSS (Cl, AS, CS) (3)

3.5 Obtaining the Transformed Structure Representing the Output Model

The output of the model is a transformed BDPSS structure, which is then translated into a UML class diagram or another output representation of the structure. The result is a sustainable structure in which the Client depends on an abstraction, while the concrete classes implement abstract functionality. However, as software systems evolve and are extended over time, unsustainable parts of the structure may reappear because of the addition of new elements or modifications to existing ones. Therefore, the proposed transformation model can be viewed as an iterative process that may be repeated when needed to preserve the sustainability of the software system structure.

4 BDPSS Element of the USSMDP Framework

The conceptual model proposed in this study is based on the Basic Design Pattern Structure of the Solution (BDPSS), which is defined within the Universal Structural–Semantic Model of Design Patterns framework [11]. The BDPSS is a formal, notation-independent model that enables the formal description of software structures. A detailed definition of BDPSS and examples of its application to three design patterns (Abstract Factory, Composite, and Observer) are provided in our previous work [11].

4.1 BDPSS Elements

The structure of a design pattern is described by the element 3-tuple: Client, Abstract Server, and Concrete Server, as follows:

$$\text{BDPSS} = (\text{Client, Abstract Server, Concrete Server}) \tag{4}$$

Each element of this 3-tuple represents a role in the proposed structure. The Client (Cl) uses a certain functionality. The Abstract Server (AS) defines the abstract functionality used by the Client, while the Concrete Server (CS) provides its concrete implementation. Within a BDPSS structure, the Client can be associated with only one Abstract Server, while one or more Concrete Servers can be derived from that Abstract Server.

A set of possible relationships is also defined between these structural elements. The relations between the Client and the Abstract Server are *has* and *using*, while the relation between the Concrete Server and the Abstract Server is *is.* The *has* relation denotes aggregation or composition, the *using* relation denotes dependency, and the *is* relation denotes inheritance or realization.

4.2 Overview of Different Structural Forms

Based on the presence or absence of individual elements of the 3-tuple and the type of relations between them, several structural forms can be defined. Some of the structural forms described in [11] are presented below.

The simplest form occurs when the structure contains only the Client, without an Abstract Server or Concrete Server. Such a structure is represented as *BDPSS(Cl, null,*

null) [11]. A particularly important form is the one in which the Client directly uses one or more Concrete Servers without the presence of an Abstract Server. This form is written as *BDPSS(Cl, null, CS), CS = (CS1, CS2, ..., CSn)* and represents the *general problem structure of design patterns* [11, 12] (see Fig. 2). By adding an Abstract Server to this structure, the solution structure of a design pattern is obtained, and it is represented by the element 3-tuple *BDPSS(Cl, AS, CS), CS = (CS1, CS2, ..., CSn)* [11, 12] (see Fig. 3).

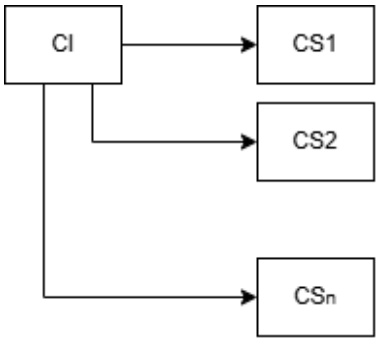


Fig. 2. Class Diagram of an Unsustainable Software System Structure [11]

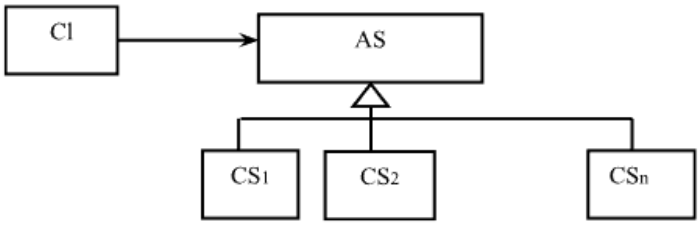


Fig. 3. Class Diagram of a Sustainable Software System Structure [11]

In addition to defining individual forms (eight in total, as described in [11]), the BDPSS allows more complex structures to be represented as compositions of multiple basic BDPSS forms. In this way, different structures can be described in a uniform manner in order to support their analysis and transformation. In this study, the BDPSS formalism was used as an intermediate step to identify problem structures and define their transformations.

5 Case Study of Applying the Transformation Model

The following section presents a case study of applying the transformation model to the Abstract Factory pattern structure, with the goal of identifying and eliminating unsustainable structures using the BDPSS formalism. Although the Abstract Factory is widely recognized as a sustainable solution that reduces the coupling between concrete classes and the Client, its structure is not sustainable when analyzed using the BDPSS

formalism. In particular, dependencies related to product families remain unsustainable because the Client is still implicitly tied to the structure of abstract products. This limitation is especially noticeable when a new abstract product type is introduced, since the Client class then needs to be modified to support the added abstraction.

The general structure of the Abstract Factory pattern is shown in Fig. 4. This UML class diagram represents the input to the transformation model.

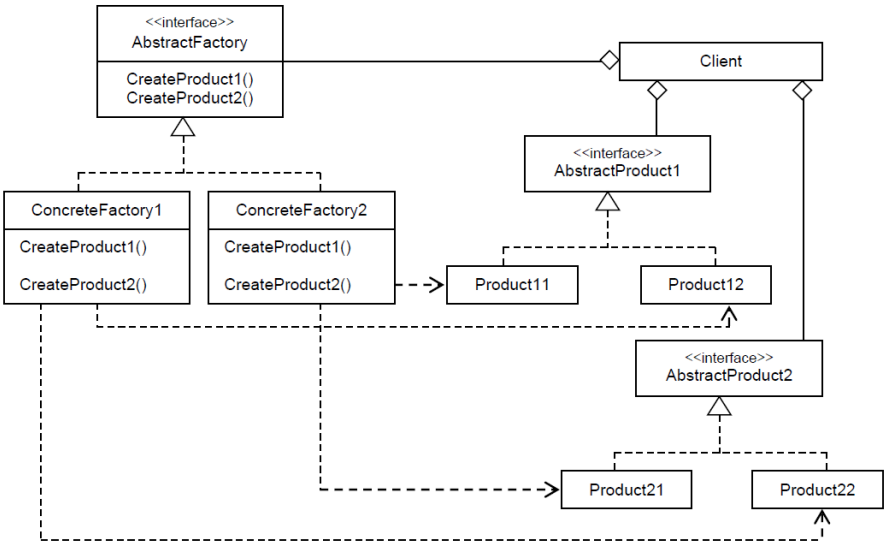


Fig. 4. Example of an Unsustainable Structure – General Structure of the Abstract Factory pattern [13, 14]

The class diagram is then translated into the structural forms of the BDPSS and can be represented as a composition of several structures:
*AbstractFactory = BDPSS(Client, AbstractFactory, ConcreteFactory) + BDPSS(Client, null, *≠AbstractProduct) + BDPSS(Client, AbstractProduct1, Product1) + BDPSS(Client, AbstractProduct2, Product2) + BDPSS_using(ConcreteFactory1, null, *≠Product_cf1) + BDPSS_using(ConcreteFactory2, null, *≠Product_cf2),* where ConcreteFactory = (ConcreteFactory1, ConcreteFactory2), AbstractProduct = (AbstractProduct1, AbstractProduct2), Product1 = (Product11, Product12) , Product2 = (Product21, Product22), Product_cf1 = (Product12, Product22), Product_cf2 = (Product11, Product21).

The identified structure reveals that several dependencies are not mediated by an Abstract Server, as indicated by the presence of null in forms such as *BDPSS(Client, null, *≠AbstractProduct), BDPSS_using(ConcreteFactory1, null, *≠Product_cf1),* and *BDPSS_using(ConcreteFactory2, null, *≠Product_cf2).* These unsustainable structures reveal that the Client and concrete factories remain dependent on the set of abstract products. Consequently, any change in product types (such as addition or modification) will cause changes in the Client class and the factories.

By applying the proposed transformation model, an additional abstraction layer is introduced over the abstract products, denoted as *AbstractAbstractProduct.* This transformation removes the direct dependency of the Client on the set of abstract products. As a result, changes in product types do not propagate to the Client. The *AbstractAbstractProduct* is introduced as an abstraction of interfaces *AbstractProduct1* and *AbstractProduct2*. Client gets a reference to *AbstractAbstractProduct*, breaking the connections to *AbstractProduct1* and *AbstractProduct2*. In this way, the structure *BPDSS(Client, *≠AbstractAbstractProduct, AbstractProduct)* is formed, according to the BDPSS formalism.

The symbol *≠ indicates that a concrete factory may use multiple abstract products, with the constraint that these products must belong to different product groups. In the presented example, the value of symbol * is equal to 2 because the Abstract Factory pattern contains two abstract product groups (*AbstractProduct1* and *AbstractProduct2*). However, in the general case, this value may represent any number of abstract product groups. The symbol *≠ means that a single factory instance cannot use multiple products from the same product category (e.g., *Product11* and *Product12*), but only products from distinct categories to form a valid complex product (e.g., *Product11* and *Product21*).

The resulting BDPSS structure can be expressed as:

*AbstractFactory = BDPSS(Client, AbstractFactory, ConcreteFactory)+BDPSS(Client, *≠AbstractAbstractProduct, AbstractProduct) + BDPSS(null, AbstractProduct1, Product1) + BDPSS(null, AbstractProduct2, Prod-uct2) + BDPSS_using(ConcreteFactory1, *≠AbstractAbstractProduct, AbstractProduct) + BDPSS_using(ConcreteFactory2, *≠AbstractAbstractProduct, AbstractProduct),* where ConcreteFactory = (ConcreteFactory1, ConcreteFactory2), AbstractProduct = (AbstractProduct1, AbstractProduct2), Product1 = (Product11, Product12), Product2 = (Product21, Product22).

In this transformed structure, all dependencies are mediated through *AbstractAbstractProduct*, and the previously identified unsustainable structures containing a null value are eliminated. Consequently, the Client is no longer dependent on a specific type of abstract product, and the overall structure can be considered sustainable.

The general transformation rule *BDPSS(Cl, null, CS) → BDPSS(Cl, AS, CS)* can be instantiated in the Abstract Factory structure through the following three transformations:

- *BDPSS(Client, null, *≠AbstractProduct) → BDPSS(Client, *≠AbstractAbstractProduct, AbstractProduct)*
- *BDPSS_using(ConcreteFactory1, null, *≠Product_cf1) → BDPSS_using(ConcreteFactory1, *≠AbstractAbstractProduct, AbstractProduct)*
- *BDPSS_using(ConcreteFactory2, null, *≠Product_cf2) → BDPSS_using(ConcreteFactory2, *≠AbstractAbstractProduct, AbstractProduct)*

The result of these transformations is represented by the transformed UML class diagram shown in Fig. 5.

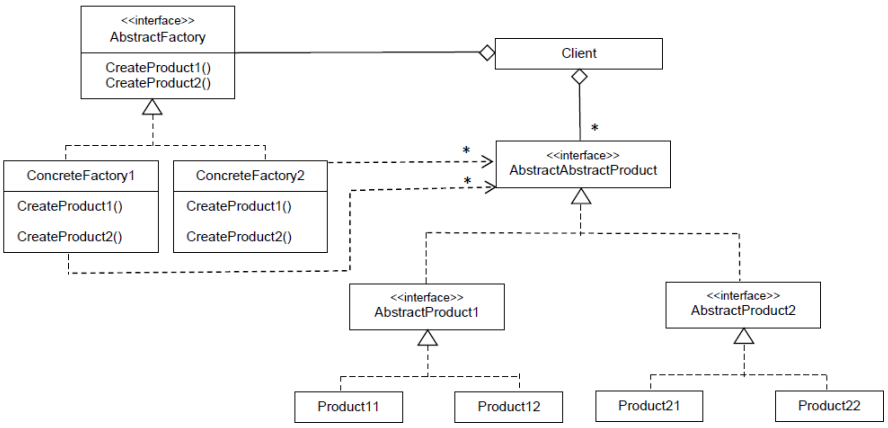


Fig. 5. Example of a Sustainable Structure – Transformed Structure of the Abstract Factory pattern

As a further improvement of the transformed structure (shown in Fig. 6), the dependency represented by the *using* relation is moved from the individual concrete factories to the *AbstractFactory*. In the structure shown in Fig. 5, each concrete factory separately depended on *AbstractAbstractProduct*. By introducing the relation *BDPSS_using(AbstractFactory, *≠AbstractAbstractProduct, AbstractProduct)*, the dependency is centralized within the *AbstractFactory* abstraction. The symbol *≠ has the same meaning as explained previously. The final BDPSS structure of the Abstract Factory pattern can be represented as:
*AbstractFactory = BDPSS(Client, AbstractFactory, ConcreteFactory) + BDPSS(Client, *≠AbstractAbstractProduct, AbstractProduct) + BDPSS(null, AbstractProduct1, Product1) + BDPSS(null, AbstractProduct2, Product2) + BDPSS_using(AbstractFactory, *≠AbstractAbstractProduct, AbstractProduct),* where ConcreteFactory = (ConcreteFactory1, ConcreteFactory2), AbstractProduct = (AbstractProduct1, AbstractProduct2), Product1 = (Product11, Product12), Product2 = (Product21, Product22).

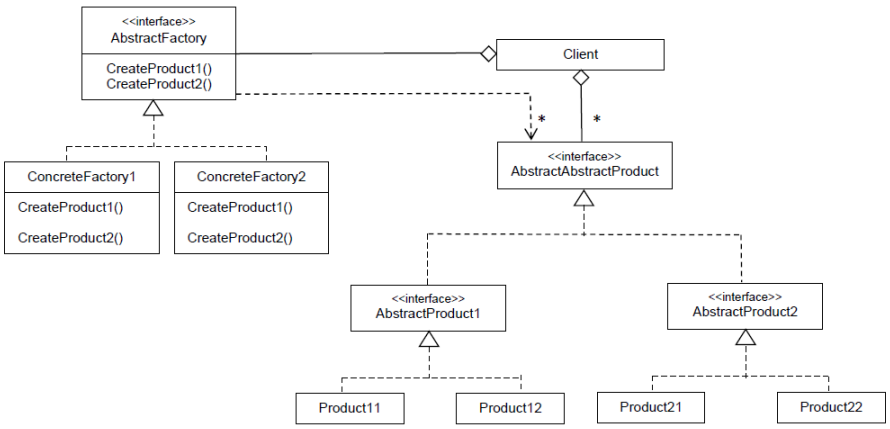


Fig. 6. Final generalized sustainable structure of the Abstract Factory pattern

This case study shows that the standard Abstract Factory structure can be considered only partially sustainable within the BDPSS formalism, since changes in the set of abstract products may still affect the Client and factories. Therefore, the transformation process is inherently iterative, where the output of one step can become the input for the next step, until all unsustainable structures are eliminated. Although the proposed transformation model reduces direct dependencies and improves extensibility, the introduction of additional abstraction layers may increase structural complexity and reduce model simplicity. As the proposed formalism is still evolving, future work will focus on refining the transformation rules and identifying additional structural cases.

6 Conclusion

This study considered the problem of identifying unsustainable structures and the way such structures can be transformed into sustainable software system structures. The term unsustainable structure was used for a structure in which one class directly depends on the functionality of other classes, so that changing or adding new functionality may cause changes in other classes. To address this problem, this study proposes a conceptual transformation model in which an unsustainable structure is translated into a sustainable structure by introducing an abstraction (Abstract Server) between the class that uses the functionality (Client) and the classes that contain the implementation of concrete functionality (Concrete Server). This transformation is represented by the formula $BDPSS(Cl, null, CS) \rightarrow BDPSS(Cl, AS, CS)$. The model uses the BDPSS as an intermediate representation of the structure, which is considered in this work through UML class diagrams. Its application enables a more formal representation of the model and the identification of problematic structures.

Future research will include the implementation of a tool that would automatically identify unsustainable structures in UML class diagrams or program code and transform them into sustainable structures. The tool would be based on the described conceptual

transformation model. In addition, future work will include an evaluation of the developed tool on different structural examples, with the aim of verifying the accuracy of structure detection and transformation.

Acknowledgments. The authors used AI-based tools only for language refinement, proofreading, and minor copyediting of the manuscript. The scientific content, research idea, dataset, interpretation of results, and conclusions were prepared by the authors.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Vlajić, S.: Projektovanje softvera (skripta) [Software Design (course script)]. University of Belgrade – Faculty of Organizational Sciences, Belgrade (2015).
2. Alreffaee, T.R., Dabdawb, M.M.A., Taha, D.B.: Measuring extensibility quality attribute using object-oriented design metrics. TELKOMNIKA Telecommunication, Computing, Electronics and Control 19(5), 1507–1516 (2021).
3. Vir, R., Dhillon, P., Dhillon, J.: Fuzzy logic approach to forecast extensibility in object-oriented design using an integrated model. International Journal of Computer Applications 94(9), 17–22 (2014).
4. Booch, G., Rumbaugh, J., Jacobson, I.: The Unified Modeling Language User Guide. Addison-Wesley, Reading (1999).
5. Genero, M., Manso, E., Visaggio, A., Canfora, G., Piattini, M.: Building measure-based prediction models for UML class diagram maintainability. Empirical Software Engineering 12, 517–549 (2007).
6. Wang, Y.: Formal Description of the UML Architecture and Extendibility. L'Objet 6(3), 1–20 (2000).
7. OMG: Unified Modeling Language (UML), Version 2.5.1. Object Management Group (2017).
8. Karasneh, B., Stikkolorum, D.R., Larios, E., Chaudron, M.R.V.: Quality Assessment of UML Class Diagrams: A Study Comparing Experts and Students. In: Proceedings of the MODELS Educators Symposium 2015, pp. 55–67, Ottawa, Canada (2015).
9. Chen, P. P.-S.: The Entity-Relationship Model—Toward a Unified View of Data. ACM Transactions on Database Systems 1(1), 9–36 (1976).
10. Roques, A.: PlantUML Language Reference Guide. PlantUML. Available at: https://plantuml.com (accessed April 15, 2026).
11. Vlajić, S., Trujić, A., Milić, M., Savić, D., Poledica, A.: Universal Structural-Semantic Model of Design Patterns. In: 2026 30th International Conference on Information Technology (IT), Žabljak, Montenegro. IEEE (2026). DOI: 10.1109/IT67293.2026.11435724.
12. Vlajić, S., Stanojević, V., Savić, D., Milić, M., Antović, I., Lazarević, S.: The General Form of GoF Design Patterns. World of Computer Science and Information Technology Journal 6(2), 12–20 (2016).
13. Vlajić, S.: *Softverski paterni* [Software Patterns]. Zlatni presek, Belgrade (2014).
14. Gamma, E., Helm, R., Johnson, R., Vlissides, J.: Design Patterns: Elements of Reusable Object-Oriented Software. Addison-Wesley, Boston (1995).
15. Bernardi, S., Merseguer, J., Petriu, D. C.: Dependability Modeling and Analysis of Software Systems Specified with UML. ACM Computing Surveys 45(1), Article 2 (2012).

16. Sunyé, G., Pollet, D., Le Traon, Y., Jézéquel, J.-M.: Refactoring UML Models. In: Gogolla, M., Kobryn, C. (eds.) UML 2001 – The Unified Modeling Language. Modeling Languages, Concepts, and Tools, Lecture Notes in Computer Science, vol. 2185, pp. 134–148. Springer, Berlin, Heidelberg (2001).
17. Marković, S., Baar, T.: Refactoring OCL Annotated UML Class Diagrams. Software and Systems Modeling 7(1), 25–47 (2008).
18. Kim, D.-K., Lu, L.: Inference of Design Pattern Instances in UML Models via Logic Programming. In: 11th IEEE International Conference on Engineering of Complex Computer Systems (ICECCS'06), pp. 9–18. IEEE (2006).
19. van Doorn, E., Stuurman, S., van Eekelen, M.: Static Detection of Design Patterns in Class Diagrams. In: CSERC '19: Proceedings of the 8th Computer Science Education Research Conference, pp. 79–88. ACM (2020). DOI: 10.1145/3375258.3375268.
20. Haendler, T.: On Using UML Diagrams to Identify and Assess Software Design Smells. In: Proceedings of the 13th International Conference on Software Technologies (ICSOFT), pp. 413–421. SciTePress, Porto (2018).

Original Research Article *https://doi.org/10.5281/zenodo.23472499*

Application of Feature Models in Modular Software Development

Isidora Vidojević[1][0009-0000-4000-2332], Vladimir Belča[1][0000-0003-0947-335X] and Đorđe Nestorović[1][0009-0008-6013-3709]

1 University of Belgrade, Faculty of Organizational Sciences

Corresponding email: isidora.vidojevic@fon.bg.ac.rs

Abstract. This paper explores the application of feature models as an approach for managing variability in the development of modular software systems. The study is motivated by the need to design flexible software solutions that can be adapted to different user requirements while preserving a unified system archi- tecture. The focus of the paper is the application of feature models in the context of con- figurable web portals for educational institutions. The objective is to demonstrate how this approach enables systematic definition and management of both com- mon and variable functionalities. The proposed solution applies established feature modeling concepts to represent relationships and constraints between features and to define an integrated model that unifies different types of educational institution portals within a single struc- ture. The results indicate that the proposed approach is a high level of flexibility, re- usability and configurability. Although managing many features introduces ad- ditional complexity, feature models represent a suitable framework for the devel- opment of modular software solutions.

Keywords: Feature model, Variability, Modular Software.

1 Introduction

Modern software systems increasingly require a high degree of adaptability which has led to a wider adoption of modular approaches in their development. Modularity allows a system to be structured as a set of components that are logically independent but still interconnected, making it easier to maintain, extend, and reuse. However, as the number of modules grows, the challenge of managing variability arises, i.e., defining and con- trolling different combinations of functionalities within the system. This challenge is particularly evident in the domain of educational portals, which represent complex information systems with many functionalities. Within such sys- tems, there is a need to simultaneously preserve a common core while supporting vari- ous specific requirements, highlighting the importance of a formal approach to varia- bility management.

The subject of this paper is the application of feature models as a mechanism for managing variability in the development of modular software solutions for educational

portals. The main problem addressed in this paper concerns the modeling of variability in a system that needs to support multiple configurations while maintaining a unified archi- tecture. The goal of this research is to demonstrate how feature models can serve as a foun- dation for defining a configurable structure of a software system intended for different types of educational institutions. Based on the defined problem, the paper aims to answer the following question: How can feature models be applied for systematic variability management in the develop- ment of modular software solutions for educational portals of different types of institu- tions? The paper is structured as follows. Section 2 introduces the basic concepts of feature models, including their structure, types of features, and relationships. Section 3 dis- cusses the application of feature models in modular software development, with a focus on variability and challenges in educational portals. Section 4 presents the feature model of a high school portal, while Section 5 introduces the integrated feature model that unifies different types of educational institutions. Section 6 provides a discussion of the proposed approach, and Section 7 concludes the paper.

2 Feature Models

2.1 Basic Concepts

Feature models represent a structured approach for describing and managing variability in software systems. They are most used in the context of Software Product Lines, where they enable systematic representation of all possible variants of a software product and their relationships [1], [2]. Their primary role is to provide a clear and structured representation of system functionalities, thereby simplifying software configuration and adaptation to different user requirements [3].

A feature model is typically represented as a hierarchical diagram, where the root node represents the system or domain under consideration, and other nodes represent its features [4]. A feature can be defined as a property or functionality of a system that is relevant to stakeholders and used to distinguish between different product variants [5]. This structure enables a clear separation of common and variable parts of the system, which is essential for managing complexity.

2.2 Types of Features and Relationships

Within feature models, different types of features and relationships are distinguished. The basic classification includes mandatory and optional features [4]. Mandatory fea- tures are present in all configurations when their parent feature is selected, whereas optional features may be included or excluded depending on specific requirements.

In addition, feature models support grouped features that allow selection among multiple options. In this context, OR groups enable the selection of one or more features, while alternative (XOR) groups allow the selection of exactly one feature [4].

2.3 Constraints in Feature Models

Beyond hierarchical relationships, feature models also allow the definition of additional constraints that describe dependencies between features. The most common constraints are *requires* and *excludes*, which enable more precise definition of valid feature com- binations and prevent inconsistent configurations [4].

The feature modeling process involves identifying common and variable aspects of the system, organizing them into a hierarchical structure, and defining relationships and constraints between features [4]. This process is iterative and based on domain analysis, stakeholder requirements and existing systems.

Feature models represent a key mechanism for managing variability in complex software systems, as they enable clear definition of possible configurations and provide a foundation for the development of flexible software solutions [2], [3]. The basic concepts and relationships in feature models are illustrated in Fig. 1.

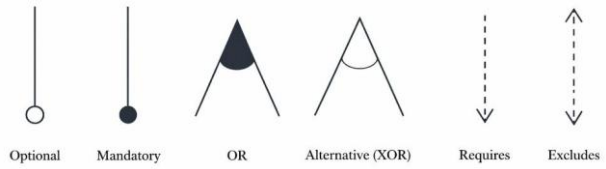


Fig. 1. Basic concepts and notation in feature models

3 Application of Feature Models in Modular Software Development

3.1 Modularity and Variability

The foundation of proposed solution lies in the principle of modularity, which involves dividing a system into mutually independent yet functionally connected components. In accordance with this principle, a software system is organized as a set of components that can be independently developed, maintained and reused, with clearly defined in- terfaces [6], [7]. This organization enables flexibility and adaptability to various user requirements.

In this paper, feature models are used as a mechanism for managing variability within a modular architecture. Features defined in the model are directly mapped to individual system modules, enabling software configuration according to different requirements.

3.2 Challenges in Educational Portals

In the domain of educational portals, there is a challenge in designing systems that sup- port different types of educational institutions, such as primary schools, high schools and universities. Although these systems share many common functionalities, such as managing news, documents, events and pages, their structure and requirements differ to some extent.

These differences necessitate the introduction of specific functionalities tailored to each level of education, which complicates the development of a unified solution. Developing separate systems for each case would lead to duplication of functionalities, increased maintenance costs and reduced reusability of components. This highlights the need for an approach to systematic variability management.

To address this challenge, an approach based on feature models is proposed, where the system is defined as a unified architecture with configurable modules. Common functionalities are modeled as mandatory features, while specific characteristics of different types of institutions are represented as variable features that can be included or excluded.

4 Feature Model of a High School Portal

As an illustration of the proposed approach, a feature model for a high school portal is defined. This model is derived as an extension of a base model that includes common functionalities of all educational institutions. The feature model of the high school portal is shown in Fig. 2 and Fig. 3.

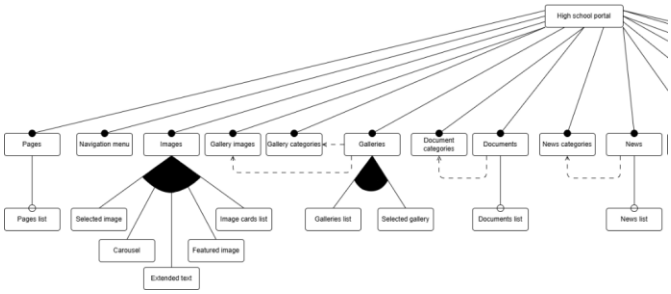


Fig. 2. Feature model of a high school portal (part 1)

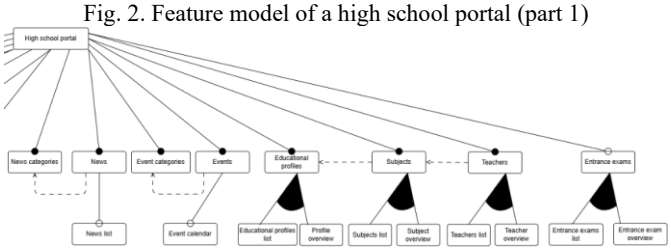


Fig. 3. Feature model of a high school portal (part 2)

As can be observed from Fig. 2 and Fig. 3, the model includes both common portal functionalities and specific features such as educational profiles and entrance exams, which are characteristic of high schools.

The structure of the feature model is organized as a hierarchical tree, whose root represents the educational institution portal. From the root node, the main functional areas of the system branch out, such as content management, users, media and educational components.

Each of these areas is further decomposed into features representing individual system modules. The model clearly distinguishes mandatory features, which are present in all system configurations, from variable features that can be included or excluded depending on the type of educational institution.

Additionally, certain groups of features are defined as alternative, meaning that only one option can be selected within the group. In this way, the model not only describes the system structure, but also specifies the constraints and rules that guide system configuration.

Compared to the base model, the high school model introduces features specific to this level of education. The key extension is the introduction of educational profiles, which represent the fundamental unit of curriculum organization in high schools. The model also includes features for representing entrance exams where applicable.

It is important to note that these features are not independent but are related to existing system elements. For example, subjects are considered within the context of educational profiles, ensuring logical consistency in representing the educational process.

This model represents one possible application of the proposed methodology. Using the same approach, models can be defined for primary schools and universities, maintaining a common structure while introducing level-specific characteristics.

5 Integrated Feature Model

To support the development of a unified portal architecture, an integrated feature model was defined by combining the individual feature models of primary schools, secondary schools, and higher education institutions. The integration process was based on a com- parative analysis of these models. First, features appearing in multiple models were identified as common features and incorporated into the integrated model. Features spe- cific to a single institution type were classified as variable features. The status of each feature was then re-evaluated within the integrated context. Features required across all institution types retained their mandatory status, while features relevant only to specific institution types were modeled as optional or alternative features. Finally, dependency constraints and exclusion rules were introduced to prevent invalid feature combinations and to ensure that valid configurations could be derived for each institution type while preserving consistency within the unified model. The integrated feature model is shown in Fig. 4 and Fig. 5.

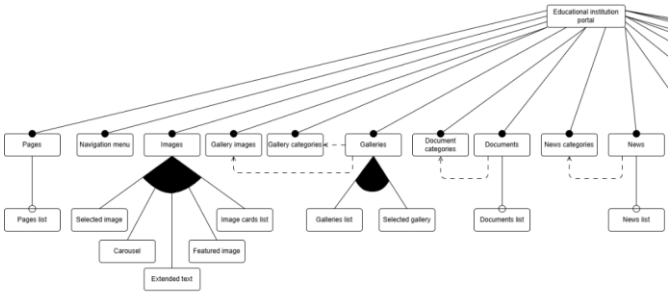


Fig. 4. Integrated feature model of educational institution portals (part 1)

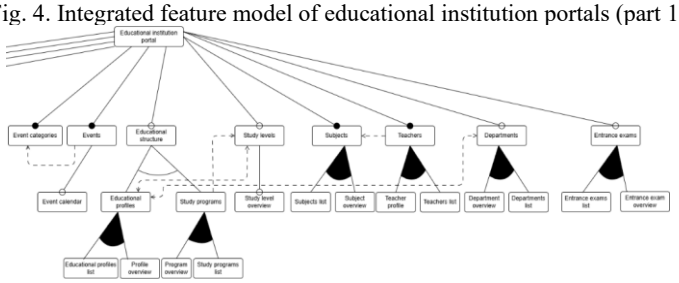


Fig. 5. Integrated feature model of educational institution portals (part 2)

As shown in Fig. 4 and Fig. 5, the integrated model unifies common and variable features, enabling system configuration according to the type of educational institution.

In this model, basic features common to all institution types retain the status of mandatory features, while specific features of individual education levels are modeled as optional or alternative features.

A particularly important aspect of the integrated model is the ability to define constraints that prevent inconsistent configurations. For example, features related to educational profiles and study programs are defined as mutually exclusive, ensuring that the system configuration corresponds to the specific type of institution.

This structure allows different system configurations to be derived from a single mode, ranging from simpler portals for primary schools to more complex systems intended for universities. At the same time, unified architecture is preserved, significantly simplifying development and maintenance.

6 Discussion

The proposed approach demonstrates that feature models represent an effective mech- anism for linking the concept of modularity with the practical implementation of soft- ware systems. Features defined in the model can be directly interpreted as system mod- ules, whose inclusion or exclusion determines the final application structure.

A key advantage of this approach is the ability to develop a unified architecture that supports different types of educational institutions. This avoids duplication of functionality, reduces maintenance costs, and ensures consistent system quality. Additionally,

the model enables easy system extension by introducing new features without disrupting the existing structure.

On the other hand, the integrated feature model introduces a certain level of complexity, particularly in defining and managing dependencies between features. An increase in the number of variable elements may affect model clarity and requires careful definition of constraints to avoid inconsistent configurations. To address this challenge, features can be organized into logical groups representing related functional areas of the system. Such grouping enables administrators to configure the system incrementally, focusing on one feature package at a time rather than managing the entire model simultaneously.

Although the paper has a conceptual focus, the proposed solution has been validated through the implementation of a prototype software system. The prototype was developed using the Payload CMS framework and the TypeScript programming language. Feature definitions, dependencies, and configuration constraints were stored and managed through the underlying database model, enabling the validation of feature selections and relationships.

Consistency is ensured through the hierarchical organization of features and the explicit definition of dependency rules. Child features are available only when their corresponding parent feature is enabled, while mutually exclusive and mandatory relationships are enforced through configuration constraints. In the prototype, these rules are validated dynamically through the administrative interface, preventing invalid feature combinations and ensuring that only consistent configurations can be stored and applied.

The implementation covers a representative subset of features that are essential for demonstrating the proposed approach. In particular, the prototype supports mandatory and optional features, mutually exclusive features, and conditional feature relationships in which child features become available only after the corresponding parent feature has been selected. These mechanisms enable the validation of feature dependencies and configuration constraints defined in the model.

Although the prototype currently implements only the key feature categories required to demonstrate the approach, the proposed architecture is not limited to this subset. The same implementation principles can be extended to support the complete feature model and additional feature types without requiring significant modifications to the underlying system architecture.

The scope of the proposed integrated model is limited to educational institution portals. During the analysis phase, three categories of institutions were identified - primary schools, secondary schools, and higher education institutions - and individual feature models were developed and integrated for these domains. Although the model was designed around these institution types, the approach is not restricted to them. Additional educational domains, such as preschool institutions, could be incorporated through further domain analysis and the definition of corresponding feature models.

A limitation of the proposed model is that not all domain-specific constraints can be expressed directly within the feature modeling notation and tool support used in this work. Certain relationships depend on the selected configuration and are therefore rep-

resented as conditional rules rather than explicit feature model constraints. For example, associations between subjects and educational profiles or study programs are applicable only in specific educational contexts and cannot be universally enforced within the integrated model. Consequently, some consistency rules must be documented and validated at the configuration level rather than being fully captured by the feature model itself.

7 Conclusion and future directions

This paper demonstrates that feature models represent an effective conceptual framework for managing variability in the development of modular software solutions in the educational domain. They enable modeling of both common and specific functionalities across different levels of education and their integration into a unified, configurable architecture.

The significance of the proposed approach lies in linking the concepts of modularity and variability modeling. Feature models enable clear separation of mandatory and variable system elements, as well as the definition of relationships and constraints between them, ensuring consistency of possible configurations.

From a software engineering perspective, the approach contributes to avoiding duplication of functionality and facilitates further system development. In the educational domain, it enables the creation of flexible portals adaptable to different types of institutions while maintaining a common foundation.

However, the increase in the number of features and their interdependencies may lead to more complex models, highlighting the need for careful design of structure and constraints to preserve clarity and usability.

Future research directions may include the development of approaches for automating the feature modeling process, reducing the need for manual model definition. Additionally, further work could explore the application of the proposed approach in other domains, as well as scalability issues of integrated feature models and methods for managing complex dependencies in systems with many variable elements.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Felfernig, A., Falkner, A., Benavides, D.: Feature Models: AI-Driven Design, Analysis and Applications. SpringerBriefs in Computer Science. Springer, Cham (2024). DOI: 10.1007/978-3-031-61874-1.
2. Nešić, D., Krüger, J., Stănciulescu, Ș., Berger, T.: Principles of Feature Modeling. In: Pro- ceedings of the 27th ACM Joint Meeting on European Software

Engineering Conference and Symposium on the Foundations of Software Engineering, pp. 62–73 (2019).
3. Benavides, D., Segura, S., Ruiz-Cortés, A.: Automated Analysis of Feature Models 20 Years Later: A Literature Review. Information Systems 35(6), 615–636 (2010).
4. Czarnecki, K.: Generative Programming: Principles and Techniques of Software Engineering Based on Automated Configuration and Fragment-Based Component Models. PhD thesis, Department of Computer Science and Automation, Technical University of Ilmenau, Ilmenau, Germany (1998).
5. Czarnecki, K., Helsen, S., Eisenecker, U.: Staged Configuration Using Feature Models. In: Software Product Lines: Third International Conference, SPLC 2004, Boston, MA, USA, August 30–September 2, 2004. Proceedings, pp. 266–283. Lecture Notes in Computer Sci- ence, vol. 3154. Springer, Berlin, Heidelberg (2004). DOI: 10.1007/978-3-540-28630-1_17.
6. Knoernschild, K.: Java Application Architecture: Modularity Patterns with Examples Using OSGi. Prentice Hall (2012).
7. Streitferdt, D., Philippow, I., Heller, C.: A Component Model for Applications Based on Feature Models. In: Proceedings of the 2nd Workshop on Software Variability Management—Software Product Families and Populations (SVM 2004) (2004).

THEMATIC SECTION

Blockchain Technology in Business and Information Systems

Resilient and Intelligent Organizations
in Times of Permanent Disruption

SYMORG 2026 • CONFERENCE PROCEEDINGS

Original Research Article *https://doi.org/10.5281/zenodo.23472551*

Cryptographic Model Chaining for Privacy-Preserving AI Ensembles and Cross-Chain Collaboration

Miloš Živadinović[1][0000−0002−0342−340X], Ivan Milenković[2][0000−0001−6252−8159], and Dejan Simić[2][0000−0002−0744−5411]

[1] AikBank, Belgrade, Serbia
https://aikbank.rs/
milos.zivadinovic@aikbank.rs
[2] Faculty of Organizational Sciences, Belgrade, Serbia
https://fon.bg.ac.rs/
{ivan.milenkovic, dejan.simic}@fon.bg.ac.rs

Abstract. The rapid proliferation of AI models across organizations has created a fundamental challenge between collaboration and protection of proprietary data. Organizations investing significant resources in model development are reluctant to share their assets without guarantees against theft, reverse engineering and unauthorized use. We propose a cryptographic framework for privacy-preserving AI model chaining, where model metadata such as training results and dataset descriptors are chained across versions and related models without exposing underlying weights or architectures. Each chain link provides verifiable proof of domain expertise through static cryptographic commitments, while zero-knowledge proofs (ZKPs) enable dynamic verification that a claimed model performs as its metadata asserts. For ensemble construction, we introduce cross-chain linking, a ladder-structure that connects two or more parent chains into a shared descendant that inherits and extends their metadata under threshold encryption keys. This structure enables selective, permission-controlled access for external consumers without revealing original models. The proposed framework opens up an approach toward mitigation of AI specific threat vectors such as model theft, reverse engineering, and intellectual property appropriation, while creating a foundation for a trust-minimized AI collaboration between untrusting organizations.

Keywords: artificial intelligence, blockchain, homomorphic encryption, zero knowledge proofs

1 Introduction

The rapid adoption of artificial intelligence and machine learning (AI/ML) across industry has introduced a growing class of security concerns that extend beyond traditional software vulnerabilities. As organizations deploy increasingly capable models trained on proprietary datasets, questions of provenance, integrity,

and verifiable performance have moved from academic interest to operational necessity. The public release of large language models (LLMs) since 2022 [17] has accelerated this trend, drawing attention to the broader ecosystem of specialized AI/ML models whose security properties remain poorly understood and inadequately addressed.

The security surface of an AI/ML model spans multiple dimensions: the protection of model weights and architecture against reverse engineering; the integrity of training data against poisoning; and the verifiability of claimed performance against independent evaluation. Existing approaches address these dimensions in isolation. Model watermarking and access-controlled APIs protect intellectual property but provide no mechanism for external verification of performance claims. Centralized audit frameworks introduce trusted intermediaries whose neutrality and availability cannot be guaranteed. Neither approach scales to the multi-party, cross-organizational collaboration that modern AI deployment increasingly demands. Furthermore, handling these issues enables a verifiable and secure approach for development of ensemble AI models involving personal sensitive data [13, 14] or the development of IoT AI devices for sensitive data [20].

Reported adversarial attacks on AI/ML systems have increased substantially since 2022 [21], including data poisoning, model inversion, and intellectual property theft. These incidents highlight the absence of a principled, trust-minimized mechanism by which model providers can assert, and model consumers can verify, the provenance and performance of AI/ML assets without exposing them. This paper presents a conceptual cryptographic framework that addresses this gap, with implementation, benchmarking and formal security proofs left for future work. We make the following conceptual contributions:

- A Merkle DAG-based model chaining protocol with selective metadata disclosure via Pedersen commitments
- A non-interactive proof of honest accumulation using Damgård-Jurik homomorphic encryption and Sigma protocols
- A cross-chain ensemble construction under threshold encryption keys that preserves the privacy of participating model owners

2 Problem statement

Securing AI/ML models introduces a fundamental verification paradox: if a model's weights and architecture are withheld to protect intellectual property, external parties cannot independently confirm that its claimed performance characteristics are genuine. Two conventional approaches exist, each with significant limitations.

The first approach is direct model evaluation, in which the prospective consumer is granted temporary inference access to assess performance on their data. This exposes the provider to intellectual property risk: a consumer may reverse-engineer architectural properties from query-response pairs or extract functional

knowledge of the model through systematic probing under the cover of evaluation. The second approach delegates verification to a trusted third party, such as an independent auditor or certification body. This avoids direct model exposure but introduces a centralized point of trust whose impartiality, availability, and computational resources cannot be cryptographically guaranteed.

The absence of a satisfactory solution to either limitation motivates the central research question of this paper: how can a decentralized framework provide AI/ML model consumers with cryptographically verifiable guarantees of provenance and performance, without requiring model providers to expose proprietary assets or delegate trust to a centralized intermediary?

3 Related work

Several instances of prior work are relevant to the proposed framework. Model watermarking techniques [22, 3] embed verifiable signatures into trained models to assert ownership, but do not provide performance guarantees or support compositional provenance across model versions. Federated learning protocols [15] enable collaborative model training without centralizing data, but do not address the problem of verifying individual model contributions or chaining their provenance. Blockchain-based integrity frameworks for AI/ML models have been explored [2, 9, 23], establishing the suitability of immutable append-only structures for model provenance; the present work extends this direction by introducing cryptographic commitments and homomorphic accumulators in place of plaintext on-chain storage.

In the cryptographic literature, Pedersen commitments [19] and their homomorphic properties have been applied to verifiable computation and privacy-preserving aggregation. The Damgård-Jurik cryptosystem [7], a generalization of Paillier [18], provides the additive homomorphic accumulation used in this work. Zero-knowledge proof systems, including Sigma protocols [6] and their non-interactive variants via Fiat-Shamir [11], have been applied to proof-of-knowledge in decentralized settings but have not previously been applied to AI/ML model provenance chains. To the best of our knowledge, this paper presents the first framework that integrates these primitives into a unified, privacy-preserving model chaining and ensemble construction protocol.

4 Preliminaries

This section introduces the cryptographic primitives underlying the proposed framework. All constructions operate in a cyclic group G of prime order $q = |G|$, with public generators $g, h \in G$, within the multiplicative group $\mathbb{Z}_p^*$ for a large prime p. The discrete logarithm problem (DLP) in this group is assumed to be computationally hard: Given g, p, $y = g^x \bmod p$ finding x is infeasible for any probabilistic polynomial-time adversary. This hardness assumption underpins the security of all primitives below.

4.1 Merkle Directed Acyclic Graphs (DAGs)

A Merkle DAG [16] is a directed acyclic graph in which each node's identifier is the cryptographic hash of its content and the identifiers of its parent nodes. This structure provides tamper-evidence: any modification to a node invalidates its hash and all descendant hashes, which is detectable by any verifier recomputing the graph. Unlike a simple hash chain, a DAG supports multi-parent nodes, enabling branching (model specialization) and merging (ensemble construction) within a single structure. Each node's hash is computed as:

$$\mathsf{node_hash}_i = H\big(C_{\mathsf{pub}} \,\|\, C_{\mathsf{priv}} \,\|\, C_{\mathsf{eval}} \,\|\, \mathsf{parent_hashes} \,\|\, \mathsf{DJ_acc}_i\big) \tag{1}$$

where H denotes a collision-resistant hash function, $||$ denotes concatenation with length-prefixed components and the remaining terms are defined in subsequent subsections.

4.2 Pedersen Commitments

A Pedersen commitment [19] allows a party to commit to a value m without revealing it, and later prove the commitment opens to exactly that value. Given public group generators g and h, where no party knows the discrete logarithm $\log_g(h)$, a commitment to m with random blinding factor $r \in \mathbb{Z}_q$ is:

$$C = g^m \cdot h^r \bmod p \tag{2}$$

The scheme satisfies two properties:

- Perfect hiding: C is uniformly distributed over the group regardless of m; even an unbounded adversary learns nothing about m from C alone.
- Computational binding: Opening C to a value $m' \neq m$ requires computing $\log_g h$, which contradicts the DLP assumption.

Pedersen commitments are additionally additively homomorphic:

$$C_1 \cdot C_2 = g^{m_1+m_2} \cdot h^{r_1+r_2} \tag{3}$$

so $C_1 \cdot C_2$ is a valid commitment to $m_1 + m_2$ with blinding factor $r_1 + r_2$, enabling homomorphic operations. In the proposed framework, three commitments are maintained per chain node: C_{pub} over public metadata, C_{priv} over private metadata, and C_{eval} over the evaluation dataset hash.

4.3 Homomorphic Encryption

Homomorphic encryption (HE) [1] is a class of public-key encryption schemes that permit computations to be performed directly on ciphertexts. For an additively homomorphic scheme:

$$\mathsf{Dec}\big(\mathsf{Enc}(m_1) \oplus \mathsf{Enc}(m_2)\big) = m_1 + m_2 \tag{4}$$

where $\oplus$ denotes the homomorphic addition operation over ciphertexts. This property allows the accumulator to be maintained and updated without decrypting the running aggregate at any intermediate step.

HE schemes are classified by the operations they support. Fully homomorphic encryption (FHE) supports arbitrary computations but introduces a noise term that grows with each operation, requiring periodic bootstrapping to refresh ciphertexts. Partially homomorphic encryption (PHE) supports a single operation type without noise accumulation and without a bound on the number of operations.

The semantic security guarantee that a ciphertext reveals nothing about its plaintext to any probabilistic polynomial-time adversary is provided by fresh randomness chosen independently at each encryption, yielding indistinguishability under chosen plaintext attack (IND-CPA) [12].

4.4 Damgård-Jurik Cryptosystem

The Damgård-Jurik (DJ) cryptosystem [7] generalises the Paillier scheme[18] to operate over $\mathbb{Z}^*_{n^{s+1}}$ for an integer expansion parameter $s \geq 1$, with plaintext space $\mathbb{Z}_{n^s}$. Security rests on the Decisional Composite Residuosity (DCR) assumption.

Given an RSA modulus $n = pq$, public key (n, s), and fresh randomness r, a plaintext m is encrypted as:

$$\mathsf{CT} = g^m \cdot r^{n^s} \bmod n^{s+1} \tag{5}$$

Security derives solely from the randomness r, not from noise. Consequently, the scheme supports an unlimited number of homomorphic additions without degradation:

$$\mathsf{CT}_1 \cdot \mathsf{CT}_2 \bmod n^{s+1} = \mathsf{Enc}(m_1 + m_2) \tag{6}$$

$$\mathsf{CT}^k \bmod n^{s+1} = \mathsf{Enc}(k \cdot m) \tag{7}$$

$$\mathsf{Enc}(-m) = \mathsf{Enc}(n^s - m) \tag{8}$$

The absence of noise limits differentiates DJ from other schemes such as CKKS [4] or BFV [10], making it well-suited to an append-only accumulator that grows indefinitely with the model chain. Numeric metadata fields are encoded as fixed-point integers and packed into a single plaintext m via a public, deterministic $\mathsf{encode}()$ function prior to encryption.

4.5 Sigma Protocols and Non-Interactive Zero-Knowledge Proofs (ZKPs)

A Sigma protocol [6] is a three-move interactive proof system ($\mathsf{Commit} \rightarrow \mathsf{Challenge} \rightarrow \mathsf{Response}$) in which a prover convinces a verifier of knowledge of a witness w satisfying a public statement, without revealing w. The protocol satisfies three properties:

- Completeness: An honest prover always convinces an honest verifier.
- Special soundness: Given two accepting transcripts with the same commitment but distinct challenges, the witness w can be extracted efficiently.
- Honest-verifier zero-knowledge: The transcript reveals nothing beyond the truth of the statement.

The Fiat-Shamir heuristic [11] transforms any Sigma protocol into a non-interactive zero-knowledge proof (NIZK) in the random oracle model by replacing the verifier's random challenge e with a hash of all public inputs and the prover's first message:

$$e = H(\mathsf{statement} \,\|\, A) \tag{9}$$

The resulting proof (A, z) can be stored at generation time and verified by any party at any future point without prover availability, which is essential for a publicly auditable chain.

4.6 Proof of Honest Accumulator Update and Node Signature

The DJ accumulator ensures the running aggregate of metadata remains encrypted. However, it does not by itself prevent a dishonest model owner from publishing accurate public metadata while secretly accumulating a different value in the ciphertext. The non-interactive proof π_{update} closes this gap.

Statement and witness. Let $m_i = \mathsf{encode}(\mathsf{public_meta}_i)$ be the publicly computable encoding of the declared metadata at node i. The prover asserts knowledge of witness r_{dj} satisfying:

$$\mathsf{DJ_acc}_i = \mathsf{DJ_acc}_{i-1} \oplus \mathsf{Enc}(m_i,\, r_{\mathsf{dj}}) \tag{10}$$

Proof generation. The prover selects fresh random values $k, s \leftarrow \mathbb{Z}_{n^s}$ and computes the commitment:

$$A = \mathsf{Enc}(k,\, s) \tag{11}$$

The Fiat-Shamir challenge is computed as:

$$e = H\big(\mathsf{DJ_acc}_{i-1} \,\|\, \mathsf{DJ_acc}_i \,\|\, m_i \,\|\, A\big) \tag{12}$$

The prover computes the response:

$$z = k + e \cdot r_{\mathsf{dj}} \bmod n^s \tag{13}$$

The proof stored at the node is the pair $\pi_{\mathsf{update}} = (A,\, z)$.

Verification. Given (A, z) and the public inputs, the verifier recomputes e via the same hash and checks:

$$\mathsf{Enc}(e \cdot m_i,\; z) \stackrel{?}{=} A \oplus \big(\mathsf{DJ_acc}_i \ominus \mathsf{DJ_acc}_{i-1}\big)^{e} \tag{14}$$

Soundness follows from the special soundness property of the underlying Sigma protocol: a valid forgery without knowledge of r_{dj} would yield an extraction of a DCR witness, contradicting the security of the DJ scheme.

Node signature. In addition to π_{update}, the owner produces a digital signature over the complete node hash:

$$\sigma = \mathsf{Sign}(\mathsf{node_hash}_i,\ \mathsf{owner_sk}) \tag{15}$$

The signature authenticates the node under the owner's public key and provides security. While π_{update} prevents internal dishonesty in the accumulator construction, σ prevents external tampering with any component of the node.

4.7 Threshold Encryption and Distributed Key Generation

A (t, n)-threshold encryption scheme [8] distributes a secret key among n parties such that any t parties can jointly decrypt, but no coalition of fewer than t parties can do so. In the cross-chain ensemble setting, a joint DJ keypair $(\mathsf{pk}_{AB}, \mathsf{sk}_{AB})$ is established between participating model owners via a Distributed Key Generation (DKG) protocol, so that no single party holds the full decryption key. Homomorphic operations over the joint ciphertext proceed identically to the single-owner case:

$$\mathsf{DJ_acc}_E = \mathsf{CT}_A \oplus \mathsf{CT}_B = \mathsf{Enc}_{\mathsf{pk}_{AB}}(m_A + m_B) \tag{16}$$

Decryption requires each keyholder to contribute a partial decryption, which are then combined to recover the plaintext aggregate:

$$m_A + m_B = \mathsf{Combine}\big(\mathsf{PartDec}(\mathsf{DJ_acc}_E,\ \mathsf{sk}_A),\ \mathsf{PartDec}(\mathsf{DJ_acc}_E,\ \mathsf{sk}_B)\big) \tag{17}$$

5 Model chaining

Model chaining represents a novel concept where AI/ML models are connected to their previous or future iterations with an immutable link [23]. The goal of model chaining is to provide a hierarchical record of developed AI/ML models with their entire lineage, as well as characteristics allowing them to be utilised for specific purposes, similar how blockchain works with previous and future nodes. Chaining is performed using Merkle DAGs, where each node's identifier is the cryptographic hash of its content and the identifiers of its parent nodes. This approach supports varied connections between models, enabling tracing of model specialisation and ensemble construction.

A key design decision is chaining only model metadata, not model weights. Including weights would make the chain impractical at scale and would expose intellectual property to anyone with chain access. Instead, model owners retain weights locally and publish a structured metadata record per version. Metadata encompasses quantitative performance indicators - training and test loss,

F1 score, accuracy, number of training epochs, as well as a commitment to the evaluation dataset used to produce those indicators. The evaluation dataset commitment serves a dual purpose: it enables downstream verification of the published metrics without requiring inference access to the model, and it prevents retroactive substitution of a more favourable evaluation set after metrics have been published, since the commitment is cryptographically bound to the chain node at the time of publication. This further allows external parties to successfully address if the provided model with its metadata suits their applicable use cases.

Numeric metadata fields are accumulated across chain nodes using Damgård-Jurik homomorphic encryption. Each new node's encoded metrics are encrypted under the owner's public key and added homomorphically to a running ciphertext accumulator, without requiring decryption at any intermediate step:

$$\mathsf{DJ_acc}_i = \mathsf{DJ_acc}_{i-1} \oplus \mathsf{Enc}(m_i, r_{\mathsf{dj}}) = \mathsf{Enc}\left(\sum_{j=1}^{i} m_j\right) \tag{18}$$

Private metadata fields not intended for public disclosure are committed via Pedersen commitments, which provide cryptographic binding without revealing the underlying values. The existence of private metadata allows partial distribution of model characteristics further enhancing security by minimizing potential discovery surfaces.

Public metadata fields are disclosed in plaintext alongside each chain node, together with the Pedersen blinding factor r_{pub} required to verify the corresponding commitment. A non-interactive zero-knowledge proof π_{update} is generated at each node addition to prove that the DJ accumulator was updated with precisely the encoded value of the declared public metadata. This ensures that a model owner cannot selectively misrepresent metrics in the public layer while accumulating different values in the encrypted aggregate.

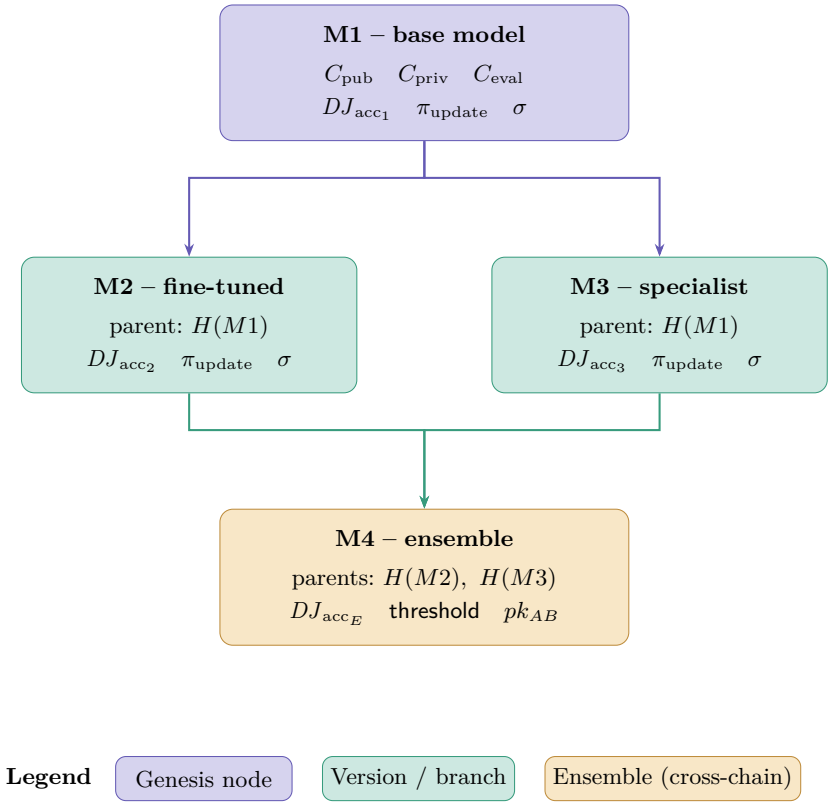


Fig. 1. Merkle DAG structure illustrating genesis, version, and ensemble nodes. $M2$ and $M3$ inherit from $M1$ via $H(M1)$ while the ensemble node $M4$ aggregates both chains under threshold keypair pk_{AB}.

6 Verifying model capabilities

External verification of a model's claimed capabilities is performed entirely from the publicly available chain data, without access to model weights or architecture. A verifier fetches the node object for the model version of interest and performs the following checks.

First, the verifier confirms that the disclosed public metadata is genuine by recomputing the Pedersen commitment from the plaintext values and the published blinding factor r_{pub}, and checking that the result matches C_{pub}:

$$g^m \cdot h^{r_{\mathsf{pub}}} \stackrel{?}{=} C_{\mathsf{pub}} \quad \text{where } m = \mathsf{encode}(\mathsf{public_meta}) \tag{19}$$

Second, the verifier confirms that the DJ accumulator was updated honestly at this node by verifying π_{update}, which establishes that the accumulator transition corresponds exactly to the declared public metadata as shown in (14)

Third, the verifier checks that the node was not modified after signing by recomputing the node hash from all its components and verifying the owner's digital signature σ:

$$\mathsf{Verify}(\sigma,\ \mathsf{node_hash}_i,\ \mathsf{owner_pk}) \stackrel{?}{=} 1 \tag{20}$$

Fourth, the verifier confirms that the evaluation dataset was registered prior to the publication of metrics. Since D_{eval} is included in $\mathsf{node_hash}_i$ and thus temporally bound to the chain at the point of node creation, retroactive substitution of the evaluation dataset is detectable by any verifier:

$$D_{\mathsf{eval}} = H(\mathsf{eval_dataset} \,\|\, \mathsf{salt}) \tag{21}$$

The consistency of the chain history is verified by following the parent hash links back to the genesis node, verifying the accumulator update proof at each step, and confirming that all nodes carry signatures from the same owner key. In cases where a verifier acquires the model weights independently, the verification can be extended: the model hash is checked against the value in the chain, and inference on the committed evaluation dataset can be performed locally to confirm that the published metrics are reproducible.

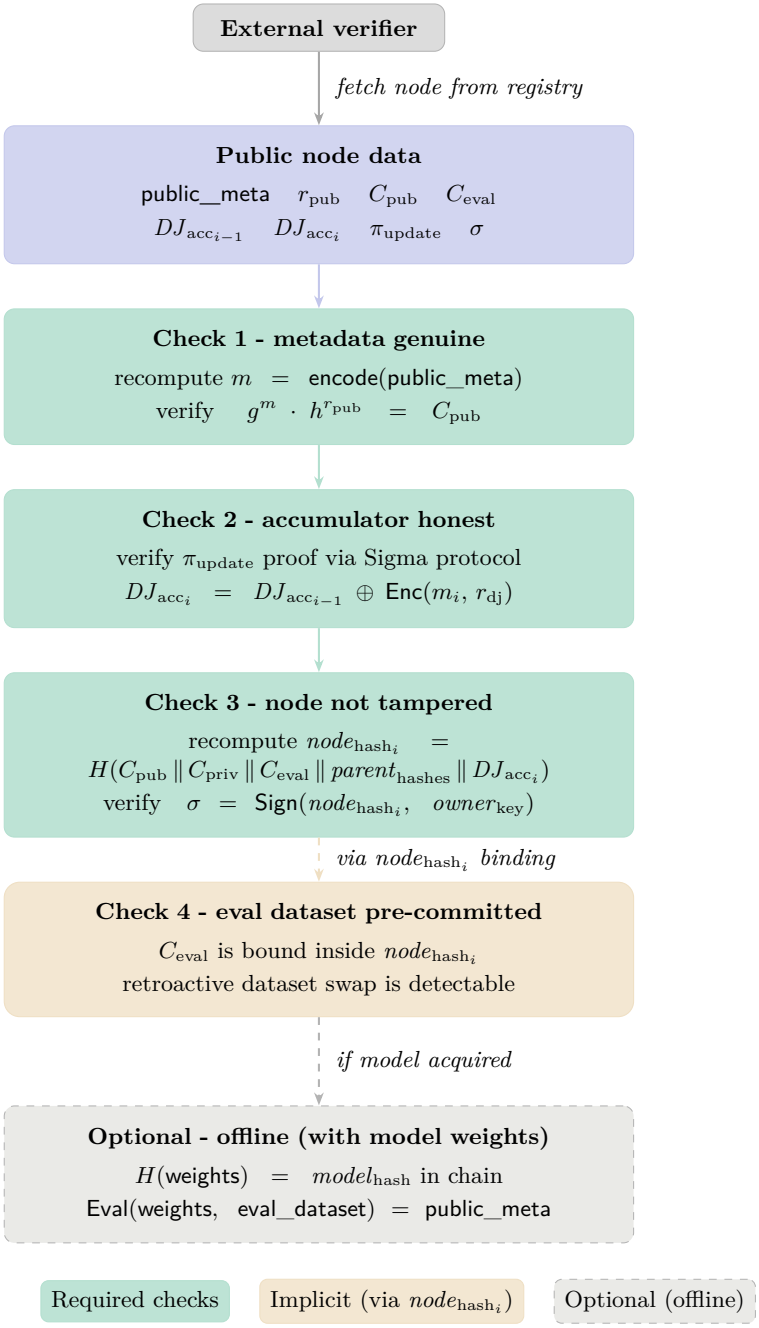


Fig. 2. External verifier protocol for a single chain node, allowing access to model's public metadata. Checks 1-3 are required; Check 4 is implicit via the $node_{\text{hash}}$ binding, while the offline check requires model weight access and is given as a plastic approach where full model is available to the external verifier.

7 Building ensembles

An ensemble model aggregates the capabilities of multiple component models, each potentially owned by a distinct organisation. In the proposed framework, ensemble construction is expressed as a DAG merge operation: a new ensemble node E is created with parent hash links to the terminal nodes of each participating model's chain. This structure, which we term *cross-chain linking*, is owner-agnostic with participating providers not required to share a common infrastructure, key material, or metadata schema beyond the protocol specification.

The ensemble node's DJ accumulator is encrypted under a joint public key established between the participating owners via a DKG protocol. Each owner contributes their chain's accumulated metrics by re-encrypting under the joint key:

$$\mathsf{CT}_A = \mathsf{ReEnc}(\mathsf{DJ_acc}_A,\ \mathsf{pk}_{AB}) \tag{22}$$

$$\mathsf{CT}_B = \mathsf{ReEnc}(\mathsf{DJ_acc}_B,\ \mathsf{pk}_{AB}) \tag{23}$$

and providing a proof of honest contribution π_{honest}, which demonstrates that the contributed ciphertext was correctly derived from their individual accumulator. The joint accumulator is then formed as:

$$\mathsf{DJ_acc}_E = \mathsf{CT}_A \oplus \mathsf{CT}_B = \mathsf{Enc}_{\mathsf{pk}_{AB}}(m_A + m_B) \tag{24}$$

The joint accumulator can be decrypted only through collaboration of all participating owners:

$$m_A + m_B = \mathsf{Combine}\big(\mathsf{PartDec}(\mathsf{DJ_acc}_E,\ \mathsf{sk}_A),\ \mathsf{PartDec}(\mathsf{DJ_acc}_E,\ \mathsf{sk}_B)\big) \tag{25}$$

This ensures that no single party can unilaterally access the aggregate or the individual contributions of others. External verification of the ensemble's capabilities follows the same protocol as single-chain verification described so far, with the additional requirement that both parent chains pass independent verification before the ensemble node is accepted. The ensemble node requires a signature from both owners over the shared $\mathsf{node_hash}_E$, so neither party can publish the ensemble unilaterally.

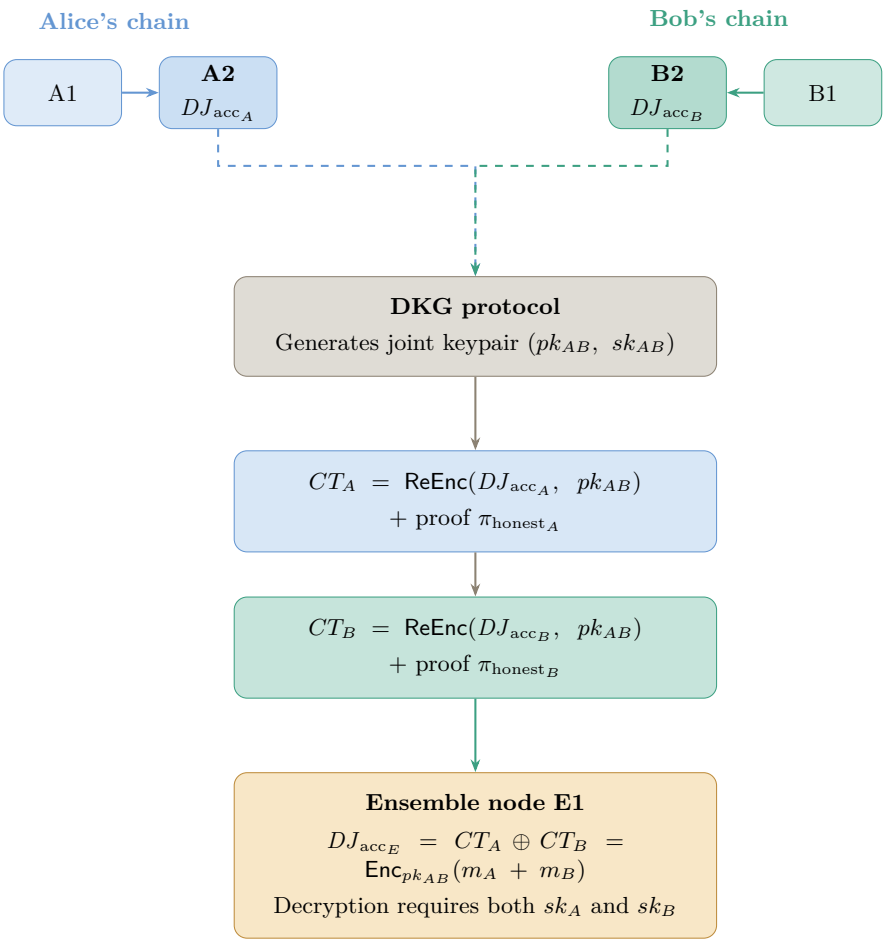


Fig. 3. Cross-chain ensemble construction via Distributed Key Generation. Both parties re-encrypt their accumulators under the joint key pk_{AB}; the ensemble node $E1$ requires both sk_A and sk_B for decryption.

8 Discussion and further work

The proposed framework introduces several directions for future investigation. The most immediate is empirical evaluation of the computational overhead introduced by Damgård-Jurik encryption and Sigma proof generation at chain write time. DJ homomorphic operations are significantly more expensive than plaintext arithmetic, and the practical throughput of the accumulator update protocol requires further research.

Scalability of the DAG structure with respect to chain depth and ensemble width also requires further analysis. The per-node verification cost grows linearly with chain length when a verifier traverses the full history. Techniques for succinct history proofs such as incrementally verifiable computation using recursive SNARKs could reduce verification cost to a constant independent of chain depth, at the expense of a more complex proof system. This represents a natural extension pathway for the journal version of this work.

Finally, the security of the framework rests on the hardness of the Discrete Logarithm Problem and the Decisional Composite Residuosity assumption, both of which are vulnerable to Shor's algorithm on a sufficiently capable quantum computer. Migration to post-quantum primitives such as lattice-based commitments and FHE schemes such as CKKS or TFHE [5] for the accumulator is an acknowledged limitation that deems exploration in further papers. The modular construction of the framework is designed to accommodate such substitutions without requiring architectural changes to the chaining or verification protocol.

9 Conclusion

This paper presented a cryptographic framework for privacy-preserving AI/ML model chaining that addresses the verification paradox at the core of secure model collaboration. By combining Merkle DAG structures with Pedersen commitments, Damgård-Jurik homomorphic encryption, and non-interactive Sigma protocol proofs, the framework enables model owners to publish selectively disclosed metadata with verifiable integrity guarantees, while external consumers can confirm provenance and performance claims without accessing model weights or relying on a trusted intermediary.

The cross-chain linking construction extends these guarantees to ensemble models involving multiple independent owners, using threshold Damgård-Jurik encryption to aggregate provenance metadata without requiring mutual disclosure of private model information. The result is a foundation for trust-minimised AI/ML collaboration applicable to federated learning deployments, multi-vendor AI pipelines, and regulated domains where model provenance must be demonstrable without compromising proprietary assets. This paper establishes the framework with conceptual introduction for further experimental evaluation, formal security proofs, and post-quantum extensions which will be done in a full journal version of this work.

References

1. Acar, A., Aksu, H., Uluagac, A.S., Conti, M.: A Survey on Homomorphic Encryption Schemes: Theory and Implementation. ACM Computing Surveys **51**(4), 79:1–79:35 (Jul 2018). https://doi.org/10.1145/3214303
2. Akaash Vishal Hazarika, Mahak Shah: Blockchain-based Distributed AI Models: Trust in AI model sharing. International Journal of Science and Research Archive **13**(2), 3493–3498 (Dec 2024). https://doi.org/10.30574/ijsra.2024.13.2.2598

3. Boenisch, F.: A Systematic Review on Model Watermarking for Neural Networks. Frontiers in Big Data **4**, 729663 (Nov 2021). https://doi.org/10.3389/fdata.2021.729663
4. Cheon, J.H., Kim, A., Kim, M., Song, Y.: Homomorphic Encryption for Arithmetic of Approximate Numbers. Lecture Notes in Computer Science (2017). https://doi.org/10.1007/978-3-319-70694-8_15
5. Chillotti, I., Gama, N., Georgieva, M., Izabachène, M.: TFHE: Fast Fully Homomorphic Encryption Over the Torus. Journal of Cryptology **33**(1), 34–91 (Jan 2020). https://doi.org/10.1007/s00145-019-09319-x
6. Cramer, R., Damgard, I., Schoenmakers, B.: Proofs of partial knowledge and simplified design of witness hiding protocols. Advances in Cryptography - Eurocrypt'94 (Santa Barbara CA, USA, August 21-25, 1994) pp. 174–187 (1994). https://doi.org/10.1007/3-540-48658-5_19
7. Damgard, I., Jurik, M.: A Generalisation, a Simplification and Some Applications of Paillier's Probabilistic Public-Key System. BRICS Report Series (Dec 2000)
8. Desmedt, Y., Frankel, Y.: Threshold cryptosystems. In: Brassard, G. (ed.) Advances in Cryptology — CRYPTO' 89 Proceedings. pp. 307–315. Springer, New York, NY (1990). https://doi.org/10.1007/0-387-34805-0_28
9. Dimitrova, P.D.S.: Blockchain Timestamping for AI Model Lifecycle Tracking. Scientific Journal of Artificial Intelligence and Blockchain Technologies **2**(1) (Jan 2025). https://doi.org/10.63345/sjaibt.v2.i1.109
10. Fan, J., Vercauteren, F.: Somewhat Practical Fully Homomorphic Encryption. IACR Cryptol. ePrint Arch. (2012)
11. Fiat, A., Shamir, A.: How To Prove Yourself: Practical Solutions to Identification and Signature Problems. In: Odlyzko, A.M. (ed.) Advances in Cryptology — CRYPTO' 86. vol. 263, pp. 186–194. Springer Berlin Heidelberg, Berlin, Heidelberg (2006). https://doi.org/10.1007/3-540-47721-7_12
12. Goldwasser, S., Micali, S.: Probabilistic encryption. Journal of Computer and System Sciences **28**(2), 270–299 (Apr 1984). https://doi.org/10.1016/0022-0000(84)90070-9
13. Hu, J., Shen, D., Chen, L., Chen, Y., Shen, S., Che, Y.: Joint heart sound denoising using DTCWT and Adaptive Sparsity-assisted signal smoothing algorithm. Computer Science and Information Systems **23**(2), 687–706 (2026)
14. Liu, L., Yin, F., Yang, G., Li, Q., Li, S., Pan, T., Liu, T., Tang, J., Ming, R., Song, Y., Feng, X., Wang, D., Zhou, X., Chen, W., Deng, J.: Development and validation of a few-shot rapid screening model for gastrointestinal cancers using AGI large vision models. Computer Science and Information Systems **23**(2), 801–826 (2026)
15. McMahan, H.B., Moore, E., Ramage, D., Hampson, S.C.D., Arcas, B.A.Y.: Communication-Efficient Learning of Deep Networks from Decentralized Data. In: International Conference on Artificial Intelligence and Statistics (Feb 2016)
16. Merkle, R.C.: A Digital Signature Based on a Conventional Encryption Function. In: Pomerance, C. (ed.) Advances in Cryptology — CRYPTO '87. pp. 369–378. Springer, Berlin, Heidelberg (1988). https://doi.org/10.1007/3-540-48184-2_32
17. OpenAI: Introducing ChatGPT. https://openai.com/index/chatgpt/ (Nov 2022)
18. Paillier, P.: Public-Key Cryptosystems Based on Composite Degree Residuosity Classes. In: Stern, J. (ed.) Advances in Cryptology — EUROCRYPT '99. pp. 223–238. Springer, Berlin, Heidelberg (1999). https://doi.org/10.1007/3-540-48910-X_16
19. Pedersen, T.P.: Non-Interactive and Information-Theoretic Secure Verifiable Secret Sharing. In: Feigenbaum, J. (ed.) Advances in Cryptology — CRYPTO

'91. vol. 576, pp. 129–140. Springer Berlin Heidelberg, Berlin, Heidelberg (1992). https://doi.org/10.1007/3-540-46766-1_9
20. Sovinc, L., Bajec, M.: Review of architectures in IoT remote patient monitoring systems. Computer Science and Information Systems **23**(1), 61–74 (2026)
21. Vassilev, A., Oprea, A., Fordyce, A., Anderson, H., Davies, X., Hamin, M.: Adversarial Machine Learning: A Taxonomy and Terminology of Attacks and Mitigations. Tech. Rep. NIST Artificial Intelligence (AI) 100-2 E2025, National Institute of Standards and Technology (Mar 2025). https://doi.org/10.6028/NIST.AI.100-2e2025
22. Zhang, J., Chen, D., Liao, J., Zhang, W., Feng, H., Hua, G., Yu, N.: Deep Model Intellectual Property Protection via Deep Watermarking. IEEE Transactions on Pattern Analysis and Machine Intelligence pp. 1–1 (2021). https://doi.org/10.1109/TPAMI.2021.3064850
23. Živadinović, M., Simić, D.: Blockchain-Based AI Model Integrity and Verification Framework. In: Unlocking the Hidden Potentials of Organization Through Merging of Humans and Digitals, Lecture Notes in Networks and Systems, vol. 1. Springer Nature Switzerland

Original Research Article *https://doi.org/10.5281/zenodo.23472609*

Intention to Adopt Blockchain Technology for Improving Assessment Quality in Higher Vocational Education in Serbia

Miloš Mijić[1][0009-0005-9648-7996] and Branko Ćebić[2][0009-0004-0576-8403]

[1] Academy of Applied Studies Novi Sad, School of Business Novi Sad, Serbia
[2] Academy of Applied Studies Western Serbia, Valjevo Department Užice, Serbia
mijic.milos3@gmail.com, branko.cebic@vipos.edu.rs

Abstract. Blockchain technology has the potential to be successfully implemented in higher vocational education in Serbia as an emerging technology that can significantly contribute to improving the quality of assessment by making it more transparent and unbiased. The application of blockchain in assessment creates opportunities for designing and implementing a reliable and transparent student evaluation system, thereby achieving a shared goal of both teachers and students—fair and transparent assessment. The aim of this study was to examine the readiness of teaching staff to adopt blockchain technology in the field of student assessment. The obtained results may provide a significant contribution to understanding the importance of blockchain and its application in managing assessment processes in higher vocational education in Serbia. For the purposes of this research, the UTAUT2 (Unified Theory of Acceptance and Use of Technology 2) model was used to identify the factors that influence individuals' acceptance of information technologies. The study identifies key constructs that affect the intention to use blockchain technology in higher vocational education in Serbia within the student assessment process. The results indicate that all analyzed constructs—performance expectancy, effort expectancy, social influence, facilitating conditions, hedonic motivation, price value, and habit—have a statistically significant positive effect on the intention to use blockchain technology. The strongest relationships with usage intention are observed for habit, hedonic motivation, performance expectancy, and social influence, while facilitating conditions and price value show weaker, yet still statistically significant, relationships.

Keywords: blockchain technology; UTAUT2 model; technology acceptance intention; higher vocational education; assessment quality.

1 Introduction

New technologies in all spheres of life are developing rapidly today. They include both existing technologies enhanced by new applications and capabilities, as well as emerging technologies that span various fields of activity. It is often the case that new tech-

nologies are not fully implemented due to challenges related to their further development, implementation, and practical application. Emerging technologies are frequently perceived as technologies with significant potential to transform the existing state of affairs. An emerging technology is defined as a "radically new and relatively fast-growing technology characterized by a certain degree of coherence that persists over time, with the potential to significantly impact socio-economic domains, including the composition of actors, institutions, and patterns of interaction among them, along with the associated knowledge creation processes. However, it's most prominent impact lies in the future; therefore, in the emergence phase, it is still somewhat uncertain and ambiguous." [1] Given the rapid and profound changes affecting society as a whole, including education, there is a growing need for the implementation of new technologies in educational institutions in order to follow these changes and adapt to new circumstances. These changes require the adaptation and alignment of educational programs, particularly in the field of contemporary technologies such as information and communication technologies, as well as in engineering and science. Among new technologies are also educational technologies, within which one of the most promising for transforming the current state of the education sector is blockchain technology.

2 Literature Review

Human behavioral dimensions of potential technology users play an important role in the successful adoption of any new technology. Therefore, it is essential that they be examined and analyzed [2]. It is considered that the assessment of user behavioral dimensions and the adoption of blockchain technology in education is still an underexplored area. Existing research findings are scarce and mostly theoretical in nature. Ronaghi [3] argues that the degree to which a technology or information system is accepted and adopted directly depends on the effectiveness of its implementation. Blockchain technology (BLC) is an emerging technology, and its application in education remains insufficiently explored. In teaching activities, blockchain technology is still in the early stages of research and development; therefore, comprehensive studies and experiments in this field are necessary [4].

According to Johnson, Maddux, Liu, and Henderson [5], the development and implementation of IT technologies in education require the introduction of new methods for student assessment and evaluation, as traditional methods can no longer meet the needs of modern teaching management. New methods imply the use of IT technologies aimed at improving the accuracy and efficiency of teaching evaluation and student learning assessment. Universities and faculties should use information technologies that contribute to promoting a scientific, professional, and objective evaluation system. Such a system requires a paradigm shift toward technology-based educational evaluation [6]. The implementation of blockchain technology in educational assessment can ensure transparency in the evaluation process, guarantee data authenticity, and improve fairness of evaluation results in higher education institutions [7, 8]. Blockchain technology, as a motivational factor for students, can be successfully applied in the teaching

evaluation process by contributing to the development of students' abilities and learning improvement [9]. Since blockchain development provides technical support to colleges and universities in building scientific evaluation systems, Ouyang and Huang [10] proposed a blockchain-based educational evaluation management system that has the potential to eliminate shortcomings in current evaluation practices. Rong Hui and Yifan Hui [11] further argue that blockchain technology plays a significant role in improving the quality of teaching evaluation in applied technology universities. A study conducted in Jordanian higher education institutions examined the possibility of applying blockchain technology to design a reliable and transparent assessment system. Smart contracts were used in the evaluation process to ensure transparency and objectivity of grading. The aim of the study was to examine assessment security, data integrity, and process efficiency. The results showed that blockchain technology has strong potential to create and guarantee a secure and fair educational environment with appropriate assessment procedures. The authors also highlight that a limitation in the application of blockchain technology is the fact that most educational blockchains are designed primarily for tracking student tasks and responses. They emphasize the importance of integrating solutions that enable adequate assessment of the extent to which students have mastered the learning content [12].

Chen et al. [13] investigated the application of blockchain technology in teaching and learning and identified three domains with significant potential for improvement: formative assessment, learning activity design and implementation, and learning process monitoring. Mehmood, Abid, Farooq, and Nawaz [14] state that student progress can be monitored using formative or summative evaluation strategies and that both offline and online assessment forms can be applied. According to Alammary et al. [15], blockchain technology can be used for competence and learning outcome management, assessment of students' professional skills, protection of educational content, creation of collaborative learning environments, enhancement of student interaction in e-learning, exam review processes, and support for lifelong learning. Blockchain can also contribute to improving exam paper auditing through reliable record-keeping systems [16]. Student evaluation is a very important part of the teaching process in higher education institutions. Since assessment involves a certain degree of subjectivity, Ramos-Sosa, Cabrera, and Moreno [17] argue that these issues can be eliminated by implementing blockchain technology for student assessment during knowledge testing. Exams can be programmed on a blockchain if databases containing questions, answers, and evaluation parameters are established. Students take exams on computers, while blockchain applications record grades and ensure objective evaluation. Student grades, representing their performance, are stored on the blockchain as part of their academic records.

Ramos-Sosa, Cabrera, and Moreno [17] propose a blockchain- and smart contract-based evaluation system capable of automated test generation and response registration in a blockchain ledger. Exam questions are selected from a question bank prepared by subject experts. Questions are designed according to difficulty levels, enabling the creation of tests with defined complexity. Responses are recorded chronologically, and the system ensures that answers and grades cannot be altered. Smart contracts [18], as simple computer programs stored on the blockchain, are automatically executed when predefined conditions are met. They are suitable for creating and storing courses, grades,

and instructor evaluations. Blockchain technology can improve exam review security through reliable record-keeping systems. Mitchell, Hara, and Sheriff [16] developed a blockchain-based decentralized application called "dAppER" for exam review. During its design, quality assurance standards for distributing exam papers among external examiners were considered. The results showed that the application significantly contributed to quality assurance management systems. Alammary et al. [15] emphasize that blockchain-based educational applications, including those used for exam review, have significant advantages over traditional assessment methods. Sharareh Alipour et al. [19] designed a new method for academic assessment based on decentralized blockchain applications. The method is applicable to various academic materials such as assignments, exams, and research papers and is designed to eliminate potential evaluator subjectivity. Survey results showed positive attitudes toward the fairness of this system compared to traditional assessment methods. The authors aim to design a model that satisfies both teachers and students, which is decentralized, transparent, based on universal and clear assessment procedures, subject to comprehensive peer review, capable of fast result distribution, and supported by a reliable infrastructure. A study on blockchain adoption in higher education institutions in the Kingdom of Saudi Arabia conducted by Alhumayzi and Mohammed [20] highlights the importance of examining behavioral intention to adopt blockchain technology among employees. It was assumed that the adoption of blockchain technology as an emerging innovation would likely face resistance. The results of an online questionnaire indicated several factors influencing the intention to use blockchain in higher education institutions in Saudi Arabia, including performance expectancy (PE), effort expectancy (EE), social influence (SI), facilitating conditions (FC), trust, perceived security (SEC), and awareness (AW). Swargiary [21] found that UTAUT2 factors such as performance expectancy and social influence significantly contribute to the intention to adopt blockchain technology in educational institutions in the United States. Behavioral intention was identified as a strong predictor of actual blockchain use. However, effort expectancy, facilitating conditions, hedonic motivation, price value, and habit were not significantly correlated with adoption intention, which contradicts findings from other studies. Swargiary explains this inconsistency by suggesting that blockchain adoption dynamics in education may differ from traditional behavioral patterns.

3 Research Methodology

The empirical research is based on a primary data collection approach to investigate the factors influencing the behavioral intention to adopt and implement blockchain technology in higher vocational education in the Republic of Serbia. The study utilizes the UTAUT2 (Unified Theory of Acceptance and Use of Technology 2) framework [22]. The research model incorporates all seven core constructs: performance expectancy, effort expectancy, social influence, facilitating conditions, hedonic motivation, price value, and habit, examining their effects on the dependent variable—behavioral intention (Fig. 1).

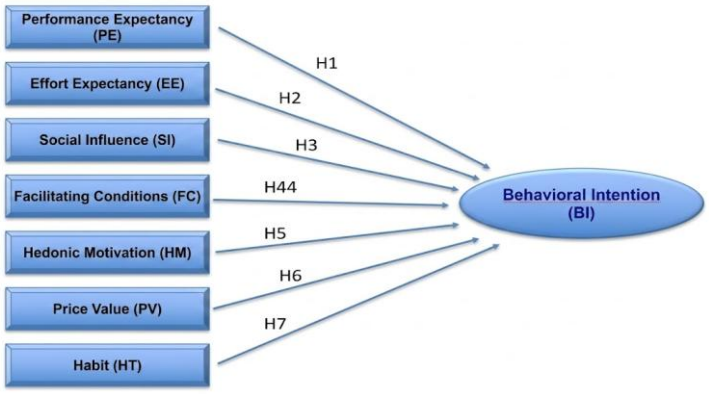


Fig. 1. Adapted UTAUT2 Technology Acceptance Model. [22]

Data collection was conducted using an online questionnaire technique administered between January and April 2026. The questionnaire consisted of two main sections. The first section gathered the socio-demographic characteristics of the respondents (gender, age, academic title, and work experience). The second section comprised 28 statements adapted from relevant literature to fit the context of blockchain adoption in student assessment. Respondents rated their agreement using a five-point Likert scale (1 – strongly disagree, 5 – strongly agree).Quantitative methodology was applied for data processing and analysis. Statistical analysis of the collected primary data was performed using SPSS version 20 software. The applied statistical procedures included descriptive statistical analysis (for sample structure), reliability testing (Cronbach's alpha), Pearson correlation analysis, and linear regression analysis. Simple linear regression was chosen over Structural Equation Modeling (SEM) due to the sample size (N=56). While SEM is generally preferred for the complex paths in the UTAUT2 framework, it typically requires a significantly larger sample to ensure model stability and statistical power. Given the exploratory nature of this study in the Serbian higher vocational education context, linear regression provides a robust and mathematically appropriate assessment of the individual impact of each construct.

4 Results and Discussion

The study involved a total of 56 respondents—including teachers, associates, and researchers—employed in the field of higher vocational education in the Republic of Serbia. The detailed socio-demographic structure of the sample is presented in Table 1.

Table 1. Demographic structure of the sample.

Description	Category	Percentage (%)	Number

Gender	Man	48.2	27
	Female	51.8	29
Age	18–25	10.7	6
	26–35	17.9	10
	36–45	30.4	17
	46–55	23.2	13
	56–65	17.9	10
Academic title	Lecturer	64.3	36
	Associate	30.4	17
	Researcher	5.4	3
Work experience	Less than 5 years	21.4	12
	5–10 years	12.5	7
	10–20 years	30.4	17
	More than 20 years	35.7	20

As detailed in Table 1, the gender distribution of the sample is relatively balanced, with female respondents accounting for 51.8% (N=29) and male respondents for 48.2% (N=27). In terms of age, the most represented group is 36–45 years old (30.4%), followed by the 46–55 age bracket (23.2%). Regarding academic roles, teaching staff (lecturers) dominate the sample at 64.3%, which is highly significant since they represent the primary stakeholders in the potential implementation of blockchain technology for student assessment. Furthermore, the sample consists of highly experienced professionals, with 35.7% of respondents having more than 20 years of work experience, and 30.4% having between 10 and 20 years. This diverse yet experienced structure provides a reliable foundation for examining attitudes toward modern technological integration in vocational education.

4.1 Reliability analysis of the instrument

The reliability of the measurement instrument was assessed using Cronbach's Alpha coefficient, i.e., the internal consistency method for each construct separately. The analysis was conducted for eight constructs used in the study: performance expectancy, effort expectancy, social influence, facilitating conditions, hedonic motivation, price value, habit, and behavioral intention. The obtained values of this coefficient range from 0 to 1. When the value is greater than 0.7, the reliability and consistency of the items are considered adequate [23], which is also consistent with the findings of George & Mallery [24], according to whom the interpretation of Cronbach's Alpha values can be as follows: $\alpha \geq 0.9$ – excellent reliability; $0.8 \leq \alpha < 0.9$ – good reliability; $0.7 \leq \alpha < 0.8$ – acceptable reliability; $\alpha < 0.7$ – low reliability. The reliability test results are presented in Table 2.

Table 2. Reliability of the measurement instrument.

Variables	*Cronbach's Alpha*
Performance Expectancy (PE)	0.931
Effort Expectancy (EE)	0.940
Social Influence (SI)	0.879
Facilitating Conditions (FC)	0.783
Hedonic Motivation (HM)	0.968
Price Value (PV)	0.941
Habit (HB)	0.955
Behavioral Intention (BI)	0.958

While some constructs exhibit Cronbach's Alpha values above 0.95, which might suggest item redundancy, these high values in this specific study are attributed to the highly standardized nature of the UTAUT2 scales and the focused context of blockchain adoption among a professionally homogenous group. This will be further addressed in the limitations section.

The obtained results indicate that all constructs demonstrate a satisfactory or high level of internal consistency, indicating adequate reliability of the instrument for further analysis. The lowest Cronbach's Alpha value was recorded for the construct facilitating conditions (FC) (0.783), which is considered an acceptable level of reliability. On the other hand, the highest value was observed for the hedonic motivation construct (HM), amounting to 0.968, indicating a very high internal consistency among the items within this construct. High Cronbach's Alpha values were also recorded for behavioral intention (BI = 0.958), habit (HB = 0.955), price value (PV = 0.941), effort expectancy (EE = 0.940), and performance expectancy (PE = 0.931), which confirms the very good reliability of these scales. The social influence construct (SI) also demonstrates high reliability, with a value of 0.879. The results presented in Table 2 confirm that the measurement instrument is reliable and suitable for further data analysis and hypothesis testing.

4.2 Correlation analysis

In order to examine the relationships between the constructs included in the research model (UTAUT2), a correlation analysis was conducted using Pearson's and Spearman's correlation coefficients. Correlation analysis was used to measure the degree of linear association between two variables, i.e., to determine the strength and direction of the linear relationship between variables. The closer the correlation coefficient is to 1 in absolute value, the stronger the correlation. The results (Table 3.) showed that there is a statistically significant positive relationship between most constructs (variables). It is particularly important to emphasize that behavioral intention to use blockchain technology (BI) is positively correlated with all other constructs in the Pearson analysis, while the Spearman analysis confirms a similar pattern of relationships, with minor differences in the strength of certain correlations.

Table 3. Correlation analysis

	PE	EE	SI	FC	HM	PV	H	BI
PE	1	0.393**	0.685**	0.174	0.820**	0.428**	0.712**	0.773**
EE	0.273*	1	0.340*	0.567**	0.355**	0.439**	0.563**	0.496**
SI	0.539**	0.151	1	0.210	0.641**	0.301*	0.651**	0.705**
FC	0.223	0.644**	0.232	1	0.191	0.109	0.331*	0.345**
HM	0.759**	0.177	0.608**	0.194	1	0.349**	0.790**	0.816**
PV	0.248	0.411**	0.117	0.274*	0.201	1	0.394**	0.346**
H	0.625**	0.488**	0.566**	0.372**	0.681**	0.273*	1	0.856**
BI	0.686**	0.370**	0.652**	0.388**	0.731**	0.178	0.808**	1

Notes: *** Statistically significant at the 1% level; * Statistically significant at the 5% level.*

Above the diagonal are Pearson correlation coefficients. Below the diagonal are Spearman correlation coefficients.

Source: Authors' calculation based on SPSS output.

The strongest relationship with behavioral intention to use blockchain technology (BI) was found for the habit construct (HB), both according to Pearson's correlation coefficient ($r = 0.856$) and Spearman's coefficient ($\rho = 0.808$). High positive correlations with behavioral intention were also observed for hedonic motivation (HM), performance expectancy (PE), and social influence (SI), indicating that these constructs are most closely associated with respondents' readiness to adopt blockchain technology. A moderate positive correlation was recorded for effort expectancy (EE), while facilitating conditions (FC) and price value (PV) showed weaker relationships with behavioral intention. In addition to the relationships with behavioral intention, significant interrelationships among individual independent constructs were also identified. Particularly notable are the correlations between performance expectancy and hedonic motivation, hedonic motivation and habit, as well as between performance expectancy and habit. These results indicate that, in respondents' perceptions, the usefulness, enjoyment, and routinization of blockchain technology are mutually reinforcing. Based on the obtained results, it can be concluded that Pearson and Spearman correlation analyses show a similar pattern of relationships, which further confirms the stability of associations between the analyzed constructs. The strongest relationships with behavioral intention are observed for habit, hedonic motivation, performance expectancy, and social influence, while the relationships with facilitating conditions and price value are weaker.

4.3 Regression analysis

In order to test the research hypotheses regarding the impact of individual constructs of the UTAUT2 model on the dependent variable, behavioral intention, a simple linear regression analysis was conducted. The dependent variable was behavioral intention to use blockchain technology, while the independent variables analyzed separately included: performance expectancy, effort expectancy, social influence, facilitating conditions, hedonic motivation, price value, and habit. The obtained results (Table 4) confirm the individual impact of the independent variables on the dependent variable, be-

havioral intention (BI). This indicates that as the values of each observed construct increase, the respondents' willingness to use blockchain technology in the student assessment process in the future also increases.

Table 4. Simple regression analysis

Independent variables	R^2	β	t
Performance Expectancy (PE)	0.598	0.773**	8.961
Effort Expectancy (EE)	0.246	0.496**	4.192
Social Influence (SI)	0.497	0.705**	7.309
Facilitating Conditions (FC)	0.119	0.345**	2.698
Hedonic Motivation (HM)	0.665	0.816**	10.364
Price Value (PV)	0.119	0.346**	2.706
Habit (HB)	0.733	0.856**	12.181

*Notes: ** Statistically significant at the 1% level (p < 0.01).*

Source: Authors' calculation based on SPSS output.

R^2* – *coefficient of determination, indicating the proportion of variance in the dependent variable (BI) explained by the independent variable.

β* – *standardized regression coefficient, indicating the strength and direction of the impact of the independent variable on the dependent variable.

t-* *t-value.

To address potential multicollinearity due to high correlations between certain constructs (e.g., HB and BI), Multicollinearity diagnostics were performed. The Variance Inflation Factor (VIF) values were within acceptable limits (below 5), confirming that the regression results remain statistically valid.

The strongest individual effect on behavioral intention was observed for the habit construct (HB), which explains 73.3% of the variance in behavioral intention (R^2 = 0.733). The value of the standardized beta coefficient (β = 0.856) confirms a strong and positive influence of habit on behavioral intention, indicating that the perception of blockchain technology becoming part of everyday professional practice represents the strongest individual predictor of future usage intention. A very strong impact was also found for hedonic motivation (HM) and performance expectancy (PE). The coefficient of determination (R^2 = 0.665) indicates that 66.5% of the variance in behavioral intention can be explained by changes in hedonic motivation (β = 0.816), while performance expectancy explains 59.8% of the variance (R^2 = 0.598; β = 0.773). These results show that the intention to adopt blockchain technology increases when respondents believe that the technology can improve the quality and efficiency of assessment, as well as when its use is perceived as interesting, enjoyable, and motivating. A significant positive effect was also found for social influence (SI), which explains 49.7% of the variance in behavioral intention (R^2 = 0.497; β = 0.705). This finding confirms that support from colleagues, the professional environment, and individuals whose opinions are valued by respondents can play an important role in shaping readiness to adopt blockchain technology. The coefficient of determination (R^2 = 0.246) and the beta coefficient (β = 0.496) indicate a positive but relatively weaker effect of effort expectancy on behavioral intention. This suggests that perceived ease of

use and understandability of the technology contribute to a higher willingness to adopt it, although this effect is weaker compared to the previously mentioned constructs. The weakest, but still statistically significant, individual effects were observed for price value (PV) and facilitating conditions (FC). Price value explains 11.9% of the variance in behavioral intention ($R^2 = 0.119$; $\beta = 0.346$), while facilitating conditions explain 11.9% of the variance ($R^2 = 0.119$; $\beta = 0.345$). These findings indicate that perceived cost-effectiveness and the availability of technical and organizational resources play a certain role, but are less pronounced as individual predictors of behavioral intention in this study.

4.4 Interpretation of hypotheses

In this study, seven hypotheses were formulated based on the assumption that the constructs of the UTAUT2 model have a positive impact on behavioral intention to use blockchain technology for improving the student assessment process in higher vocational education in Serbia. Based on the results of the simple linear regression analysis, it can be concluded that all proposed hypotheses were supported.

Hypothesis H1 assumes that performance expectancy has a positive impact on behavioral intention to use blockchain technology. The regression results showed that performance expectancy has a statistically significant positive effect on behavioral intention ($\beta = 0.773$), explaining 59.8% of the variance in the dependent variable. Accordingly, H1 is supported. This finding indicates that respondents are more willing to adopt blockchain technology if they believe it can contribute to greater efficiency, reliability, and quality of the assessment process.

Hypothesis H2 proposes that effort expectancy has a positive impact on behavioral intention to use blockchain technology. The results indicate that effort expectancy has a statistically significant positive effect on behavioral intention ($\beta = 0.496$), with 24.6% of explained variance. Therefore, H2 is supported. This means that respondents show greater readiness to adopt blockchain technology when they perceive it as understandable, easy to use, and relatively simple to learn.

Hypothesis H3 assumes that social influence has a positive impact on behavioral intention to use blockchain technology. The obtained results confirm this assumption, as social influence has a statistically significant positive effect on behavioral intention ($\beta = 0.705$), explaining 49.7% of the variance. Thus, H3 is supported. This result indicates that support from colleagues, institutions, and the professional environment plays an important role in shaping respondents' readiness to adopt blockchain technology.

Hypothesis H4 refers to the assumption that facilitating conditions have a positive impact on behavioral intention to use blockchain technology. The regression analysis showed that this construct also has a statistically significant positive effect ($\beta = 0.345$), although with relatively lower explanatory power compared to other constructs ($R^2 = 0.119$). Accordingly, H4 is supported. This means that the availability of technical infrastructure, support, and organizational resources contributes to behavioral intention, but to a lesser extent than some other factors.

Hypothesis H5 assumes that hedonic motivation has a positive impact on behavioral intention to use blockchain technology. The results show a very strong and statistically

significant positive effect of this construct (β = 0.816), explaining 66.5% of the variance in behavioral intention. Therefore, H5 is supported. This finding suggests that respondents are more willing to adopt the technology if they perceive its use as interesting, stimulating, and professionally engaging.

Hypothesis H6 proposes that price value has a positive impact on behavioral intention to use blockchain technology. The regression results indicate that price value has a statistically significant positive effect on behavioral intention (β = 0.346), explaining 11.9% of the variance. Accordingly, H6 is supported. Although this effect is weaker compared to most other constructs, the results suggest that perceived cost–benefit evaluation still plays a role in forming behavioral intention.

Hypothesis H7 assumes that habit has a positive impact on behavioral intention to use blockchain technology. The results show that habit is the strongest individual predictor of behavioral intention (β = 0.856), explaining 73.3% of the variance. Therefore, H7 is supported. This finding indicates that respondents' readiness to integrate blockchain technology into their regular professional practice is crucial for forming their future usage intention.

Based on the overall results, it can be concluded that all research hypotheses were confirmed, meaning that all analyzed constructs of the UTAUT2 model have a statistically significant positive impact on behavioral intention to use blockchain technology in higher vocational education. However, the strength of their influence is not equal. The strongest effects are observed for habit, hedonic motivation, performance expectancy, and social influence, while facilitating conditions and price value are relatively weaker predictors.

Table 5. Presents the results of hypothesis testing.

Hypothesis	Status
H1: Performance expectancy has a statistically significant impact on behavioral intention to use blockchain technology for improving assessment quality.	Hypothesis supported
H2: Effort expectancy has a statistically significant impact on behavioral intention to use blockchain technology for improving assessment quality.	Hypothesis supported
H3: Social influence has a statistically significant impact on behavioral intention to use blockchain technology for improving assessment quality.	Hypothesis supported
H4: Facilitating conditions have a statistically significant impact on behavioral intention to use blockchain technology for improving assessment quality.	Hypothesis supported
H5: Hedonic motivation has a statistically significant impact on behavioral intention to use blockchain technology for improving assessment quality.	Hypothesis supported
H6: Price value has a statistically significant impact on behavioral intention to use blockchain technology for improving assessment quality.	Hypothesis supported
H7: Habit has a statistically significant impact on the intention to use blockchain technology to improve rating quality.	Hypothesis supported

4.5 Discussion of results

The aim of this study was to examine the intention to adopt blockchain technology for improving the quality of student assessment in higher vocational education in Serbia, using the extended UTAUT2 model. The obtained results showed that all observed constructs have a positive and statistically significant impact on behavioral intention to use blockchain technology, although the strength of this impact varies across factors. This finding confirms that the adoption of new technological solutions in education cannot be explained by a single aspect, but rather results from the simultaneous influence of multiple factors. One of the most important findings of the study is that habit represents the strongest individual predictor of behavioral intention to use blockchain technology. This result indicates that respondents' willingness to perceive the application of blockchain systems as part of their regular professional practice is crucial for forming their future usage intention. In the context of higher vocational education, this means that mere awareness of the technology is not sufficient; instead, its adoption depends on whether it is perceived as a solution that can naturally be integrated into everyday academic and administrative activities. In other words, the more respondents are willing to view blockchain technology as a routine part of their work, the higher the likelihood that they will express an intention to use it. Hedonic motivation also showed a very strong influence, which represents a particularly interesting finding of this study. Although it is often assumed in educational and institutional contexts that rational and functional factors dominate, the results show that the subjective experience of using the technology—namely the perception that it is interesting, stimulating, and professionally appealing—significantly contributes to its acceptance. This can be interpreted by the fact that new technologies, especially those perceived as innovative and modern, may generate additional interest and intellectual engagement among respondents. In this sense, the adoption of blockchain technology is not based solely on perceived usefulness, but also on the perception that its use conveys a sense of professional modernity and innovation. Performance expectancy also demonstrated a significant and strong effect, which is consistent with the basic logic of technology acceptance models. This result indicates that respondents show a higher intention to use blockchain technology if they believe it can improve the quality, reliability, transparency, and efficiency of the student assessment process. In the context of higher education, this is particularly important, as the assessment process has not only an educational, but also an institutional and ethical dimension. The perception that blockchain can enhance the security and credibility of assessment likely encourages respondents to view it as a useful and justified solution for improving existing procedures. Social influence also showed a strong and statistically significant effect on behavioral intention. This finding indicates that the support and attitudes of colleagues, associates, institutions, and the broader professional environment play an important role in the adoption of blockchain technology. In academic settings, where technological and organizational changes are often implemented collectively and institutionally, the perception that relevant stakeholders support a certain solution can be decisive for its acceptance. This means that the intention to use blockchain technology is not solely an individual decision but is largely shaped by the social and professional context in which respondents operate. A moderate but

statistically significant effect was found for effort expectancy, indicating that ease of use and understandability of the technology matter for adoption, but not to the same extent as the previously mentioned factors. This result may be explained by the fact that respondents, as members of an academic and professional environment, may not primarily evaluate technology through ease of use, but rather through its functional value and potential benefits. Nevertheless, the statistical significance of this construct confirms that a clear, accessible, and user-friendly implementation of blockchain systems would contribute to greater adoption willingness. The weakest, although still significant, effects were observed for facilitating conditions and price value. This suggests that the availability of technical support, infrastructure, and organizational resources, as well as the perceived cost–benefit ratio, are relevant factors, but do not play a central role in shaping behavioral intention in this study. One possible explanation is that respondents still perceive blockchain technology more as a conceptual and strategic solution rather than as a fully operational system whose costs and infrastructural requirements can be precisely assessed. Therefore, usefulness, interest, and integration into everyday practice appear to be more influential than practical considerations of resources and costs. The obtained results also indicate that the adoption of blockchain technology in higher vocational education is a multidimensional process. Behavioral intention is shaped not only by perceptions of usefulness or ease of use, but also by whether the technology fits into professional routines, whether it is socially supported, and whether it generates a sense of professional motivation and engagement among users. Therefore, the observed pattern of results suggests that successful implementation of blockchain technology in student assessment would require not only technical system preparation, but also institutional promotion, user education, and gradual integration into everyday academic processes. Based on all the above, it can be concluded that the results confirm the relevance of the UTAUT2 model for explaining blockchain technology adoption in higher vocational education. Habit, hedonic motivation, performance expectancy, and social influence stand out as the key factors shaping behavioral intention, while effort expectancy, facilitating conditions, and price value make a somewhat weaker but still significant contribution. These findings provide a solid foundation for further theoretical and empirical research on the potential application of blockchain technology in improving the quality and transparency of educational processes.

Contribution: Based on the literature reviewed by the authors, this study is among the first to examine the intention to adopt blockchain technology for assessing the acquisition of teaching content in higher vocational education in Serbia.

Limitation: A significant limitation of this study is the sample size; therefore, future research should be conducted on a larger and more representative sample to generalize the findings. Another limitation is the relatively small number of existing studies dealing with the intention to adopt and implement blockchain technology specifically in the grading and teaching process. Furthermore, as with any emerging technology, blockchain may face initial resistance from both teachers and students due to a general reluctance to accept new digital tools. From a methodological standpoint, the exceptionally high Cronbach's Alpha values in some constructs may indicate potential item redundancy, suggesting that future studies should refine and shorten the measurement scales. Similarly, the high correlation between certain UTAUT2 constructs suggests an overlap

in respondent perception. This potential multicollinearity should be explored further in future research using larger samples and more complex structural models, such as Structural Equation Modeling (SEM).

5 Conclusion

This study successfully identified the key constructs influencing the behavioral intention to adopt blockchain technology for student knowledge assessment within the Serbian higher vocational education sector. The empirical findings demonstrate that all analyzed UTAUT2 constructs exercise a statistically significant positive impact on behavioral intention. Notably, habit (HB), hedonic motivation (HM), and performance expectancy (PE) emerged as the most powerful predictors. This indicates that for educators and institutions in this specific context, the decision to adopt blockchain is driven not merely by technical ease of use, but by the technology's potential to integrate into daily routines and its perceived professional value and attractiveness. For successful institutional implementation, higher vocational schools should focus on fostering a supportive digital culture rather than just providing infrastructure. Training programs should emphasize the practical benefits of blockchain in reducing administrative burdens and ensuring grading integrity (Performance Expectancy), while making the transition engaging and user-friendly (Hedonic Motivation) to facilitate routine integration (Habit) into daily academic workflows. While this study provides valuable exploratory insights, future research should overcome the sample size limitation by expanding the respondent pool across multiple regions. This would allow for more complex structural testing, including gender and age moderation, as well as Structural Equation Modeling (SEM) to further validate these findings.

References

1. D. Rotolo, D. Hicks, and B. R. Martin, “What is an emerging technology?,” Research Policy, vol. 44, no. 10, pp. 1827–1843, 2015, doi: 10.1016/j.respol.2015.06.006.
2. M. H. Ronaghi and J. Mahmoudi, “The relationship between corporate governance and IT governance in public organizations,” Journal of Information Technology Management, vol. 7, no. 3, pp. 615–634, 2015.
3. M. H. Ronaghi, “Open-source software migration under sanctions conditions,” International Journal of System Assurance Engineering and Management, vol. 12, no. 6, pp. 1132–1145, 2021, doi: 10.1007/s13198-021-01329-y.
4. W. Gräther, S. Kolvenbach, R. Ruland, J. Schütte, C. Torres, and F. Wendland, “Blockchain for education: Lifelong learning passport,” in Proceedings of the 1st ERCIM Blockchain Workshop 2018, 2018.
5. D. L. Johnson, C. D. Maddux, L. Liu, and N. Henderson, Evaluation and Assessment in Educational Information Technology. Boca Raton, FL, USA: CRC Press, 2021.

6. D. G. H. Divayana, P. W. A. Suyasa, and N. K. Widiartini, “An innovative model as evaluation model for information technology-based learning at ICT vocational schools,” Heliyon, vol. 7, no. 2, 2021, doi: 10.1016/j.heliyon.2021.e06347.
7. A. Machado, M. Sousa, and Á. Rocha, “Blockchain technology in education,” in Proceedings of the 4th International Conference on E-Commerce, E-Business and E-Government, 2020, pp. 130–134, doi: 10.1145/3409929.3416793.
8. B. Razia and B. Awwad, “A comprehensive review of blockchain technology and its related aspects in higher education,” in Technologies, Artificial Intelligence and the Future of Learning Post-COVID-19: The Crucial Role of International Accreditation, 2022, pp. 553–571, doi: 10.1007/978-3-030-93921-2_29.
9. R.-G. Chivu, I.-C. Popa, M.-C. Orzan, C. Marinescu, M. S. Florescu, and A.-O. Orzan, “The role of blockchain technologies in the sustainable development of students’ learning process,” Sustainability, vol. 14, no. 3, p. 1406, 2022, doi: 10.3390/su14031406.
10. S. Ouyang and X. Huang, “Education evaluation management based on blockchain technology,” Mobile Information Systems, vol. 2022, Article ID 7513365, 2022, doi: 10.1155/2022/7513365.
11. R. Hui and Y. Hui, “Application of blockchain technology in teaching evaluation in applied technology universities,” Journal of Combinatorial Mathematics and Combinatorial Computing, vol. 123, pp. 361–372, 2024, doi: 10.61091/jcmcc123-36.
12. F. Ayasrah, A. Shdouh, and K. Al-Said, “Blockchain-based student assessment and evaluation: A secure and transparent approach in Jordan's tertiary institutions,” Kurdish Studies, vol. 11, pp. 2036–2049, 2023, doi: 10.58262/ks.v11i02.144.
13. G. Chen, B. Xu, M. Lu, and N.-S. Chen, “Exploring blockchain technology and its potential applications for education,” Smart Learning Environments, vol. 5, no. 1, pp. 1–10, 2018, doi: 10.1186/s40561-017-0050-x.
14. A. Mehmood, M. Abid, M. S. Farooq, and N. A. Nawaz, “Curriculum, teaching and learning, and assessments for the introductory programming course,” IEEE Access, vol. 8, pp. 125961–125981, 2020, doi: 10.1109/ACCESS.2020.3008321.
15. A. Alammary, S. Alhazmi, M. Almasri, and S. Gillani, “Blockchain-based applications in education: A systematic review,” Applied Sciences, vol. 9, no. 12, p. 2400, 2019.
16. I. Mitchell, S. Hara, and M. Sheriff, “dAppER: Decentralised application for examination review,” in 2019 IEEE 12th International Conference on Global Security, Safety and Sustainability (ICGS3), 2019, pp. 1–14.
17. M. d. P. Ramos-Sosa, D. Cabrera, and B. Moreno, “Blockchain and smart contracts for education,” MPRA Paper No. 101518, Munich Personal RePEc Archive, 2020. [Online]. Available: https://mpra.ub.uni-muenchen.de/101518/ [Accessed: March 2026].
18. M. S. Stanković and J. S. Trajković, “Opportunities and challenges of applying blockchain technology and tokenization in education,” 2021. [Online]. Available: https://www.researchgate.net/... [Accessed: March 2026].
19. S. Alipour, S. Elahimanesh, S. Jahanzad, P. Morassafar, and S. P. Neshaei, “A blockchain approach to academic assessment,” in CHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI EA ’22), New Orleans, LA, USA, 2022, doi: 10.1145/3491101.3519682.

20. M. Alhumayzi, “Acceptance of blockchain technology by higher education institutions in the Kingdom of Saudi Arabia,” Ph.D. dissertation, Aston University, Birmingham, U.K., 2024.
21. K. Swargiary, “Predicting blockchain technology adoption in education: An application of the UTAUT2 model,” Preprints, 2024, doi: 10.20944/preprints202405.0210.v1.
22. V. Venkatesh, J. Y. L. Thong, and X. Xu, “Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology,” MIS Quarterly, vol. 36, no. 1, pp. 157–178, 2012.
23. J. F. Hair, W. C. Black, B. J. Babin, and R. E. Anderson, Multivariate Data Analysis, 8th ed. Andover, U.K.: Cengage Learning EMEA, 2018.
24. D. George and P. Mallery, SPSS for Windows Step by Step: A Simple Guide and Reference, 11.0 Update, 4th ed. Boston, MA, USA: Allyn & Bacon, 2003.

Original Research Article *https://doi.org/10.5281/zenodo.23473215*

Development of DApp for Fashion E-Commerce Supply Chains

Danica Sovtić[1[0009-0002-8601-1642]], Aleksandra Trpkov[1[0009-0007-7211-046X]], Marta Magovčević[1[0009-0000-3596-094X]], Božidar Radenković[1[0000-0003-2111-7788]], Sergei Prokhorov[2[0000-0002-1874-8602]] and Aleksandra Labus[1[0000-0002-7716-5845]]

[1] University of Belgrade, Faculty of Organizational Sciences, Belgrade, Serbia
[2] S.I. Vavilov Institute for the History of Science and Technology of Russian Academy of Science, Moscow, Russia
danica.sovtic@gmail.com, aleksandratrpkov1@gmail.com, martamagovcevic01@gmail.com
bozidar.radenkovic@fon.bg.ac.rs, sergei.prokhorov@gmail.com,
aleksandra.labus@fon.bg.ac.rs

Abstract. The growing complexity of fashion e-commerce supply chains has increased the need for reliable mechanisms that ensure transparency of business transactions among stakeholders, product traceability, and verification of product authenticity. Traditional centralized systems often limit transparency and reduce trust among supply chain participants, while also creating vulnerabilities to counterfeit fashion products. In response to these challenges, this paper presents the development of a decentralized application (DApp) to improve transparency, traceability, and verification processes in fashion e-commerce supply chains. The research focuses on modelling the system architecture, data flows, and key functionalities within the developed DApp. The proposed architecture is designed to support implementation in smaller fashion e-commerce supply chains, where transparent coordination among designers, manufacturers, distributors, and sellers is particularly important. The developed DApp is intended to enable verification of the originality and authenticity of fashion products within such supply chains. Through the proposed DApp, blockchain-based records are used to store transactions related to product development and testing, marketing activities, sales revenue, and operational expenses. By maintaining transparent and verifiable records of these activities, the application enables reliable tracking of product-related events throughout the supply chain. The proposed solution contributes to increasing consumer trust in the authenticity and origin of fashion products, enhancing transparency of business processes, reducing the risk of counterfeit goods, improving collaboration among supply chain stakeholders, and supporting more efficient and secure transactions in fashion e-commerce ecosystems.

Keywords: Blockchain; Decentralized Application; Fashion E-Commerce; Supply Chain

1 Introduction

The development of digital platforms and electronic commerce has improved the availability of products over time. It also enabled the growth of global trade. However, all these development processes have increased the complexity of fashion supply chains, which include many different actors. Some of them are suppliers of raw materials, manufacturers, logistics providers, distributors, retailers and end consumers. In such an environment, it is a challenge to manage the flow of information and monitor the movement of products throughout the network [1].

Lack of transparency between participants is a key challenge. Information about suppliers, production processes, as well as the origin of products is distributed through different systems. That is why it is difficult for interested parties to access and verify them. Companies, moreover, are unwilling to share such data. The reason for this is the protection of the competitive position, but also of internal business processes. As a consequence, the lack of transparency limits the ability of participants to verify the conditions of production and the origin of the product. And that affects consumer confidence and their perception of sustainability [2].

Another important problem in the fashion industry is the presence of counterfeit products on the global market. International trade and online sales channels have made it possible for counterfeit goods to look like original ones, but also to reach consumers more easily. This can lead to damage to the brand's reputation. Also, it affects the reduction of consumer confidence, and can lead to financial losses. Counterfeiting leads to weak protection of intellectual property and can often reduce the value of original design work [3].

Blockchain technology is recognized as a potential solution for improving transparency and traceability in supply chains. It enables the creation of distributed and immutable records that can be shared among many different participants. Blockchain also enables reliable tracking of product origin, then checking authenticity, and contributes to building trust between participants [2].

The goal of this research is the development of a decentralized application that is intended to improve transparency, traceability and authenticity verification in small fashion supply chains within electronic commerce. The development of this application can be seen as a response to the identified challenges and technological possibilities. The proposed system has several functions, such as recording key events related to product development, then production processes and product sales. In this way, transparency is ensured and the possibility of checking information about fashion products during their entire life cycle.

The proposed decentralized application is adapted to small fashion brand e-commerce systems. The system is based on a limited number of clearly defined participants, simple interactions with smart contracts, then selective recording of key events and a simple interface intended for users who do not have technical knowledge. The contribution of this work is reflected in the modelling of the modular and accessible architecture of the decentralized application. It is also suitable for small fashion businesses that are limited technically and financially.

The paper is organized in the following way. The second chapter contains an

overview of the relevant literature on the application of blockchain technology in fashion e-commerce supply chains. Then, the third chapter presents the modelling of the proposed system and the architecture. The fourth chapter refers to the development of the decentralized application, as well as the implementation of key functionalities. At the very end, the main results of the work are summarized in the conclusion, and directions for future research are indicated.

2 Blockchain and Decentralized Applications in E-Commerce Supply Chains

2.1 Blockchain Applications in E-Commerce Supply Chains

The use of blockchain technology in the context of e-commerce supply chain management has become a significant topic of interest over the past few years because of the benefits that the technology can offer in ensuring transparency, traceability, and trust. The conventional systems in which supply chain management operates rely on centralized data systems, incomplete information flow, and lack of visibility among other issues [2].

Blockchain brings an immutable and decentralized ledger system that allows for safe and transparent recording of transactions among many organizations. Transactions are always encrypted and can be altered only when all parties agree within the network. This greatly reduces chances of any manipulation or fraud. It becomes extremely essential to have blockchain technology in an environment like e-commerce that involves many transactions between individuals who do not necessarily trust one another. In such a scenario, smart contracts help in automating and enforcing the predetermined transaction rules and conditions without any intermediary. Smart contracts help in executing rules such as payment release after delivery or even checking on product integrity [4].

Blockchain technology when applied can improve the performance of the supply chain. It is possible to track goods in real time. Also, it is possible to provide a common and reliable source of information for all participants. Blockchain-based systems enable the recording of key events. Some of the events can be production, transport, storage and sales. Thus, a consistent history can be formed that can check the movement of the product [5]. This helps in making the traceability process easier. Using the blockchain technology together with digital identifiers, it is easy to link the physical item to its digital record, thus making it possible to verify the authenticity of the product [1].

Hybrid blockchains are often employed in e-commerce systems, in which essential elements like transactions, smart contracts, and validations are handled via on-chain technology, while larger and more adaptable data, such as product information, customer details, and media files, can be processed off-chain. This solution offers increased scalability and efficiency for the system, while maintaining the security and integrity of blockchain data [6]. The existing literature brings to attention a number of advantages of using blockchain technology for e-commerce supply chain

management, including increased transparency, traceability, mutual trust between the parties involved, and improved transaction processing efficiency. Nevertheless, questions regarding scalability, interoperability with legacy systems, and regulation still exist and need to be addressed by future studies.

2.2 Transparency and Traceability in Fashion E-Commerce Supply Chains

Supply chains in fashion and e-commerce are complex. They include global distributed operations and many different participants in manufacturing, logistics and retail. This complexity leads to divided information systems, limited visibility, as well as difficulties in tracking the origin of products. Transparency and traceability are still challenges in the fashion industry. Especially since consumers increasingly expect reliable information about the authenticity and sustainability of products.

Moreover, an additional issue involves the lack of means by which to trace the movement of products from inception until their disposal. Information concerning materials acquisition, manufacturing processes, and the product's movement through the channels of distribution will be fragmented, making verification of the product's origin and data integrity very difficult. Such fragmentation breeds mistrust and the possibility for errors, such as counterfeit production [7].

Blockchain technology offers a very good solution in tackling these problems since allows for the formation of an immutable ledger of activities that occur within the supply chain. With blockchain, all parties are allowed access to a shared and verified transactional history, which helps improve transparency in the process. All activities related to the product through the entire lifecycle, from manufacturing up to selling, can be captured using the blockchain [8].

The fashion industry needs traceability because there is a prevalence of counterfeit products. Blockchain technology provides an opportunity to develop unique digital identity cards for products that can then be tied to actual physical products using methods such as QR code or RFID technology. Using these methods, consumers or interested third parties can easily access information about the product's provenance, how it was made, and its history of ownership [9] [10]. Blockchain adoption along with Internet of Things (IoT) technology is very useful when it comes to tracking and managing production of goods, as IoT sensors provide the ability to collect and analyse real-time information concerning product quality, which is especially important for fashion supply chain management. With IoT sensors providing automatic data collection regarding temperature, humidity, and transportation conditions, data is recorded on blockchain, which increases its validity and authenticity [11].

2.3 Decentralized Applications for E-Commerce and Supply Chain Management

Decentralized applications are blockchain-based systems that operate without a central authority. They enable safe, transparent and automated interactions between participants in an e-commerce environment [4]. In the context of fashion supply

chains, decentralized applications enable key processes, such as transaction execution, product traceability, and authenticity verification. Functionalities are enabled through smart contracts. They serve to automate sales contracts, payments and transfer of ownership. What is important is that reliance on intermediaries is reduced, but on the other hand, trust between interested parties is strengthened [12]. A common approach involves recording key events in the supply chain. This creates a history of product movement that is resistant to unauthorized changes. Such systems in fact provide visibility in real time, but also facilitate coordination between the various participants involved in the supply chain processes [8, 10].

One of the major uses of DApps in e-commerce and supply chains is the verification of authenticity of products. In the current context, most of the approaches use systems that leverage digital identities, which are unique and associated with each physical product to be verified with the help of blockchain records [13]. This includes platforms like VeChain, which allow for product traceability through a unique digital identity assigned to each product via the blockchain and QR code technology [14]. The other notable example includes Provenance, which enables transparency in the supply chain by capturing product life cycle information and facilitating consumer verification through the blockchain system [11].

On the other hand, most of the blockchain-based traceability applications are designed with respect to more comprehensive supply chain ecosystems, business enterprises, or generic product origin verification. The designed decentralized application is intended to be implemented within fashion e-commerce supply chains of smaller size, wherein fewer actors take part and less formality and fewer resources can be involved. This means that all details associated with the operation cannot be stored in the blockchain. Instead, only certain points of the product's lifecycle will be logged into the blockchain including requests for materials, responses from suppliers, product approvals, testing by laboratories, quality control, and payments by customers. In order to further establish the proposed DApp's positioning among other blockchain-enabled tracing systems, a comparative analysis among the various selected platforms and research-oriented frameworks along with that of the proposed system is presented in Table 1 below. This comparison includes the application domain, main purpose, role of the blockchain in the solution, and the importance of each system in terms of understanding the proposed DApp.

Table 1. Comparative overview of blockchain-based traceability solutions and the proposed DApp

Solution/ approach	Domain	Main objective	Blockchain role	Relevance for comparison with the proposed DApp
Vichai platform	Product and supply chain traceability	Product authenticity verification and lifecycle tracking	Uses blockchain-based product identities and verification mechanisms, often connected with QR codes or similar identifiers	Enterprise-oriented platform for broad traceability use cases

Solution/ approach	Domain	Main objective	Blockchain role	Relevance for comparison with the proposed DApp
Provenance platform	Product provenance and supply chain transparency	Verification of product origin, sourcing, and production information	Records and Shares product provenance data to improve transparency for consumers and supply chain participants	General provenance platform focused on supply chain transparency and consumer trust
Ready-to-wear clothing traceability framework [17]	Ready-to-wear clothing and textile supply chains	Traceability of garments across production, distribution, and retail stages	Uses blockchain to record product and process information throughout the clothing supply chain	Research-based framework focused on ready-to-wear clothing traceability
Blockchain-based fashion supply chain transparency architecture [14]	Fashion industry supply chain transparency	Improving transparency, accountability, sustainability, and ethical compliance	Examines blockchain integration into enterprise architecture to support verifiable and durable supply chain records	Qualitative and architectural approach focused on fashion supply chain transparency
Proposed DApp	Small fashion e-commerce supply chains	Traceability, authenticity verification, and selected transaction recording	Uses modular smart contracts for selected lifecycle events, procurement, approvals, laboratory references, and customer payments, with operational data stored off-chain	Lightweight prototype tailored to small fashion e- commerce businesses, with role-based interfaces and selective on-chain recording

Nevertheless, there are several issues associated with scalability, interoperability with legacy systems, difficulty of implementation, and lack of user adoption faced by current traceability approaches based on blockchain technology. Based on information from Table 1, the main part of the currently used traceability approaches is aimed at designing more universal enterprise platforms or proving the concept of blockchain implementation in the fashion industry. These approaches do not take into consideration some specifics of a smaller fashion e-commerce ecosystem, which implies that there are fewer stakeholders involved, less formalized business processes, and fewer available technical resources. Therefore, it can be claimed that there is a need for fashion industry-specific DApp models, which should take into account various business processes associated with fashion products' creation, testing, quality approval, marketing, and sales.

3 Modelling a Decentralized Application for Small Fashion E-Commerce Supply Chains

3.1 System Architecture of the Proposed Decentralized Application

Fashion supply chains, especially in smaller fashion e-commerce businesses, often face challenges due to a lack of transparency, disconnected information flows, and the complexity of coordinating activities among multiple participants [15]. This section presents the system architecture of the proposed blockchain-based DApp for fashion product traceability and authenticity (Fig. 1). A decentralized application (DApp) connects the main stakeholders involved in the fashion product lifecycle and supports a transparent, secure, and structured exchange of data throughout the process [16].

The proposed architecture is organized around three main parts. On the left side, the model identifies the key stakeholders participating in the ecosystem: the fashion designer, material supplier, manufacturer, laboratory technician, quality tester, accountant, and marketing assistant. These stakeholders represent the main actors involved in product development, material procurement, production, testing, financial monitoring, and market preparation [17]. The reason for this stakeholder configuration is the limited scale of the proposed solution. Rather than depicting a supply chain that is very large and distributed and involves many intermediaries, the model concentrates on the important stakeholders that usually are pertinent to small online stores dealing in fashionable clothing. It simplifies the implementation of the model by making the interaction more straightforward for all participants of the process.

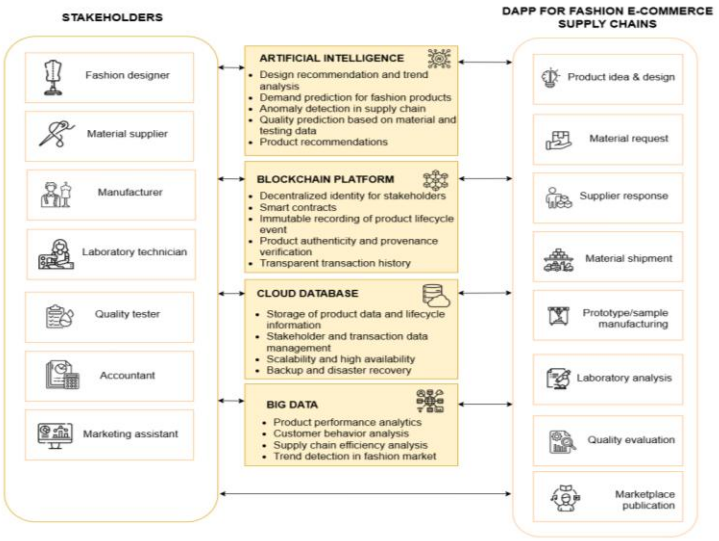


Fig. 1. System architecture of the blockchain-based decentralized application for fashion product traceability and authenticity.

The central part of the architecture presents the technological backbone of the proposed DApp. The architecture includes the blockchain, cloud database, and big data layers, and the artificial intelligence layer is shown as an extension that is yet to be developed in the future. In the next version of the architecture, the artificial intelligence layer can perform tasks such as designing and analysing trends, predicting demand, detecting anomalies within the supply chain process, and predicting the quality of products from lab results. This would assist in decision-making and help in optimizing crucial phases within the product life cycle including design, validation, and product launch in fashion e-commerce [18]. Since this feature is yet to be incorporated into the prototype, it is deemed as part of future system development.

Blockchain platform constitutes the key trust layer of the system. It makes decentralized identity management, smart contracts, consensus protocols, digital credentials management, and audit transparency possible, and is also responsible for enhancing data security and integrity [19]. The trust layer helps to capture stakeholder interactions in an immutable way to ensure the transparency of relevant transactions.

The database layer is where all the operational data necessary for the DApp to function are stored. This includes stakeholders' data, transactions, scalability, and disaster recovery. The big data layer assists in analytics through analysis of performance, sales trends, and consumer behaviour in the fashion ecommerce ecosystem. This includes metrics on number of sales, revenue trends, and popular products. The layer will also facilitate supply chain performance monitoring and analyse patterns that emerge at different lifecycle stages of the product. These layers form the technological foundation necessary for safe data storage and analytics [20].

On the right side of Fig. 1, the architectural framework depicts the various functions of the DApp. This involves product ideas and design, requesting materials, supplier response, delivery of the materials, manufacturing of prototypes or samples, laboratory testing, quality assessment, and market publishing. In this way, the model links the stakeholders to the concrete digital processes they perform within the application.

3.2 Process Modelling and Data Flow

The presented BPMN model describes the complete workflow within the proposed decentralized fashion e-commerce application, from the emergence of an idea for a new product to its purchase and delivery to the end customer (Fig. 2) [4]. The BPMN model represents a role-based sequence of business activities, where each horizontal pool represents one system participant, while the activities within it show the tasks for which that participant is responsible [21]. The process flow is followed from left to right within each pool, while the dashed lines between pools indicate the exchange of messages, requests, results, and approvals among different actors [22]. The diagram clearly illustrates the collaboration of all actors in the value chain, as well as the exchange of messages, documents, and decisions that enable the product to pass through all necessary stages of development, verification, and market placement [23].

The process starts with the fashion designer, who develops a product idea,

prepares a model draft, and defines the required materials through a bill of materials (BOM). A material supplier receives a request and confirms materials availability and delivery feasibility. Once the request is approved, the materials are provided to the manufacturer, who confirms their adequacy and develops a product sample. As soon as the product sample is made successfully, it is transferred to the fashion designer for consideration. Once the fashion designer confirms the sample conformity to requirements and approves it, the product goes through a testing stage. At this stage, the product sample goes to the laboratory for testing, and the results are evaluated and summarized into the report. With the help of the received information, quality tester approves the product. When the product is approved and goes through all development stages successfully, it is made available for sale and distributed on the online store. The product then becomes available to consumers, who can review its information and complete a purchase. The purchase will be completed upon a successful transaction and the completion of the process through delivery of the product and membership to the loyalty program.

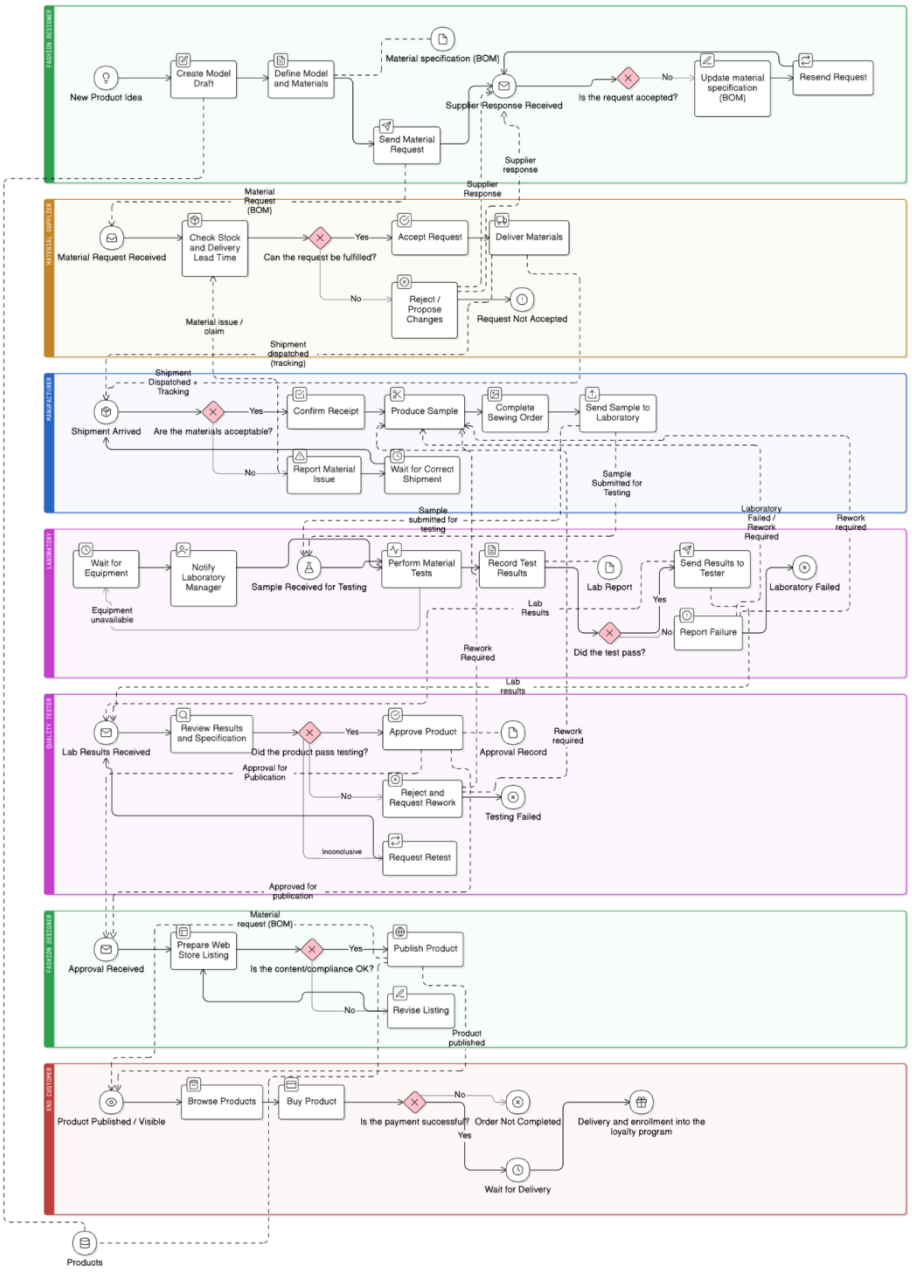


Fig. 2. BPMN diagram for the workflow within the proposed fashion e-commerce DApp

3.3 Sequence Diagram of Application Interactions

The presented sequence diagram (Fig. 3) illustrates a basic happy path scenario of the proposed decentralized application for product traceability and authenticity verification in a fashion e-commerce environment. A happy path scenario refers to an ideal and successful process flow in which all participants perform their activities without errors, rejections, or returns to previous steps. The diagram shows only one complete and successful product journey, from collection and product model creation, through material procurement, production, laboratory and quality validation, to product publication in the online shop, customer purchase, and recording key events in the blockchain and the supporting database.

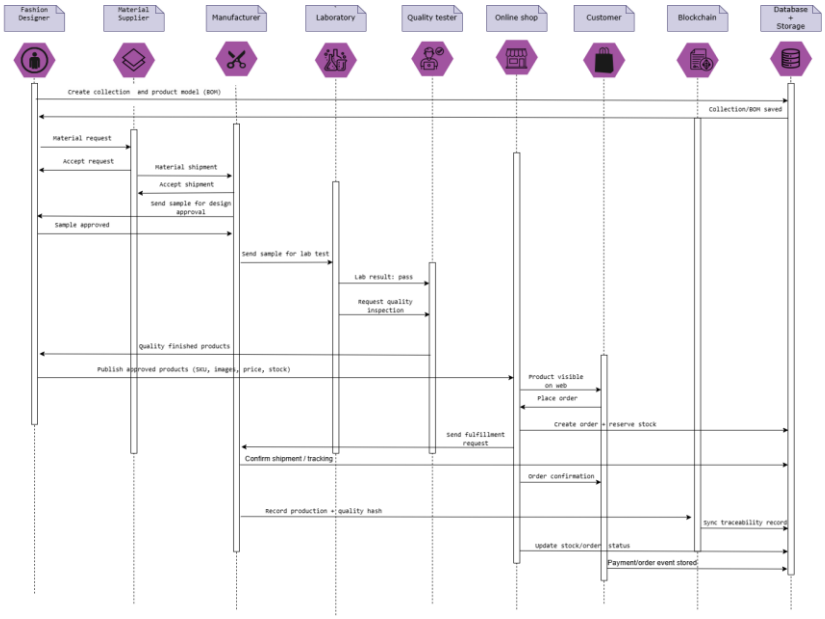


Fig. 3. Sequence diagram of interactions between decentralized application participants

Within this scenario, not all process steps are implemented directly through smart contracts. Smart contracts are used only for those events that represent key business evidence and that must be recorded in a permanent, transparent, and immutable way [24]. They do not function by storing full operational information, but by collecting only the essential aspects of the operations for traceability, verification, and accountability [25]. This selective use of smart contracts is a critical component of the design for small-scale e-commerce fashion contexts. Given that small firms may lack the resources to sustain more complex blockchain infrastructure, this DApp stores just those events where immutability and external verification are necessary. Other data, like product descriptions in their entirety, test reports, pictures made within the company will continue to be stored off-chain. It cuts down on storage fees, makes the smart contract less complex, and facilitates adoption of blockchain traceability.

The process of creating a material request could be managed using the

MaterialRequestContract. This ensures that all the stages of material requests are traceable, from the very beginning of their creation to delivery. The designer's sample approval could be managed with the SampleApprovalContract [26]. The lab verification procedure can be performed using a LabCertificationContract. In this particular scenario, the blockchain saves only the result of verification and hash of laboratory report while the actual report remains stored outside the blockchain [27]. Thus, the verification decision made by the laboratory can be considered both verifiable and unchangeable [28]. In terms of final quality control, a special QualityGateContract should be created, the purpose of which will be registering whether the batch in question is accepted for sale or corrections need to be made. The blockchain stores information about the final quality assurance decision without storing actual documentation in the chain. Moreover, the main order events generated within the online shop can be saved by creating an OnlineShopOrderAnchorContract. The model does not require the entire checkout process to be executed through smart contracts. Instead, some essential events such as the order creation, its status changes, shipping details, or references to payments should be recorded [29].

To summarize, the use of smart contracts is limited to evidence, confirmation, and traces of certain procedures rather than storing detailed documentation within the blockchain [30]. A clear separation occurs between the blockchain level and database/storage level where complete documentation is saved [28].

4 Development of the Decentralized Application for Fashion E-Commerce Supply Chains

4.1 Implementation Environment and Technology Stack

The prototype developed was realized as a decentralized web application that unifies a traditional application architecture and a trust layer provided by a blockchain technology solution. Such a design choice was justified by the necessity to make an implementation practical for day-to-day use in small-scale fashion e-commerce settings and allow selected operations and transactions to be registered on a blockchain. As for the user interface of the prototype, the development was done via React and Vite technologies, offering access based on roles. For the backend and off-chain data layer, Node.js/Express and Supabase technologies were utilized. This subsystem of the architecture stores the operational data necessary for proper functioning of the application, such as product details, product images, collections, inventory data, orders, laboratory-related data, and documentation. Concerning the implementation of the blockchain part of the solution, the smart contracts written in Solidity v.0.8.20 were developed. The implementation and deployment of such were performed via the Remix IDE and the Ethereum Sepolia testnet. MetaMask was chosen as a wallet environment that facilitates blockchain transactions and management of participants' addresses. Communication between the frontend and blockchain layers was provided through the ethers.js library. Regarding the architecture, it can be said that it provides for a hybrid data storage strategy, which

implies that selected data should be registered on-chain, while other data is stored off-chain. Thus, the blockchain serves only for selected confirmations and approvals, transactions, payments, and state changes, but not the entire data of the application.

4.2 Security Considerations and System Limitations

The suggested DApp does not serve as a comprehensive cybersecurity solution that would ensure protection of all elements of the fashion e-commerce ecosystem from threats. Instead, the main security contribution offered by the technology is related to ensuring the integrity, transparency, and verifiability of certain product lifecycle events. Such an approach to building cybersecurity into blockchain technology is justified by the hybrid architecture of the prototype, where the blockchain acts as a trust layer that stores critical pieces of evidence. Data regarding ongoing operations are maintained in a traditional application database off-chain. Consequently, the prototype focuses on securing crucial moments of the product lifecycle, but leaves existing application-level and database-level security solutions untouched [31].

The first security measure implemented via blockchain is related to maintaining the immutability and verifiability of critical evidence, for example, records that correspond to important milestones in the product lifecycle. They include material request creation, material request funding, supplier acceptance, records related to shipment, product approval, laboratory verification references, and customer payment events. Once such evidence is recorded on the blockchain, it becomes tamper-resistant and can be checked against potential inconsistencies using transaction hashes and explorers [31]. Smart contracts contribute to ensuring security at the transaction level by enforcing strict control over the state transition of a particular transaction and the correctness of its parameters. In the material procurement scenario, the DesignerSupplierContract ensures that the state of the request corresponds to its current status in the database (and hence, that it can be either funded or accepted); that only the designer is able to fund the request and that the transaction amount corresponds to what was requested; that only the supplier is able to accept a funded request. The OrderPayment contract performs checks concerning the amount of payment and validity of the order ID prior to storing the record and transferring the transaction funds. Thus, smart contracts allow limiting manual manipulation of important steps and minimizing the probability of inconsistent process execution and unauthorized actions. In supply chains, smart contracts provide for self-enforced transactions, relationship automation, and transaction-wise governance [32].

In contrast to blockchain data, not all information and processes involved in the operation of the DApp are automatically protected by blockchain security measures. Following the hybrid architecture, any sizable and frequently changing data (product descriptions, photos, internal notes, user data, product inventory data, laboratory test reports) are off-chain, meaning that no blockchain layer protection is provided for their safety. While the use of blockchain helps mitigate many risks, several types of threats cannot be prevented by blockchain alone. For instance, blockchain cannot help detect or stop attempts of unauthorized changes, deletion, alteration, and leakage of off-chain data. Similarly, the prototype provides no protection from potential threats

related to user account and private key compromise, incorrect usage of cryptocurrency wallets, and the physical copying and substitution of product identifiers such as QR codes or product tags [31].

4.3 Functionalities of the Developed Decentralized Application

The developed decentralized application serves as a conceptual solution for improving transparency, traceability, and authenticity verification in small-scale fashion e-commerce supply chains. In particular, the app incorporates various participants in the life cycle of fashion products, such as designers, suppliers of materials, manufacturers, laboratory staff, quality controllers, and final consumers. The main objective of the project is to capture and track key events throughout the product lifecycle, from design to post-sale verification. Using blockchain to record the data and storing the data off-chain will ensure data integrity and scalability [33]. The functionality provided by the DApp includes the following activities performed in accordance with real-life fashion supply chain processes: product registration, product development, testing and marketing activity record keeping, sale transaction tracking, and product verification. All functionalities are implemented via user interfaces for different stakeholder roles and linked to respective blockchain transactions for ensuring traceability and immutability of relevant events. The DApp was developed using a DevOps-oriented approach, including continuous integration, automated testing, and iterative release of functional units. Such a development methodology ensures smooth cooperation between developers and deployments as well as enhances system dependability and scalability [16, 34].

The following subsections describe the main functionalities of the developed application and illustrate them through interface screenshots. These functionalities correspond to the processes defined in the BPMN model and system architecture.

Product Registration. Product registration represents the initial stage of the product lifecycle implemented in the DApp. It provides the possibility for the fashion designer to logically organize and register products from the level of fashion collection creation to detailed product specification.

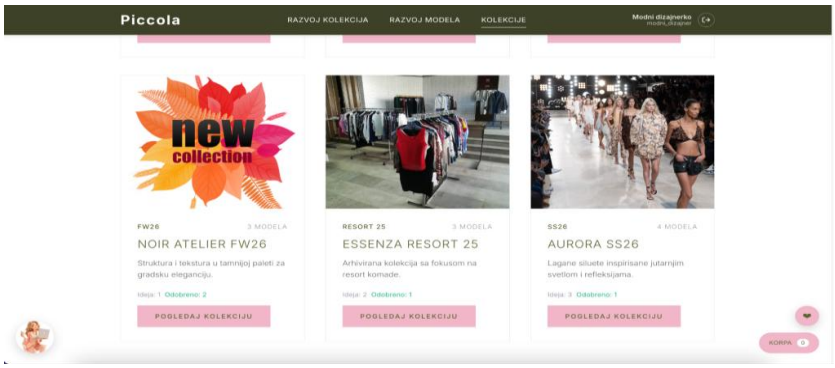


Fig. 4. Overview of fashion collections and product grouping within the DApp

First, the application offers the fashion designer the opportunity to manage fashion collections, i.e., the grouping of products in accordance with seasonal, inspirational, and/or marketing factors. As shown in Fig. 4, the presented interface contains information about all created collections with various product models. Each collection includes information about the product name, its description, and the current stage of development. Each fashion collection contains several products/models, which can be independently registered and controlled by the designer. The registration of fashion products is carried out using the product development interface shown in Fig. 5. Using this interface, the fashion designer defines all features of the product in detail, including conceptual features (inspiration and description of product development), technical properties (material, cuts, sizes), aesthetic characteristics (color palettes and product variants), price information, and internal notes.

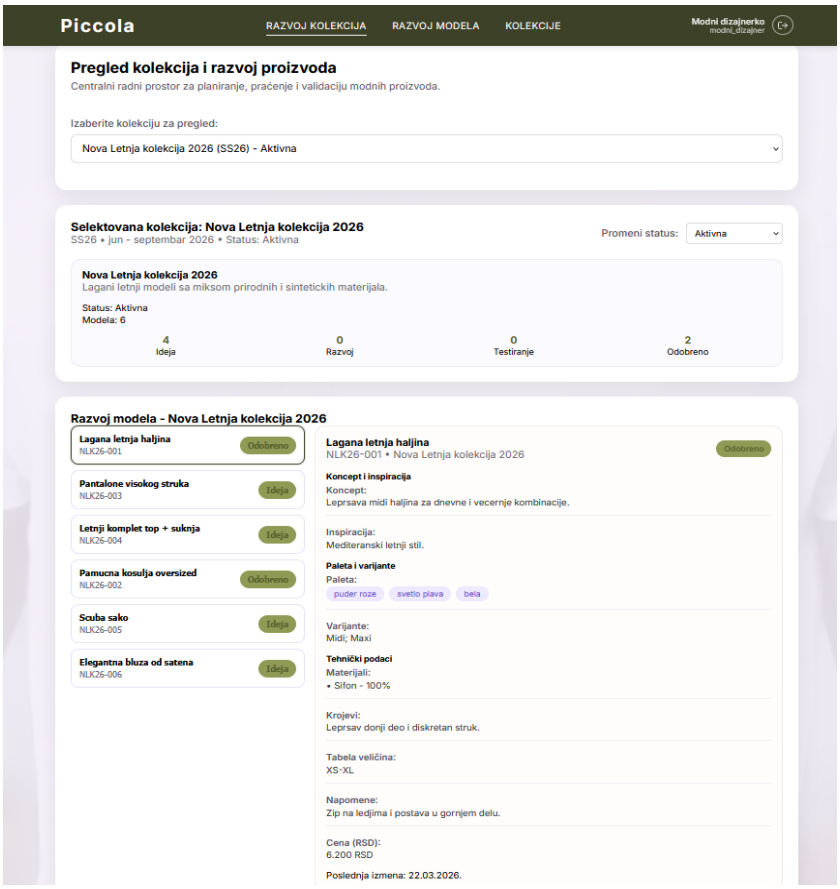


Fig. 5. Product definition interface with detailed attributes and design specifications

All products will be associated with the unique identifier that allows tracking it through all the following stages of the supply chain cycle. During this stage, the products will be in the early stages of their lifecycle, idea and development. The

products are still unknown for the stakeholders outside of the company, as well as end users of the goods. Thus, the product development process will include transitions that allow the transition of the products from idea to validation and approval.

If the product passed all subsequent steps – purchasing of materials, manufacturing process, laboratory tests, and quality control, then the product gets the status of an approved one. The product becomes published on the marketplace page. In this way, the transition of the product from internal process to being known for stakeholders and end consumers happens, and it becomes visible and available at the next step. This step is important as it implies the availability of the required data for tracking and authentication of the products. In terms of the functioning of the DApp, this module will become the initial stage for all the following processes. The registration of the product provides an opportunity to define it, place within the collection, enrich with required information. In such a way, the registration of products provides prerequisites for effective interactions.

Tracking of Product Development and Testing. Product Development and Testing Tracking is one of the key functions of the proposed decentralized application, since it helps track and coordinate any activities associated with the process of developing a product concept into a product.

The next stage following product registration involves defining and procuring necessary materials. For that purpose, the fashion designer creates requests for materials needed for the product according to the previously defined bill of materials. There is an interface for the fashion designer where he or she can set up necessary materials, their quantity, deadline for delivery, and supplier of those materials (Fig.6).

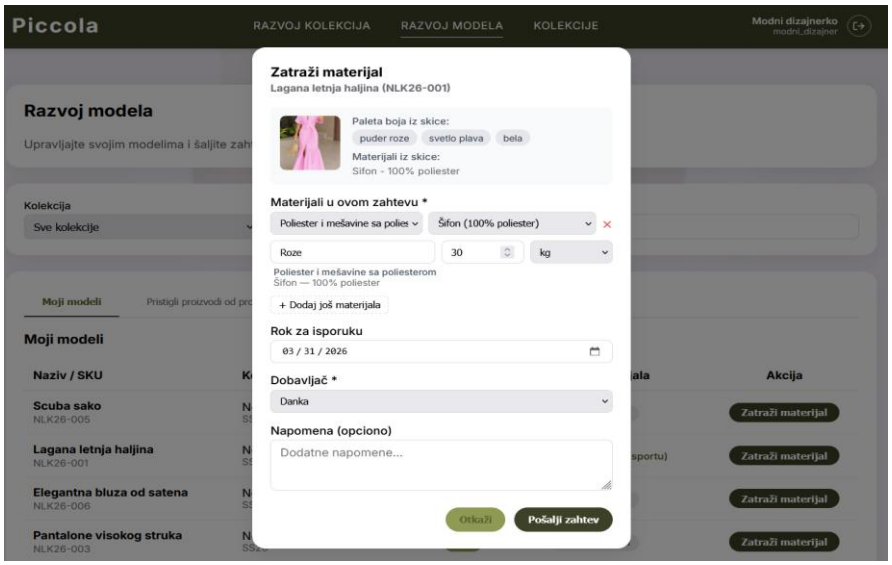


Fig. 6. Material request interface for defining required materials and initiating procurement process.

These requests are further transmitted to the material providers for evaluation based on the current stock availability and feasibility of delivery. The material provider functionalities of the application incorporate functionalities for handling inventory and requests to facilitate seamless coordination of procurement operations. There is an interface through which the material provider manages material stocks (Fig. 7).

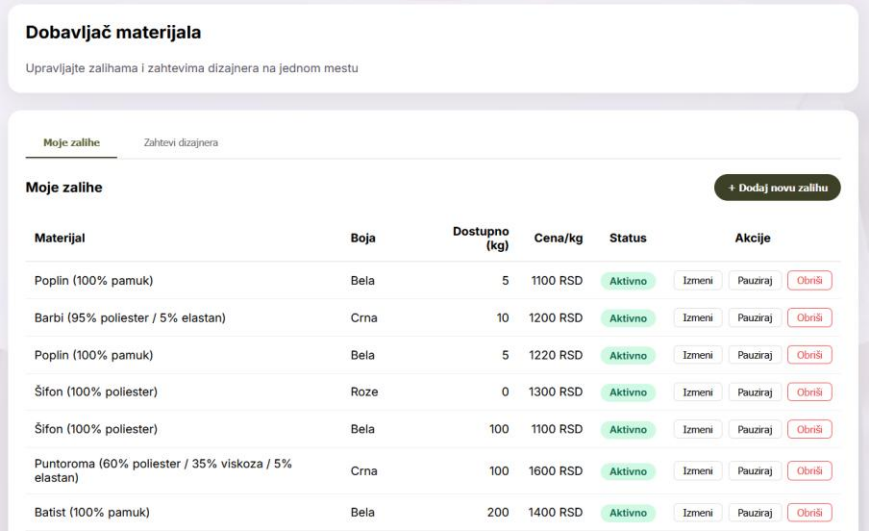


Fig. 7. Supplier inventory management interface with available material stocks

Using this interface, the material provider can track material availability, amount of material, colour, price, and status of material among other information. This facilitates tracking of current resources, thus setting the foundation for responding to procurement requests. Apart from managing the inventory, the system allows suppliers to analyse and process the orders sent by designers. As seen in Fig. 8, the details contained in each order include those regarding the materials required, the quantity required, the deadline for delivery, and the model of the product for which the materials are intended. In addition, the interface highlights the contrast between the order quantity and the current inventory, making it easier for the supplier to assess if he can fulfil the order. From there, he may choose to either accept or make further contact concerning changes that could be made, including modifications in quantity or deadlines.

When the necessary resources are validated and received, the manufacturing process starts. The manufacturer receives the resources and commences manufacturing the product prototype according to the specifications set by the fashion designer.

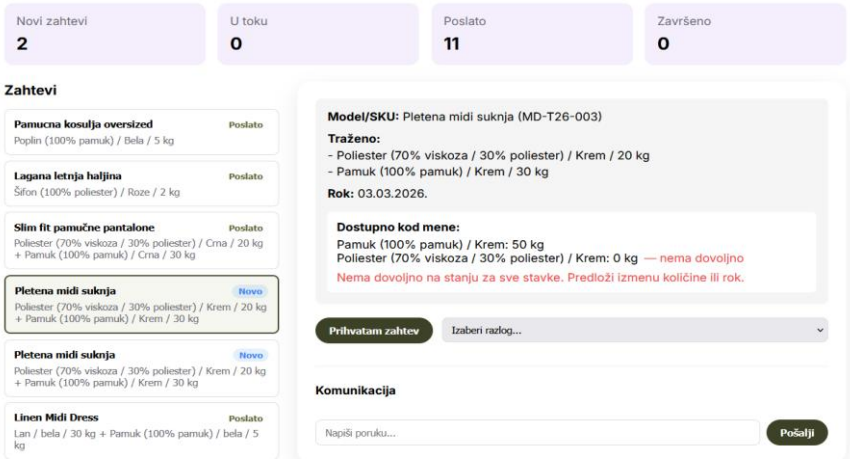


Fig. 8. Supplier interface for reviewing and responding to incoming material requests

Fig. 9 shows that the software offers an interface for handling and validating the received shipments and materials.

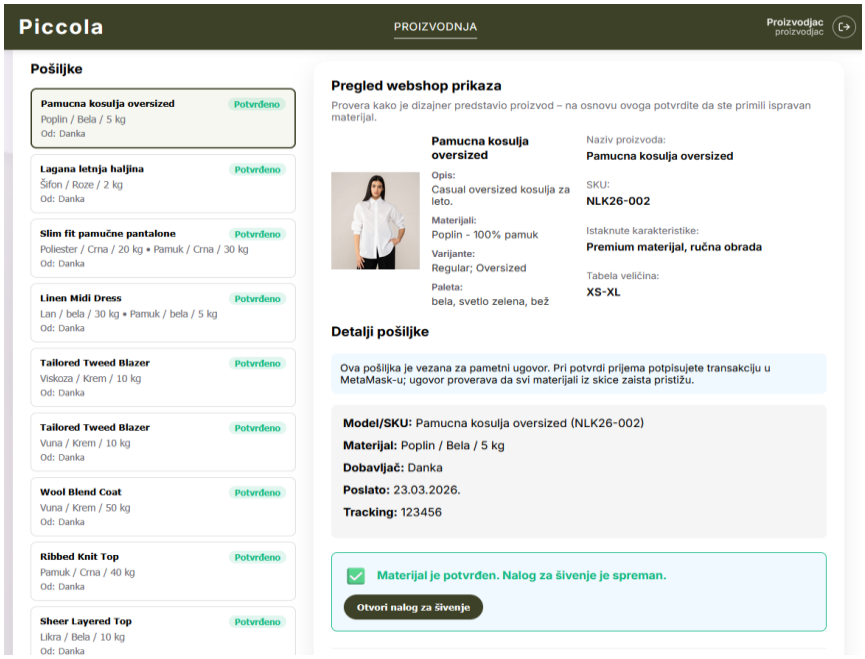


Fig. 9. Manufacturer interface for shipment verification and confirmation of received materials prior to production.

The manufacturer can examine the shipment information such as the model number of the product, the specifications of the materials, the source of the materials, and the dates of shipment and delivery. The software is designed to validate whether the

materials received match the materials expected according to the product specification. One of the important elements of this stage is the confirmation of materials receipt, which is registered in the system and may be connected to a blockchain transaction. The manufacturer guarantees that the received inputs are according to the requirements defined earlier and creates a strong basis for further operations. Thus, confirmation of receipt increases data integrity and avoids any discrepancies related to the production input planning and implementation. Further, the manufacturer starts creating a product sample. Tracking of production is carried out using the application, and thus, all parties can observe how production activities go on and keep track of their progress. After completion of a sample, it is sent to the fashion designer for preliminary verification concerning such aspects as its compliance to the project's design criteria and functionality. Once the designer approves the sample's compliance to all specified criteria, a product moves to the next stage, namely to the testing stage. This step allows managing changes in product's status in a systematic way using the application.

As shown in Fig. 10, the system presents an overview of all products in a collection, together with the current statuses of each product, whether it is an idea, being tested, approved, or sent back for further improvements. Using this application, the designer will have the ability to transfer any product to the next step such as testing, sending it back, or validating the product further. Upon entering the testing stage, the product is then sent for lab tests in which both material composition and physical characteristics of the products are analysed based on the predetermined testing criteria.

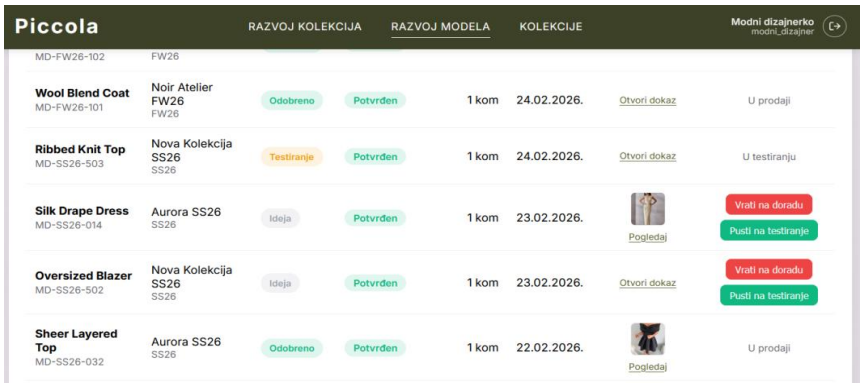


Fig. 10. Product lifecycle status management interface with transitions between development phases.

As shown in Fig. 11, the lab module allows for the execution of test methods that are appropriate in verifying material compositions, assessing product durability, among others. Data generated from these tests will be captured in the system as well. The recorded results of the tests conducted in the lab can be used as a starting point for the quality assurance process that follows. During this stage, the quality tester determines whether the product meets all the required technical and quality criteria.

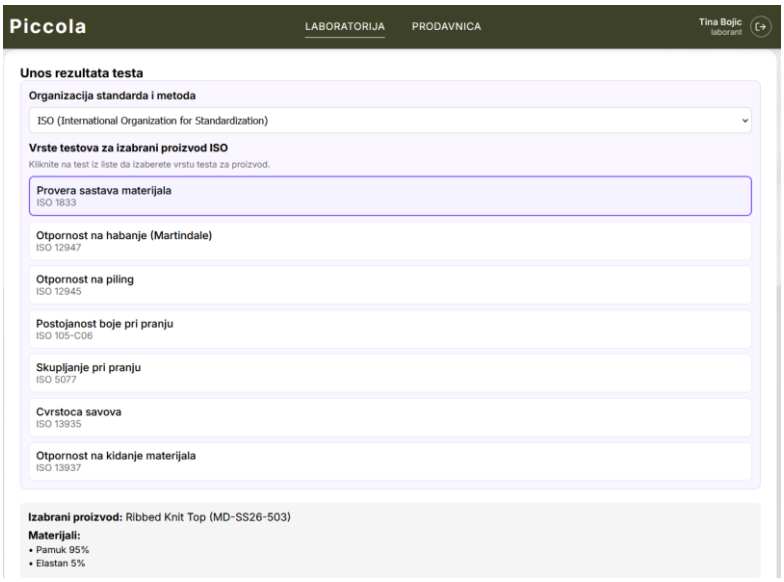


Fig. 11. Laboratory interface for entering and managing product testing results based on standardized procedures.

As illustrated in Fig. 12, the software offers an interface through which one can analyse the information about the tested product.

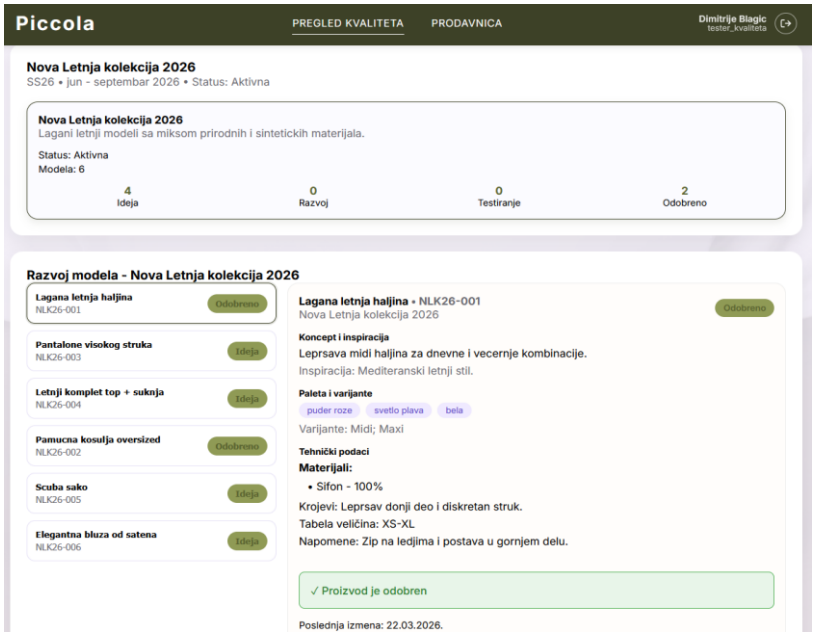


Fig. 12. Quality assurance interface with final product approval and lifecycle status overview

Thus, depending on the results obtained, a certain product can be either allowed to continue the journey through all stages of development or sent back for improvement.

The system stores this information, which guarantees that only products meeting all requirements are allowed to enter the commercialization stage. Throughout the whole period of development and testing, the application ensures a chronological list of all actions performed with regards to each product separately. It allows recording any activity related to requesting materials, the reaction of the supplier, manufacturing process, laboratory test, quality approval of product etc., each step being tied to the product itself and stored in its life cycle history. It will provide an opportunity to trace the product development process and verify all actions taken during the process.

Authenticity Verification. Product authenticity is another vital feature that can be performed through the created DApp. This feature serves as one of the significant advantages of the offered system, as it deals with the problem of transparency and fake products in fashion supply chains. The customer interface enables consumers to get familiar with product-related details such as the design features, material content, and production information.

Other than just being able to get basic details about the product, it allows for checking the authenticity of the products by giving information regarding their recorded life cycle. Every single product comes with its digital ID along with its record, and these records detail everything from its testing process, development stage, and approval. Moreover, these records can even be linked with blockchain-related transactions.

While the entire process of validation is discussed more thoroughly in the next section, it becomes clear that the notion of authenticity allows tying all the above processes, starting from the registration and development of the product up to its sale, into one coherent history of the product which can be verified.

In the proposed decentralized application, each product is provided with a unique digital identity as well as a lifecycle record that customers can view through the product user interface. The blockchain record is not sufficient proof that the physical item that a customer is holding is actually the same product that corresponds to the digital record. This relies, to a great extent, on the accuracy of the product identification device, such as QR code, RFID tag, or NFC tag.

4.4 Recording Transactions Using Smart Contracts

This developed DApp incorporates a series of smart contracts to track and manage important transactions in the supply chain of the fashion e-commerce industry. This series of smart contracts makes up the blockchain component of this system and serves as autonomous agents that implement pre-defined business logic without intermediaries [35]. Transactions on the blockchain for essential milestones like material acquisition, product verification, and client payment become irreversible by employing smart contracts. Every activity done by the parties involved gets verified based on predetermined conditions, and after completion, gets saved in a non-editable format [33].

Other than executing transactions, smart contracts also issue events that provide crucial data related to each transaction that can be used to track the process of operations throughout the whole life cycle of the product [35]. All the parties involved will be able to monitor the whole process from the inception until the end. As for the storage of transactional information, which is vital, it is kept on the blockchain to ensure the integrity of the data; however, when it comes to the storage of operational information, then it will be kept off the blockchain to ensure scalability.

The implemented smart contract architecture reflects interactions between different stakeholders and supports multiple phases of the product lifecycle. The developed smart contracts include:

- DesignerManufacturerContract – controls collaboration between designers and manufacturers.
- DesignerSupplierContract – controls material acquisition between designers and suppliers.
- SupplierManufacturerContract – controls material delivery from suppliers to manufacturers.
- ProductApprovalContract – captures product evaluation and approval procedures.
- OrderPayment – captures customer payments and sales transactions.

Every contract is aimed at capturing the particular type of event in the blockchain. Moreover, two contracts deserve special attention due to their complexity and functionality. These are DesignerSupplierContract, which is responsible for material acquisition, and OrderPayment, responsible for capturing customer payments. The DesignerSupplierContract is used to control collaboration between the fashion designer and the supplier in terms of material acquisition. This contract uses a certain lifecycle of the material request, which includes several states, namely: Pending, Funded, Accepted, Rejected, and Refunded. Such a lifecycle allows tracking the material request state at any moment.

The DesignerSupplierContract begins when the designer makes a material request. Then, the designer is obliged to put the money to the smart contract. Thus, it is guaranteed that there will be funds that should be paid to the supplier. It should be mentioned that the process of funding includes quite restrictive logic requirements. The process of funding includes quite restrictive logic requirement:

```
require(req.status == RequestStatus.Pending,
  "Zahtev nije u Pending statusu");
require(req.designer == msg.sender,
  "Samo dizajner moze deponovati");
require(msg.value == req.totalPriceWei,
  "Iznos mora biti jednak totalPriceWei");
```

The conditions above ensure that:

- the request is in its right status,
- the designer alone deposits money for the request,
- the designer deposits the exact amount of funds required.

This ensures that there will be no financial mistakes in the procurement process.

After the funds have been deposited, the supplier accepts the request. While accepting, further validations are carried out by the contract:

```
require(req.status == RequestStatus.Funded,
  "Zahtev nije u Funded statusu");
require(req.supplier == msg.sender,
  "Samo dobavljac moze prihvatiti");
require(address(this).balance >= amount,
  "Nedovoljno sredstava u ugovoru");
```

Once validated, the contract immediately transfers the funds to the supplier:

```
(bool success, ) = payable(req.supplier).call{value: amount}("");
        require(success, "Transfer ka dobavljacu nije uspeo")
```

These rules will ensure that:

- the supplier does not receive payment before accepting the request,
- the designer should have enough money in the deposit,
- the transaction is executed automatically and transparently.

In case of rejection of a material request, the contract provides for a refund of funds to the designer as another security measure. This contract is in charge of managing the customer payments and registering them on the blockchain. With its help, it becomes possible to perform transactions between the customer and the brand without intermediaries. If the customer decides to make a payment, then the payForOrder function is used. At the initial stage, several simple validations are carried out, whether the payment value is greater than zero and whether the order id exists.

If these conditions are satisfied, then the contract emits an event with information about the main aspects of the transaction – it can be called a record on the blockchain. In it, the id of the order, the address of the person making the payment, the payment amount, and the moment in time are stored.

Afterwards, the payment itself goes into the wallet of the brand owner and is checked for successful transfer. If it was unsuccessful for some reason, then the transaction is reverted. The complete payment logic is shown below:

```
function payForOrder(bytes32 orderId) external payable {
        require(msg.value > 0, "Iznos mora biti veci od 0");
        require(orderId != bytes32(0), "orderId ne sme biti prazan");


        emit OrderPaid(orderId, msg.sender, msg.value, block.timestamp);
```

```
        (bool sent, ) = payable(owner).call{value: msg.value}("");
        require(sent, "Slanje vlasniku brenda nije uspelo");
    }
```

Such a strategy guarantees that all client payments are conducted with maximum security, documented efficiently, and settled instantly without compromising on the integrity of financial transactions.

4.5 Transactions Between Supply Chain Stakeholders

The proposed DApp enables the transaction between multiple parties taking part in the fashion e-commerce supply chain such as the designer, supplier, manufacturer, and customer. These transaction activities take place via the use of smart contracts within the blockchain network [33, 36]. While there may be many interactions between different types of stakeholders in this system, this paper will discuss the process flow of transactions between the fashion designer and the supplier of raw materials as an example. The transaction process starts when the designer creates a material request and deposits the required funds into the smart contract. The supplier is guaranteed to get paid in advance for accepting the request. Fig. 13 below shows the confirmation of the transaction on behalf of the designer using a MetaMask blockchain wallet.

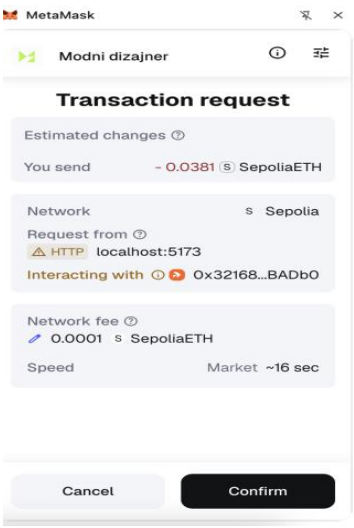


Fig. 13. Transaction confirmation interface showing outgoing payment from the designer

After the transaction is confirmed, the funds are locked in the smart contract. Having examined the transaction, the supplier approves it, and therefore the process of transferring money starts. The smart contract itself sends the sum of money stored earlier to the supplier's wallet (Fig. 14).

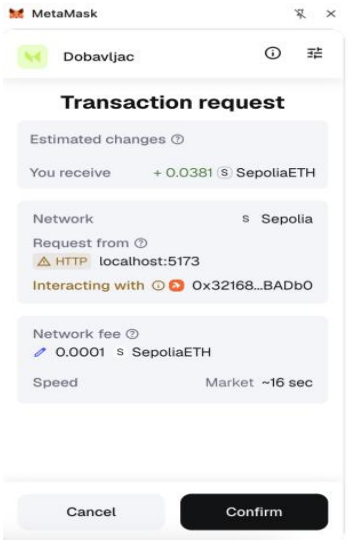


Fig. 14. Wallet interface showing incoming payment to the supplier

After the execution of the transaction, it will be permanently stored on the blockchain. As shown in Fig. 15, the transaction details are available via blockchain explorers like Etherscan, which provides information like the transaction hash, sender's address, receiver's address, amount sent, timestamp, and confirmation status.

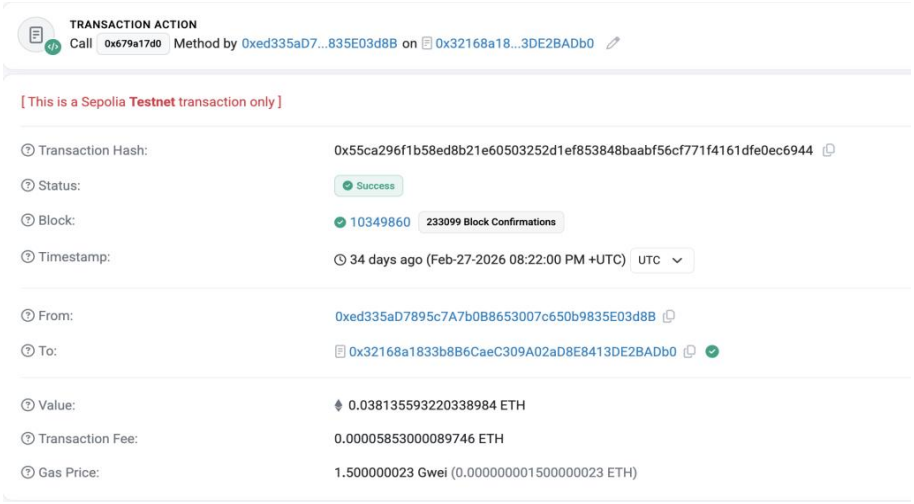


Fig. 15. Blockchain transaction record displayed in Etherscan

Through this procedure, transparency is guaranteed on both ends of the transaction, that is, the sending of the transaction by the designer and the receiving of the transaction by the supplier. Automation of the transaction using smart contracts makes intermediation unnecessary and reduces risks of conflict among stakeholders.

Regarding the system point of view, these processes illustrate how blockchain technology provides security, automation, and tracking of financial transactions in the supply chain.

4.6 Customer Interface and Product Authenticity Verification

The DApp provides a customer-oriented interface that allows customers to view product information, verify authenticity, and review the product lifecycle. As shown in Fig. 16, the products are displayed in the market interface, where the customers can see some important details about the products such as product name, price, size, and materials. The interface serves as the point of engagement for the customers who visit this site to buy products [37].

The lifecycle perspective involves critical activities like the creation of the product, acquisition of materials, production, testing, and the release of the product in the market. Every activity is timestamped with respect to the concerned stakeholder, thereby giving a timeline of how the product was developed and also aiding its authenticity. Certain lifecycle activities are added to the blockchain for immutability. Every occurrence is time-stamped and attributed to the concerned stakeholder, thus offering a clear picture of the product's evolution process. The feature acts as the basis for establishing the authenticity of the product. The mechanism guarantees that the product undergoes all necessary processes of production and testing through providing the consumer with legitimate information about its lifecycle. Selected processes during the lifecycle are logged into the blockchain, making sure that the important information remains immutable.

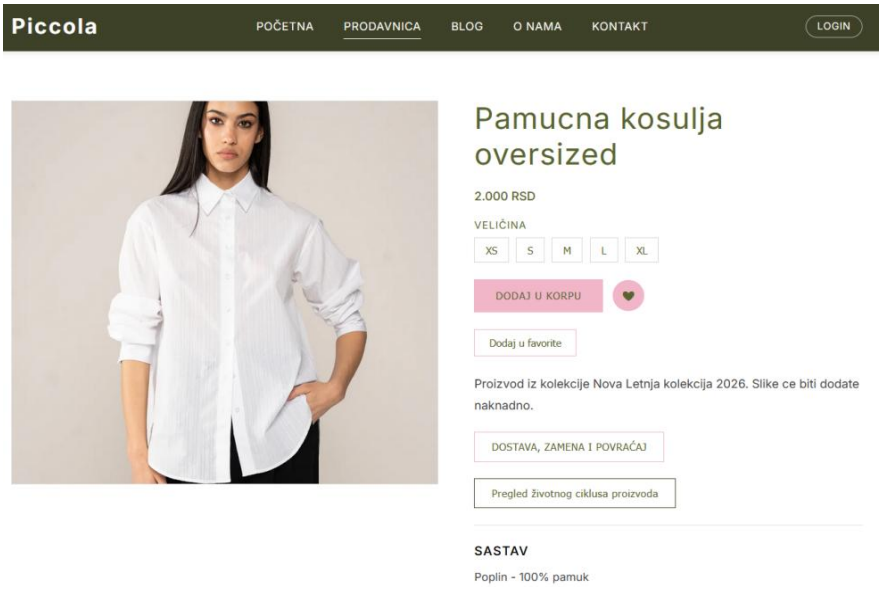


Fig. 16. Product display interface with detailed product information available to customers

Through this approach, the DApp offers a very high degree of transparency and trust that enables the customers to buy items as well as ensure their source, authenticity, and genuineness. This helps reduce counterfeiting in the fashion e-commerce system.

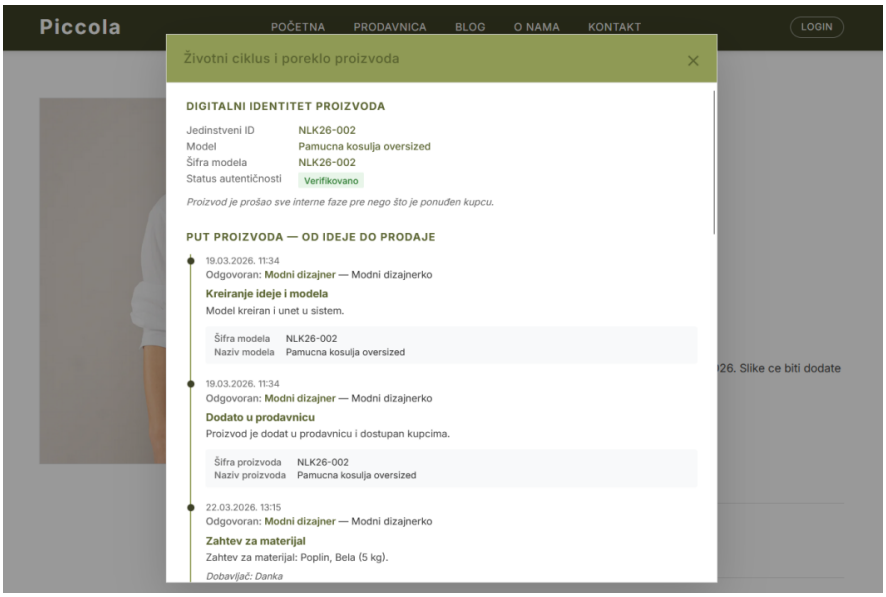


Fig. 17. Product lifecycle view showing chronological record of events and digital identity for authenticity verification.

4.7 Prototype Evaluation

To provide an initial empirical assessment of the created DApp prototype, its major operations related to blockchain were evaluated in the context of deployment costs, gas usage by the primary state-changing functions of the main smart contracts, and transaction latency for the flow of material procurement. Read-only functions and helpers were excluded from the analysis of the gas usage since they do not create any blockchain transaction when invoked from the frontend. The purpose of this evaluation was not to conduct extensive benchmarking, but to demonstrate the feasibility of the presented approach to implementing a blockchain-powered DApp in a small fashion e-commerce scenario.

To test the implemented smart contracts, a synthetic scenario for the small fashion e-commerce supply chain was designed with one fashion designer/brand owner, one material supplier, one manufacturer, one laboratory role, one quality tester, and several customer orders. These activities include product registration, creation of material requests, funding of such requests, accepting such requests, confirmation of shipment, product approval, and payments from customers. This scenario is a simplified representation of a real-world fashion e-commerce supply chain where there is only a few participants and not all lifecycle events are registered on the

blockchain.

The deployment costs for each implemented smart contract were obtained based on transaction receipt information retrieved from the Ethereum Sepolia testnet. Considering that the cost evaluation is performed in a test environment, transaction fees are denominated in test ETH, and they serve for comparing deployment and execution costs. Table 2 summarizes the amount of deployment gas consumed and transaction fees for each smart contract used in the created prototype. Thus, Table 2 provides an estimation of the initial blockchain cost associated with the deployment of the most important parts of the smart contracts' layer.

Table 2. Deployment cost of the implemented smart contracts on the Sepolia testnet

Smart contract	Network	Deployment gas used	Transaction fee (test ETH)
DesignerSupplierContract	Sepolia	1,971,657	0.005107108649728482
SupplierManufacturerContract	Sepolia	1,694,161	0.004310935350127742
DesignerManufacturerContract	Sepolia	1,189,910	0.00306525699509631
ProductApproval	Sepolia	2,087,858	0.005364402400253976
OrderPayment	Sepolia	550,571	0.001389635706548565

As indicated in Table 2, the deployment gas cost varies, depending on the level of contract logic complexity that is being deployed. This means that the highest deployment cost was for the product approval contract since it has the most complex validation logic, whereas the order payment contract registered the lowest deployment cost due to its relatively simple payment registration and forwarding mechanism. It is necessary to note that these numbers are not related to production costs, as they were obtained during evaluations on the Sepolia testnet.

Once deployed, the gas usage for the main state-changing functions was estimated. These functions were selected based on their capability to perform relevant lifecycle actions, by confirming approvals or payments and recording state changes on the blockchain. As presented in Table 3, the gas required for performing operations as well as the cost of performing these transactions on the Ethereum network is provided. In addition, the latency value is provided for the material procurement flow, as it is the primary blockchain-enabled transaction flow in the application.

Table 3. Gas consumption and transaction latency of the main smart contract operations on the Sepolia testnet

Contract	Function	Purpose	Gas used	Transaction fee	Latency (s)
Designer-Supplier	`createRequest`	Material request creation	275,133	0.000767028559253781	13.38
Designer-Supplier	`fundRequest`	Funding of material request	39,020	0.000099853385 1327	5.26
Designer-Supplier	`acceptRequest`	Supplier acceptance and fund transfer	46,308	0.000116993440318656	14.13
Supplier-Manufacturer	`createShipment`	Shipment record creation	392,266	0.000970396354 79584	-
Supplier-Manufacturer	`acceptShipment`	Shipment acceptance	52,631	0.000135899584 088131	-
Product-Approval	`approveProduct`	Product quality approval	100,053	0.000262057392 84528	-
Order-Payment	`payForOrder`	Customer payment	36,043	0.000093690609 058795	-

As shown in Table 3, shipment creation demanded the largest amount of gas among the assessed procedures since it keeps shipments, material hashes, and material lines on the blockchain. In relation to the procurement of materials, creating material requests was more costly in terms of gas than funding and acceptance. This was the case since the former procedure involved the initiation and storage of data about materials. Funding and acceptance were less gas intensive, as they only updated the status of the request and performed transactions.

Latency measures are related to the confirmation period on the Sepolia testnet and could depend on the testnet conditions. It measures pertain exclusively to the material procurement flow, as it was chosen as the main transaction workflow of the prototype. Generally, it can be concluded that the implemented DApp can perform all necessary blockchain transactions on the Sepolia testnet at reasonable costs of execution for a prototype application. In this regard, the hybrid approach in which only some lifecycle events get stored on the blockchain while operational data stays off-chain proved to be relevant [38].

5 Conclusion

This paper addresses key limitations in existing fashion retail supply chain processes, particularly the lack of transparency and reliable verification of product origins,

through the development of a DApp. The proposed solution enables the recording of product lifecycle events and related business activities in a tamper-resistant and accessible ledger, tailored to the needs of smaller supply chain environments. The system makes use of blockchain and smart contracts for the sake of integrity and automation, but at the same time takes into account the practical considerations like costs and scalability.

The innovation behind this proposal is not in creating another blockchain-based provenance solution or platform, but in adopting the former for specific operational conditions of small fashion e-commerce supply chains. The proposed approach includes a reduced number of stakeholders, modular smart contracts, selective transactions and records on the blockchain ledger, off-chain storage of detailed documentation, and separate user interfaces for particular stakeholders involved.

A disadvantage of the prototype is that the AI layer discussed as a part of the solution architecture was neither developed nor empirically tested in the context of this prototype. The focus is on the recording of major lifecycle events through the blockchain, smart contract execution, and product authentication functionality. Therefore, any AI-based solutions, such as demand forecasting, anomaly detection, quality predictions, and decision-making support should be seen as a part of future work on this project. It can be said that the further work should involve the testing of the proposed prototype in the real-world environment, along with assessing its transaction costs, scalability, usability, and stakeholder acceptance. In addition, the future research can be conducted on implementing and assessing the AI layer, as well as using IoT technology to ensure better traceability of products in the physical-digital world.

Acknowledgments. This research was supported by the University of Belgrade – Faculty of Organizational Sciences and, in part, by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia through institutional funding (grant number: 200151). The authors would like to acknowledge that ChatGPT was used only for language editing and grammar improvement. Full responsibility for the originality, accuracy, and integrity of the content remains with the authors.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. https://www.zotero.org/google-docs/?ell6wO Tam, F.Y.: Blockchain technology adoption in the fashion supply chain: a review on the practices of fashion brands. *Journal of Global Operations and Strategic Sourcing* 18(4), 782–808 (2025). https://doi.org/10.1108/JGOSS-01-2025-0003
2. https://www.zotero.org/google-docs/?ell6wO Badhwar, A., Islam, S., Tan, C.S.L.: Exploring the potential of blockchain technology within the fashion and textile supply chain with a focus on traceability, transparency, and product authenticity: A

systematic review. *Frontiers in Blockchain*, 6, Article 1044723 (2023). https://doi.org/10.3389/fbloc.2023.1044723

3. https://www.zotero.org/google-docs/?ell6wO OECD, European Union Intellectual Property Office: Mapping Global Trade in Fakes 2025: Global Trends and Enforcement Challenges. Illicit Trade (2025). https://doi.org/10.1787/94d3b29f-en
4. https://www.zotero.org/google-docs/?ell6wO Sovtić, D., Trpkov, A., Radenković, M., Popović, S., Labus, A.: Examining Readiness to Buy Fashion Products Authenticated with Blockchain. *Journal of Theoretical and Applied Electronic Commerce Research* 20, 119 (2025). https://doi.org/10.3390/jtaer20020119
5. https://www.zotero.org/google-docs/?ell6wO Dolgui, A., Ivanov, D., Sokolov, B.: Reconfigurable supply chain: the X-network. *International Journal of Production Research* 58(13), 4138–4163 (2020). https://doi.org/10.1080/00207543.2020.1774679
6. https://www.zotero.org/google-docs/?ell6wO Kim, H.M., Laskowski, M.: Toward an ontology-driven blockchain design for supply-chain provenance. Intelligent Systems in Accounting, Finance and Management 25(1), 18–27 (2018). https://doi.org/10.1002/isaf.1424
7. https://www.zotero.org/google-docs/?ell6wO Dutta, P., Choi, T.-M., Somani, S., Butala, R.: Blockchain technology in supply chain operations: Applications, challenges and research opportunities. *Transportation Research Part E: Logistics and Transportation Review* 142, 102067 (2020). https://doi.org/10.1016/j.tre.2020.102067
8. https://www.zotero.org/google-docs/?ell6wO Agrawal, S., Tiwari, S., Singh, R., Kumar, K.: The impact of blockchain technology on supply chain efficiency: A review. E3S Web of Conferences, 552, 01013 (2024). https://doi.org/10.1051/e3sconf/202455201013
9. https://www.zotero.org/google-docs/?ell6wO Saberi, S., Kouhizadeh, M., Sarkis, J., Shen, L.: Blockchain technology and its relationships to sustainable supply chain management. International Journal of Production Research 57(7), 2117–2135 (2019). https://doi.org/10.1080/00207543.2018.1533261
10. https://www.zotero.org/google-docs/?ell6wO Wannenwetsch, K., Ostermann, I., Priel, R., Gerschner, F., Theissler, A.: Blockchain for supply chain management: A literature review and open challenges. Procedia Computer Science 225, 1312–1321 (2023). https://doi.org/10.1016/j.procs.2023.10.119
11. https://www.zotero.org/google-docs/?ell6wO Casino, F., Dasaklis, T.K., Patsakis, C.: A systematic literature review of blockchain-based applications: Current status, classification and open issues. Telematics and Informatics 36, 55–81 (2019). https://doi.org/10.1016/j.tele.2018.11.006
12. https://www.zotero.org/google-docs/?ell6wO Xu, X., Weber, I., Staples, M.: Architecture for blockchain applications. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-03035-3
13. https://www.zotero.org/google-docs/?ell6wO Chen, Y., Bellavitis, C.: Blockchain disruption and decentralized finance: The rise of decentralized business models. Journal of Business Venturing Insights 13, Article e00151 (2020). https://doi.org/10.1016/j.jbvi.2019.e00151
14. https://www.zotero.org/google-docs/?ell6wO Hindarto, D., Alim, S., Hendrata, F.: Uncovering blockchain's potential for supply chain transparency: Qualitative study on the fashion industry. Sinkron: Jurnal dan Penelitian Teknik Informatika, 8(2), 1107–1115 (2024). https://doi.org/10.33395/sinkron.v8i2.13590
15. https://www.zotero.org/google-docs/?ell6wO Shaik, V.A., Malik, P., Singh, R., Gehlot, A., Juyal, A., Ghafoor, K., Shrestha, S.: Implementation of digitalized

technologies for fashion industry 4.0: Opportunities and challenges. Scientific Programming 2022, Article ID 7523246, 1–17 (2022). https://doi.org/10.1155/2022/7523246
16. https://www.zotero.org/google-docs/?ell6wO Radenković, M., Despotović-Zrakić, M., Bogdanović, Z., Barać, D., Labus, A., Simić, M.: *Razvoj softvera orijentisanog na procese, odabrana poglavlja: DevOps, mikroservisi, kontejneri*, 2nd revised and expanded edn. (2024)
17. https://www.zotero.org/google-docs/?ell6wO Bullón Pérez, J.J., Queiruga-Dios, A., Gayoso Martínez, V., Martín del Rey, Á.: Traceability of ready-to-wear clothing through blockchain technology. Sustainability, 12(18), Article 7491 (2020). https://doi.org/10.3390/su12187491
18. https://www.zotero.org/google-docs/?ell6wO Francisco, K., Swanson, D.: The supply chain has no clothes: Technology adoption of blockchain for supply chain transparency. Logistics, 2(1), Article 2 (2018). https://doi.org/10.3390/logistics2010002
19. https://www.zotero.org/google-docs/?ell6wO Tripathi, G., Nautiyal, V., Ahad, M., Feroz, N.: Blockchain Technology and Fashion Industry: Opportunities and Challenges. In: *Proceedings of International Conference on Intelligent Computing and Smart Communication 2020*, pp. 193–204. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-69395-4_12
20. https://www.zotero.org/google-docs/?ell6wO Lee, I., Mangalaraj, G.: Big data analytics in supply chain management: A systematic literature review and research directions. Big Data and Cognitive Computing, 6(1), Article 17 (2022). https://doi.org/10.3390/bdcc6010017
21. https://www.zotero.org/google-docs/?ell6wO Kechagias, E.P., Gayialis, S.P., Konstantakopoulos, G.D., Papadopoulos, G.A.: An application of a multi-criteria approach for the development of a process reference model for supply chain operations. Sustainability 12(14), Article 5791 (2020). https://doi.org/10.3390/su12145791
22. https://www.zotero.org/google-docs/?ell6wO Gayialis, S.P., Kechagias, E.P., Papadopoulos, G.A., Panayiotou, N.A.: A business process reference model for the development of a wine traceability system. Sustainability 14(18), Article 11687 (2022). https://doi.org/10.3390/su141811687
23. https://www.zotero.org/google-docs/?ell6wO Tomaskova, H., Maresova, P., Penhaker, M., Augustynek, M., Klimova, B., Fadeyi, O., Kuca, K.: The business process model and notation of open innovation: The process of developing medical instrument. *Journal of Open Innovation: Technology, Market, and Complexity*, 5(4), Article 101 (2019). https://doi.org/10.3390/joitmc5040101
24. https://www.zotero.org/google-docs/?ell6wO Tokkozhina, U., Martins, A.L., Ferreira, J.C.: Use of blockchain technology to manage the supply chains: Comparison of perspectives between technology providers and early industry adopters. Journal of Theoretical and Applied Electronic Commerce Research 17(4), 1616–1632 (2022). https://doi.org/10.3390/jtaer17040082
25. https://www.zotero.org/google-docs/?ell6wO Taherdoost, H.: Smart contracts in blockchain technology: A critical review. Information 14(2), Article 117 (2023). https://doi.org/10.3390/info14020117
26. https://www.zotero.org/google-docs/?ell6wO Kshetri, N.: Blockchain's roles in meeting key supply chain management objectives. International Journal of Information Management 39, 80–89 (2018). https://doi.org/10.1016/j.ijinfomgt.2017.12.005
27. https://www.zotero.org/google-docs/?ell6wO Agrawal, T.K., Kumar, V., Pal, R., Wang, L., Chen, Y.: Blockchain-based framework for supply chain traceability: A case

example of textile and clothing industry. *Computers & Industrial Engineering* 154, Article 107130 (2021). https://doi.org/10.1016/j.cie.2021.107130

28. https://www.zotero.org/google-docs/?ell6wO Tomić, M., Sovtić, D., Trpkov, A., Rodić, B., Labus, A.: Blockchain-based healthcare ecosystem. *E-Business Technologies Conference Proceedings* 3, 203–210 (2023)
29. https://www.zotero.org/google-docs/?ell6wO Sitorus, S.L., Safariningsih, R.T.H., Karsa, A.H.A.N., Komara, M.A., Evans, R.: Optimizing smart contracts and blockchain for sustainable digital fashionpreneurship. *Blockchain Frontier Technology*, 5(1), 37–48 (2025). https://doi.org/10.34306/bfront.v5i1.814
30. https://www.zotero.org/google-docs/?ell6wO Trpkov, A., Sovtić, D., Tomić, M., Labus, A., Rodić, B.: Stakeholders' readiness for adopting blockchain in the fashion industry. *Facta Universitatis, Series: Electronics and Energetics* 37(1), 1–28 (2024). https://doi.org/10.2298/FUEE2401001T
31. https://www.zotero.org/google-docs/?ell6wO Karaduman, Ö., & Gülhas, G. (2025). Blockchain-enabled supply chain management: A review of security, traceability, and data integrity amid the evolving systemic demand. *Applied Sciences*, 15(9), 5168
32. https://www.zotero.org/google-docs/?ell6wO Groschopf, W., Dobrovnik, M., & Herneth, C. (2021). Smart contracts for sustainable supply chain management: Conceptual frameworks for supply chain maturity evaluation and smart contract sustainability assessment. *Frontiers in Blockchain*, 4, 506436
33. https://www.zotero.org/google-docs/?ell6wO Trpkov, A., Sovtić, D., Labus, A.: Methodology for introducing blockchain in e-commerce supply chain. In: Unlocking the Hidden Potentials of Organization Through Merging of Humans and Digitals. *Springer*, Cham (2025). https://doi.org/10.1007/978-3-032-08093-6_2
34. https://www.zotero.org/google-docs/?ell6wO Sovtić, D., Trpkov, A., Radenković, M., Jevtić, M., Labus, A., Marković, L.: Methodological approach to applying DevOps for the development of a decentralized application. In: *6th International Workshop on Engineering Technologies and Computer Science* (EnT). pp. 1–7. IEEE (2025)
35. https://www.zotero.org/google-docs/?ell6wO Stefanović, M., Pržulj, Đ., Stefanović, D., Ristić, S., Čapko, D.: The proposal of new Ethereum request for comments for supporting fractional ownership of non-fungible tokens. *Computer Science and Information Systems* 20(3), 1133–1155 (2023). https://doi.org/10.2298/CSIS230127038S
36. https://www.zotero.org/google-docs/?ell6wO Cho, Y., Lee, C.: The effects of process innovation and partnership in SCM: Focusing on the mediating roles. *Computer Science and Information Systems* 21(2), 453–472 (2024).
37. https://www.zotero.org/google-docs/?ell6wO Petković, G., Dokić, A., Vasić, V.: How does organisational ICT implementation affect e-commerce potential? Empirical evidence from Serbia. Management: *Journal of Sustainable Business and Management Solutions in Emerging Economies* 29(1), 21–40 (2024). https://doi.org/10.7595/management.fon.2021.0037
38. https://www.zotero.org/google-docs/?ell6wO Motla, R., Saxena, N., Kumar, A., Sana, S.S.: Enhancing supply chain efficiency through blockchain technology: A centralized multidistributor and multi-retailer approach. *Yugoslav Journal of Operations Research* 35(3), 483–500 (2024). https://doi.org/10.2298/YJOR231115026M

Original Research Article *https://doi.org/10.5281/zenodo.23473299*

Hybrid Risk Modeling Framework for Comparing PoW and PoS Consensus Mechanisms

Jovan Petrović[1][0009-0000-7951-8432], Dušan Mitrović*[1][0000-0001-5299-9707], Bojan Marčeta[1][0000-0003-2459-8302]

[1] University of Belgrade – Faculty of Organizational Sciences
*Corresponding author, e-mail: dusan.mitrovic.fon@gmail.com

Abstract. Consensus algorithms serve as the foundational architecture for all modern blockchain networks. By transitioning beyond traditional peer-to-peer networking, these protocols establish enhanced security frameworks across decentralized systems. They enable dispersed network users to agree on state transitions on a blockchain, even when acting maliciously toward one another. Many of the most popular consensus algorithms (e.g., Proof-of-Work [PoW] vs. Proof-of-Stake [PoS]) have been extensively compared qualitatively in the literature, while existing quantitative approaches are typically fragmented and do not provide a unified framework for systematically comparing their resilience against different types of attacks from both technical and economic perspectives. In order to compare the advantages of PoW and PoS using a structured risk assessment model, this study offers a hybrid modeling approach. The cost of carrying out an attack, the probability of successfully completing an attack, the time it takes to detect a successful attack, and the degree of difficulty encountered during recovery from an attack are the four main criteria that constitute an integrated Risk Score function. The proposed framework enables evaluation of security properties of these protocols under different attack scenarios, once the relevant parameters are defined for those scenarios. The corpus of knowledge pertaining to cryptoeconomics and incentives, as well as the most recent sustainability models for blockchain governance and PoS security modeling, were used to qualitatively assign model parameters. In order to identify relative vulnerabilities between consensus techniques, the resulting risk scores will be graphically represented and contrasted.

Keywords: Blockchain, Consensus Mechanisms, Risk Analysis

1 Introduction

Risk assessment in blockchain systems has been shown to be directly dependent on consensus mechanisms, which were demonstrated in early blockchain architectures [10],[11]. As previously documented, Proof-of-Work (PoW) and Proof-of-Stake (PoS) have been analyzed thoroughly through formal analysis of their security properties and systematic reviews [1], [13]; similarly, literature has provided wide documentation of the operational tradeoffs between PoW and PoS [9]. However, these two protocols are

often analyzed independently with respect to cryptographic or economic assumptions alone. Therefore, network architects and researchers need a standardized methodology that provides reliable comparisons between different protocols under critical situations.

Recent development of "MEV-driven" strategies for malicious behavior [5] shows how an attacker can be successful not only because they find a technical bypass, but also due to a combination of cost, time and the ability of the network to recover from attacks in a timely fashion [12]. Consequently, evaluating network resilience requires moving beyond isolated metrics toward an integrated risk framework.

Therefore, we provide a framework to quantify security in blockchain systems using a structured risk model. Specifically, the core contribution is the Risk Score (RS), a composite function designed to provide standardized comparability across all consensus mechanisms. Our approach is different from traditional analysis in that it focuses on four measurable dimensions:

1. **Economic Hurdle:** The capital needed to manipulate the consensus of the given network.
2. **Execution Probability:** The likelihood for an attack to succeed given the network's current state.
3. **Confirmation Delay:** The time difference between the start of an exploit and the response of the network.
4. **Recovery Difficulty:** The network's ability to restore the ledger to a valid state after a compromise.

By applying this model to the PoW and PoS mechanisms, we provide a comparative basis for spotting vulnerabilities which are obscured in standard qualitative reviews. It is crucial to emphasize that this framework is illustrative in nature, designed to provide a structural scoring methodology rather than a definitive empirical security evaluation. This model is hybrid because it combines economic incentive-based variables and security characteristics, all within a single quantitative framework. The proposed framework does not just describe how protocols work, but rather provides a tool for measuring how they can potentially fail.

2 Related Work

The majority of recent studies to determine whether consensus mechanisms found in blockchain systems are secure compared to one another, have been comparative surveys which classify all available consensus models as well as evaluate them using parameters such as performance, security, power consumption, etc. Andrews et al. analyzed the trade-off between the operational aspects and the security of both Proof-of-Work (PoW), and Proof-of-Stake (PoS) from a security point of view [2]. Similar in nature, Shaikh et al. provided a comprehensive review of these two protocols. The authors also identified what they refer to as "general attack surfaces" with respect to different architectures [9]. Arafat expanded upon this perspective, and provided an expansive survey of consensus protocols focused on energy costs, and transition toward

alternative mechanisms [3]. As such, collectively, these studies provide a solid qualitative base for cross-protocol comparisons of security.

In addition, most of the work done in this area focuses on understanding how the economic/incentive behaviors of individuals participating in blockchain-based systems apply particularly to PoS based systems [15], [16]. Bruno et al. investigated specific incentive structures with regard to sustainability and long-term stability under rational behavior assumptions [4]. They demonstrated how economic incentives influence validators' behavior and ultimately the security status of decentralized systems. On a more formal basis, Li et al. developed a cryptoeconomic framework for modeling and evaluating PoS systems that incorporate economic incentives directly into security assumptions at the protocol-level [6]. As such, these contributions demonstrate how critical it is to include economic modeling when determining the security associated with consensus mechanisms.

Prominent metrics include heuristics like the Nakamoto coefficient, which quantifies network decentralization by calculating the minimum number of independent entities needed to execute a successful disruption under given operational constraints [7]. Other concepts such as cost-of-corruption and resilience-based metrics have also been introduced in literature to quantify the economic viability of attacks [2], [6]. While these approaches represent significant steps towards quantification, they typically only provide a narrow scope of application and typically consider only one aspect of security (i.e. probability of successful attack confirmation delay recovery complexity) rather than combining those factors into a single framework.

There is no doubt about the separation between (i) qualitative comparative studies of consensus mechanisms and (ii) economically-motivated models of protocol incentives and security; however, there is clearly a deficiency in the form of a unified, risk-modeling quantitative framework that encompasses both technical and economic security attributes. The deficiency represents motivation for the approach outlined in this paper, which seeks to link descriptive analysis and formalized risk evaluation using a composite Risk Score methodology.

3 The Proposed Risk Modeling Framework

The core contribution of this paper is a unified quantitative model. It is designed to evaluate the resilience of consensus mechanisms, especially in malicious environments. It differs from traditional approaches because it primarily conceptualizes security as a continuous risk function driven by economic incentives and temporal response characteristics, rather than qualitative comparisons [1], [6]. This perspective agrees with recent study efforts in cryptoeconomic modeling of blockchain systems that treat security as an emergent property of incentive structures, instead of purely cryptographic guarantees [3], [4].

3.1 The Composite Risk Score Function

To provide quantitative grounds for evaluating consensus protocols, we introduce the Risk Score (RS). This function takes into consideration multiple dimensions of an attack vector into a single, comparable metric:

$$RS = \frac{P + R + lnln\,(1 + T)}{1 + lnln\,(1 + C)} \quad (1)$$

To maximize methodological precision, the operational definitions, mathematical boundaries, and literature-based sourcing for each parameter are standardized in Table 1.

Table 1. Parameter Notation, Operational Definitions, and Sourcing.

Parameter	Name	Boundary	Operational Context	Metric Nature and Sourcing
P	Execution Probability	[0,1]	Likelihood of exploit success given the network's architectural and validator state.	Stochastic metric; theoretically measurable via hash-rate/stake distribution profiles as documented in [1], [6], [14].
R	Recovery Difficulty	[0,1]	Systemic effort and friction required to restore ledger integrity post-breach.	Qualitative index; derived from protocol-level governance rules and slashing mechanics analyzed in [3], [6].
T	Confirmation Delay	[0,1]	Time window required to achieve deterministic finality or state consensus (applied to PoW).	Temporal metric; directly measurable via native block generation times and network finality depth [1], [9].
C	Economic Hurdle	[0,1]	Total financial expenditure necessary to aggregate resources required for the attack.	Financial metric; empirically derived from hardware markets and liquid capital requirements as modeled in [3], [6].

The model assumes that the parameters have orthogonal relationships; therefore, each parameter acts as a separate variable. The decision to take the natural logarithm of both T and C models specific marginal effects. For T, it results in a decreasing marginal effect for increasing times of detection. For C, it creates a less-than-unity power function between capital costs and entry barriers. Both beliefs are consistent with

common assumptions used in the literature to define the incentives found within Proof-of-Stake (PoS) cryptoeconomic models.

To standardize the assignment of values and eliminate arbitrary estimation, Table 2 translates the qualitative boundaries established in blockchain security and cryptoeconomic literature into a structured quantitative scoring rubric.

Table 2. Literature-Grounded Parameter Scoring Rubric.

Score Range	Operational Meaning	P	R	T	C
0.00-0.33	Low Risk / High Resilience	Strictly bounded by mathematical honest-majority or cryptographic stake assumptions [1], [6].	Automated in-protocol recovery mechanics (e.g., immediate slashing or programmatic finality) [6], [15].	Achieved within minimal block depth or near-instantaneous deterministic finality [1], [9].	Attack costs exceed standard feasibility thresholds (e.g., >50% of total network capital) [2], [16].
0.34 - 0.66	Moderate Risk / Exposure	Exploit possible during transient adversarial alignment or pool concentrations [7], [14].	Requires out-of-protocol intervention, such as social consensus coordination or targeted hard forks [3], [4].	Settlement dependent on deep block confirmations or prolonged block intervals [9], [12].	Capital costs are significant but achievable for well-funded coalitions or nation-states [3], [6].
0.67 - 1.00	High Risk / Severe Vulnerability	Exploit highly probable due to severe structural centralization or systemic protocol flaws [1], [14].	Catastrophic governance deadlocks or permanent, irreversible ledger state corruption [4], [6].	Detection or confirmation is delayed indefinitely due to deep stealth reorganizations [1], [14].	Financial barriers are negligible relative to the economic capacity of the adversary [2], [16].

3.2 Normalization and Scale Mapping

Given that the raw Risk Score is heavily dependent on the magnitude of input parameters, some form of normalization would be a necessity to enable meaningful

comparison across heterogeneous consensus mechanisms. We define the normalized Risk Score as:

$$RS_{norm} = \frac{RS - RS_{min}}{RS_{max} - RS_{min}} \quad (2)$$

where RSmin and RSmax represent theoretical bounds chosen to normalize the framework across the defined parameter range. Comparative studies of consensus protocols frequently utilize normalization techniques to enable cross-protocol evaluation with a unified metric [1], [9].

To support interpretability for both technical and non-technical stakeholders, the continuous value $RS_{norm} \in [0,1]$ is mapped onto a Likert-type ordinal scale, as shown in Table 3.

Table 3. Mapping of Normalized Risk Score to Risk Levels.

RS_{norm} Range	Risk Level	Score
0.00-0.20	Very Low Risk	1
0.20-0.40	Low Risk	2
0.40-0.60	Medium Risk	3
0.60-0.80	High Risk	4
0.80-1.00	Very High Risk	5

There is an important point about RS and RSnorm as they represent strict risk metrics, not security or safety scores. The scale is monotonically increasing with risk: a higher RS score directly indicates a more vulnerable system with greater risk exposure. These methods have been used to measure how vulnerable a system is based on a set of parameters or assumptions made by each method. The higher the RS score; the more vulnerable the system will be. The use of a two layer approach provides both quality (qualitative) and quantity (quantitative) in measuring systemic risk profiles. In many ways this is similar to what has been seen in other areas of blockchain security. Blockchain technology allows for translation of technical data into a format that provides insight to decision makers, which is often required in order for effective governance [1], [4]. Additionally, earlier comparative studies demonstrated that abstracting away layers of detail provides the ability to compare multiple protocols. This approach maintains consistency throughout the evaluation process [1], [5].

3.3 Framework Validation Protocol

To verify whether the generated risk scores are mathematically meaningful and stable, the framework implements a sensitivity-based validation strategy. To isolate which parameters exert the highest leverage on the final risk score, we compute the first-order partial derivatives of the continuous RS function with respect to each independent variable:

$$\frac{\partial RS}{\partial P} = \frac{1}{1 + ln\ ln\ (C + 1)}$$

$$\frac{\partial RS}{\partial R} = \frac{1}{(C+1)}$$

$$\frac{\partial RS}{\partial T} = \frac{1}{(T+1)\cdot(C+1))}$$

$$\frac{\partial RS}{\partial C} = -\frac{P+R+ln\,ln\,(T+1)}{(C+1)\cdot[(C+1)]^2[ln\,ln\,(C+1)]} = -\frac{RS}{(C+1)\cdot(C+1))}$$

The mathematical formulation confirms that RS scales linearly with changes in execution probability (P) and recovery difficulty (R). Conversely, the derivative $\frac{\partial RS}{\partial T}$ exhibits a factor of $\frac{1}{T+1}$, formally proving the intended diminishing marginal risk response as confirmation delays grow. Similarly, the negative sign in $\frac{\partial RS}{\partial C}$ mathematically demonstrates the non-linear deterrent effect of the economic hurdle (C). At the baseline values, P and R are the most influential parameters, followed by T, while C acts as a deterrent term.

4 Comparative Analysis and Framework Validation: PoW vs. PoS

Validating the utility of the proposed framework requires applying the Risk Score (RS) function to the two most prominent consensus mechanisms today: Proof-of-Work (PoW) and Proof-of-Stake (PoS). The analysis focuses on a standard majority attack scenario and evaluates how each protocol performs across the four proposed dimensions.

4.1 Parameter Estimation

None of the values for P, R, T or C in this assessment were empirically measured. Consequently, the resulting scores are illustrative and transparently subjective, meaning that the final outputs depend heavily on how these baseline metrics are designated. The values used to make those assessments were therefore instead based on heuristic (rule-of-thumb) estimations made using qualitative findings and comparative discussions from previous literature regarding blockchain systems as well as general characteristics that have been observed among major blockchain networks such as Bitcoin and Ethereum.

- **Proof-of-Work (PoW):** PoW is characterized by a high economic hurdle (C = 0.90) relative to hardware and electricity costs [3][9][16]; moderate recovery difficulty (R = 0.50), since PoW depends primarily on the "heaviest chain" rule [14]; and low execution probability (P = 0.35) reflecting hash-rate concentration risks [7].
- **Proof-of-Stake (PoS):** PoS has an entirely different profile in terms of a very low recovery difficulty (R = 0.15) due to cryptoeconomic penalty structures [3][6] and lower execution probability (P = 0.25) in mature networks [2]. Additionally, PoS has a much lower confirmation delay (T = 0.30); due to protocol-level security assumptions and validator reward dynamics [6]. Finally, the economic hurdle (C = 0.75) was modeled differently because it pertains to liquid capital rather than physical infrastructure [4],[15],[16].

4.2 Calculation of Raw Risk Scores (RS)

Using Equation (1), we calculate the raw risk values for both mechanisms:

$$RS_{PoW} = \frac{0.35 + 0.50 + lnln\,(1 + 0.70)}{1 + lnln\,(1 + 0.90)} = \frac{0.85 + 0.5306}{1 + 0.6418} \approx 0.8409$$

$$RS_{PoS} = \frac{0.25 + 0.15 + lnln\,(1 + 0.30)}{1 + lnln\,(1 + 0.75)} = \frac{0.40 + 0.2623}{1 + 0.5596} \approx 0.4247$$

4.3 Determination of Theoretical Boundaries

In order to place these results within a global security context, we derive the absolute theoretical boundaries of the proposed RS function for the parameter range $[0,1]$:

Theoretical minimum (RSmin) is achieved at P=R=T=0, C=1 and represents a theoreticaly perfectly safe system:

$$RS_{min} = \frac{0 + 0 + lnln\,(1)}{1 + lnln\,(1 + 1)} = 0$$

Theoretical maximum (RSmax) is achieved at P=R=T=1, C=0 and represents the theoretically most vulnerable system:

$$RS_{max} = \frac{1 + 1 + lnln\,(1 + 1)}{1 + lnln\,(1 + 0)} = \frac{2.6931}{1} = 2.6931$$

4.4 Normalization and Risk Level Assessment

The final normalized risk (RSnorm) is obtained by scaling the raw results against the calculated theoretical extremes:

$$RS_{norm(PoW)} = \frac{0.8409 - 0}{2.6931 - 0} \approx 0.31$$

$$RS_{norm(PoS)} = \frac{0.4247 - 0}{2.6931 - 0} \approx 0.16$$

Following the completion of the normalization process, the scores from each mechanism are then combined into one composite summary comparing them. The

RSnorm (Risk Score Normalized) ratings for both mechanisms are reasonably low. This indicates that, while either protocol may have an issue with specific aspects of its security posture, neither has enough issues to put the entire security posture at unacceptable risk.

While the use of this model will primarily provide the RSnorm rating, the most useful component of this model is the ability to gain additional insights in detail. While making a decision based solely upon the RSnorm rating may be difficult, the model can break down the rating into various dimensions. This allows stakeholders to compare technical resiliency or economic barriers based on their risk tolerance.

Table 4. Comparative Risk Score Summary.

Mechanism	P	R	T	C	RS_{norm}	Risk Level
PoW	0.350	0.50	0.70	0.90	0.31	Low (2)
PoS	0.255	0.15	0.30	0.75	0.16	Very Low (1)

Mapping these technical attributes to a unified scale allows the framework to move beyond qualitative descriptions and provides a valuable tool for stakeholders and network architects to assess which mechanism best aligns with their specific risk tolerance requirements.

4.5 Results and Visualization

To visualize the multi-dimensional security profile of each mechanism, we employ a **spider (radar) diagram**. This visualization shows the trade-offs inherent in each protocol's design.

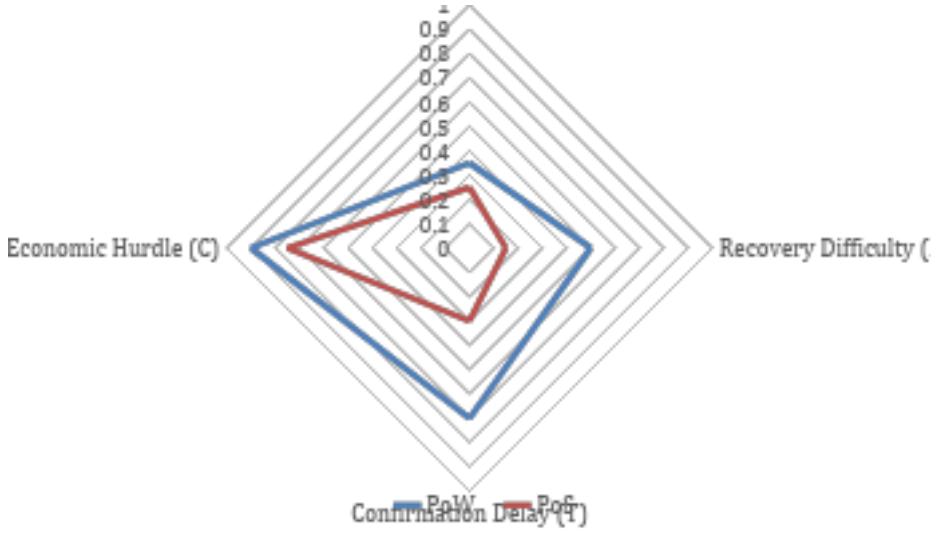


Fig. 1. Comparative analysis of PoW and PoS Security Dimensions. Values are normalized on a scale of 0 to 1, where higher values represent increased magnitude within each specific dimension.

Discussion. The radar diagram illustrates the structural trade-offs inherent in each protocol's design architecture. While both mechanisms achieve comparable baseline security profiles, they distribute their resilience differently across the four conceptual dimensions. PoW relies heavily on a maximized economic hurdle (C) backed by physical infrastructure and energy expenditure, but exhibits higher vulnerability in confirmational delay (T). Conversely, PoS minimizes baseline recovery difficulty (R) through programmatic cryptoeconomic design, shifting its primary risk exposure toward potential adjustments in execution probability (P) under scenarios of high stake concentration.

Both PoW and PoS fall within the "Very Low" or "Low" risk categories according to the RSnorm classification. Within the assumptions of the proposed model, this indicates relatively low risk profiles during normal operation. This baseline resilience does not imply absolute immunity; rather, it indicates that even within a low-risk profile, each architecture exhibits localized, structural vulnerabilities when subjected to targeted attack vectors. Potential Vulnerabilities:

- **PoW:** The primary vulnerability exists in confirmation delay, because if an attacker has enough computational capability, he/she can produce a competing longest chain at the same rate as the legitimate network. This will make it very difficult to determine whether a malicious reorganization has occurred until after the divergence occurs.
- **PoS:** While the PoS vulnerability is most exposed to changes in Execution Probability under conditions of high stake concentrations, this is somewhat offset by low Recovery Difficulty.

One of the key benefits of the proposed method lies outside of the simple quantitative comparisons. Different attack vectors, such as eclipse attacks or sybil-based governance exploits could result in dramatically different inputs to the parameters P, R, T, and C, which would ultimately affect the variation in risk score values. Additionally, applying this methodology to other types of consensus mechanisms would likely demonstrate even greater variations in risk scores than those found here.

5 Conclusion

This paper developed an integrated risk modeling method using a Risk Score (RS) model for evaluating the comparative risks of both the Proof-of-Work (PoW) and Proof-of-Stake (PoS) consensus methods. Departing from legacy descriptive approaches, this paper delivers a quantifiable methodology rooted in four distinct dimensions of risk: the financial barriers to entry, the probability of exploit execution, the timeframe required for threat identification, and the friction associated with post-attack systemic recovery.

Our study's primary contribution is a formalized analytical tool that allows users to compare the distinct risk profiles of different consensus mechanisms. The parameters adapt to the specific context in which they are applied, allowing the framework to remain dynamic across the range of potential uses for each blockchain network.

Although the utility of our proposed model is grounded in its ability to allow comparisons among competing consensus mechanisms based on a common logic, it isolates specific architectural vulnerabilities, such as PoW's exposure to confirmation delay during stealth reorganizations and PoS's sensitivity to stake concentration. Future studies will explore the use of our proposed model to analyze additional consensus mechanisms that utilize hybrid Byzantine Fault Tolerance (BFT)-based models and other emerging consensus mechanisms in order to demonstrate the applicability of our proposed model to a wider variety of consensus mechanisms.

Disclosure of Interests. The authors hereby declare that they have no competing interests.

References

1. Álvarez, I.A., Gramlich, V., Sedlmeir, J.: Unsealing the secrets of blockchain consensus: a systematic comparison of the formal security of proof-of-work and proof-of-stake. arXiv preprint arXiv:2401.14527 (2024). DOI: 10.48550/arXiv.2401.14527.
2. Andrews, K., Ngo, L.B., Amiruzzaman, M.: A detailed comparative analysis of blockchain consensus mechanisms. arXiv preprint arXiv:2511.15730 (2025). https://arxiv.org/abs/2511.15730
3. Arafat, S.M.: A study of blockchain consensus protocols. Cybersecurity **9**, 120 (2026). DOI: 10.1186/s42400-025-00417-x.
4. Bruno, B., Murano, A., Vespri, V.: Incentives for sustainable governance in blockchain-based organizations. Sustainability **17**(21), 9728 (2025). DOI: 10.3390/su17219728.

5. Daian, P., Goldfeder, S., Kell, T., Li, Y., Zhao, X., Bentov, I., Breidenbach, L., Juels, A.: Flash Boys 2.0: Frontrunning, transaction reordering, and consensus instability in decentralized exchanges. arXiv preprint arXiv:1904.05234 (2019). https://arxiv.org/abs/1904.05234
6. Li, S.-N., Xu, J., Tasca, P., Tessone, C.J.: Proof-of-stake cryptoeconomics design: a general framework of modeling and evaluation. SSRN Electronic Journal (2024). DOI: 10.2139/ssrn.5072963.
7. Lin, Q., Li, C., Zhao, X., Chen, X.: Measuring decentralization in Bitcoin and Ethereum using multiple metrics and granularities. arXiv preprint arXiv:2101.10699 (2021). https://arxiv.org/abs/2101.10699
8. Motepalli, S., Jacobsen, H.-A.: Decentralization in PoS blockchain consensus: quantification and advancement. arXiv preprint arXiv:2504.14351 (2025). https://arxiv.org/abs/2504.14351
9. Shaikh, M., Wiil, U.K., Ebrahimi, A., Memon, Y.: An overview and comparison of blockchain consensus mechanisms. Int. J. Networked Distrib. Comput. 14, art. 4 (2026). DOI: 10.1007/s44227-025-00087-8.
10. Nakamoto, S.: Bitcoin: A Peer-to-Peer Electronic Cash System. White paper (2008). Available at: https://bitcoin.org/bitcoin.pdf. Accessed 23 April 2026.
11. Buterin, V.: Ethereum: A Next-Generation Smart Contract and Decentralized Application Platform. White paper (2014). Available at: https://ethereum.org/en/whitepaper/. Accessed 23 April 2026.
12. Kiffer, L., Neu, J., Sridhar, S., Zohar, A., Tse, D.: Nakamoto Consensus under Bounded Processing Capacity. In: CCS 2024: Proceedings of the 2024 ACM SIGSAC Conference on Computer and Communications Security, pp. 363–377. Association for Computing Machinery (2024). DOI: 10.1145/3658644.3670347.
13. Pass, R., Seeman, L., Shelat, A.: Analysis of the blockchain protocol in asynchronous networks. In: Coron, J.-S., Nielsen, J.B. (eds.) Advances in Cryptology – EUROCRYPT 2017 – 36th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Paris, France, April 30–May 4 2017, Proceedings, Part II. LNCS, vol. 10211, pp. 643–673. Springer, Cham (2017). DOI: 10.1007/978-3-319-56614-6_22.
14. Garay, J., Kiayias, A., Leonardos, N.: The Bitcoin Backbone Protocol: Analysis and Applications. In: Oswald, E., Fischlin, M. (eds.) Advances in Cryptology – EUROCRYPT 2015 – 34th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Sofia, Bulgaria, April 26–30, 2015, Proceedings, Part II. LNCS, vol. 9057, pp. 281–310. Springer, Cham (2015). DOI: 10.1007/978-3-662-46803-6_10.
15. Saleh, F.: Blockchain without Waste: Proof-of-Stake. The Review of Financial Studies 34(3), 1156–1190 (2021). DOI: 10.1093/rfs/hhaa075.
16. Gans, J.: Proof of Work Versus Proof of Stake. In: The Economics of Blockchain Consensus: Exploring the Key Tradeoffs in Blockchain Design, pp. 69–83. Palgrave Macmillan, Cham (2023). DOI: 10.1007/978-3-031-33083-4.

Original Research Article *https://doi.org/10.5281/zenodo.23474581*

Mapping AI agents in Enterprise Cybersecurity: A Taxonomy of Architectures, Autonomy and Adoption Patterns

Petar Ilić[1][0009-0000-1668-1291], Uroš Šošević[1][0000-0003-0322-8097] and Dušan Mitrović[1][0000-0001-5299-9707]

[1] University of Belgrade – Faculty of Organizational Sciences, 11010 Belgrade, Serbia
{petar.ilic, uros.sosevic, dusan.mitrovic}@fon.bg.ac.rs

Abstract. The growth of AI agents into Enterprise Cybersecurity Environments, has occurred much faster than the classification or governance of those agents. This paper addresses this gap in two ways. First, it introduces a taxonomy for classifying AI agent Applications, using four key attributes of those applications; Functional Role, Level of Autonomy, Architectural Design, and Security Lifecycle Position. Second, it uses that taxonomy to analyze seven different platforms (five commercial, two research grade), which allowed it to identify trends or patterns that are often obscured when reviewing each platform individually. The results indicate that this is an environment that is currently dominated by institutional conservatism: most commercial deployments occur with relatively low levels of autonomous functionality; they favor architectures that allow for human oversight; and are nearly exclusively defensive in their orientation. Autonomous Multi-Agent configurations are being developed largely as part of academic research settings, indicating a significant disconnect between prototypes developed within Universities/Research Institutions and how enterprises deploy these types of systems.

Keywords: AI agents; Large Language Models, enterprise cybersecurity; application taxonomy; agent autonomy; multi-agent systems; cybersecurity operations; organizational implications.

1 Introduction

Artificial intelligence (AI) has changed the way we use Information Technology. Some even refer to it as a new industrial revolution due to its transformative impact on technology, economy, and society. With capabilities extending beyond simply automating processes, AI has opened new avenues of opportunity in its application [1]. One of the greatest changes in this area is the development of AI agents which can perceive and understand their surroundings, formulate objectives, and perform multiple steps of action with some degree of independence [2]. Utilizing advancements in LLMs and RL, these types of agents are now going from research laboratory testing to being implemented in operational environments such as Enterprise Cybersecurity [3].

Cybersecurity represents one of the most pressing and resource demanding challenges facing modern organizations [4]. Number, speed and complexity of cyber-attacks continues to grow faster than the ability of human security professionals and traditional automated security tools to respond [4]. So, SOCs (Security Operations Center) continue to be burdened with chronic alert fatigue, personnel shortfalls, and increasing exposure due to widespread adoption of cloud computing, remote workers, and growing numbers of Internet-connected devices [5]. In this context, the integration of AI agents into enterprise cybersecurity workflows is no longer a speculative prospect. The integration of AI agents is an active and rapidly evolving reality [6].

However, despite the rapid growth in the adoption of AI agents within the Enterprise Cybersecurity domain there remains a lack of a common and organized framework to describe the various ways in which AI agents are currently being used to support different cybersecurity functions, the architectures that are being followed in supporting those functions, and the impacts that organizations may experience as a result of adopting AI agents [7]. Commercial vendors, independent research teams, and open-source communities are developing solutions at pace, yet these efforts remain fragmented, with little systematic comparison or classification available in the academic literature [8].

To fill that void, the focus of this research will be on developing a Taxonomy of AI agent applications within enterprise cybersecurity. The Taxonomy will provide a way to classify existing and emerging AI agent-based security applications into four distinct categories (Functional Role/functionality; Level of Autonomy; Architectural Design; and Position within the Security Lifecycle). Following the development of the Taxonomy, this paper will review several commercially available and research-oriented security solutions based upon their alignment with the previously defined taxonomic categories. Additionally, the review discusses key deployment considerations including governance requirements, human oversight models, and the risk tolerance dynamics that shape current enterprise adoption patterns. They are typically overlooked in technologically focused studies.

The remaining sections of this report follow this structure. Section Two provides the theoretical basis for this research through definitions related to AI agents and their relationship to cybersecurity environments. Section Three discusses the Taxonomy used to develop classifications for AI-based security applications. Section Four utilizes the developed Taxonomy to evaluate a variety of real-world AI-based security application examples. Section Five addresses the organizational and management-related aspects of implementing AI agents. Finally, Sections Six and Seven discuss key challenges facing researchers who seek to apply AI agents to enterprise cybersecurity problems and possible directions for future research, respectively.

2 Theoretical Foundation: AI agents as a concept

Artificial Intelligence (AI) represents a broad field within computer science focused on the development of systems capable of performing tasks that typically require human intelligence, such as learning, reasoning, problem-solving, and decision-making. AI is conceptualized as the study of rational agents that perceive their environment and act

to maximize goal achievement [9]. However, defining AI remains inherently complex and contested. Recent research highlights that AI does not have a single universally accepted definition, there are many definitions depending upon the technology being developed, philosophy supporting that technology, and societal implications of that technology [10].

One common way to define AI is that AI is any system(s) able to interpret external information/data, learn from that information/data and apply that learned knowledge to achieve a particular goal through adaptable behaviors [11]. Another author defines AI as an entity (and/or entities) that perceives its environment, processes environmental input(s), adapts those inputs into a set of flexible actions and achieves long-term goal attainment [10]. The latter author's viewpoint is very similar to the agent-centered paradigm for viewing AI.

Based on the agent-centered paradigm, we now see AI agents evolving rapidly with the advent of large language models and Reinforcement Learning methods. An AI agent is defined as a system able to perceive its environment through a variety of means, maintain an internal state, formulate objectives and perform sequential actions over time with some degree of independence [2]. One of the distinguishing features of current AI agents compared to prior automated systems was the ability to evaluate the context of each situation and select among a potentially unlimited number of available options to adapt their actions accordingly. These characteristics: perception, memory, planning and action, create the functional architectures used in all of the Agent-Based Systems referenced in this article.

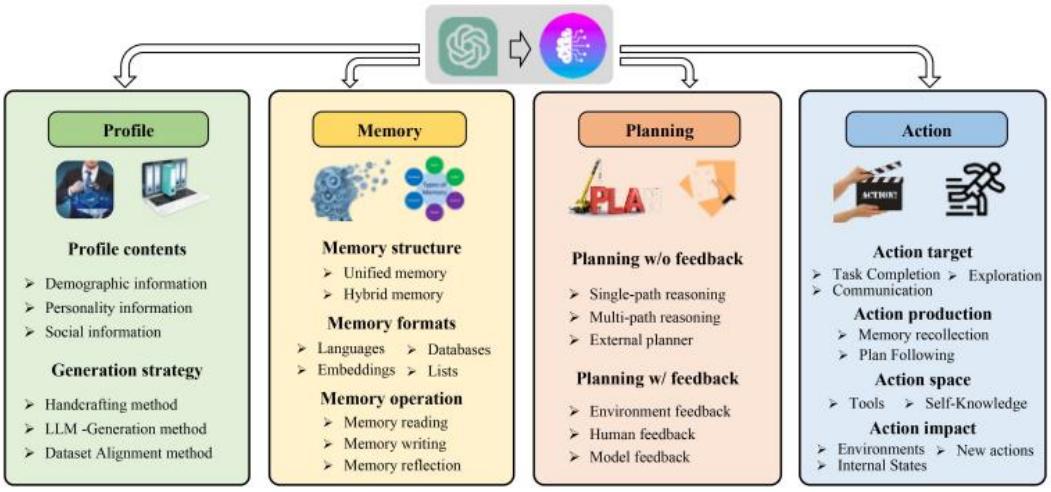


Fig. 1. A unified framework for the architecture design of agents [2]

AI agent's importance shows most clearly when we talk about cybersecurity. Cybersecurity includes all the practices, technologies and procedures needed for protecting information systems, computer networks, and data from an unapproved or malicious attempt to alter, steal, destroy, expose, disrupt or otherwise disable their normal function [12]. Traditional security methods are often insufficiently capable of addressing the

increasing number of sophisticated and frequent cyberattacks. Modern attackers use automated tools, multiple forms of polymorphic malware, and advanced persistent threats (APTs) making it necessary for defenders to develop new and equally sophisticated countermeasures. It is this large gap between the scope of today's threat environment and the current capability of traditional defense strategies which makes agentic AI systems appear so relevant and timely. Agentic AI systems extend beyond reactive models by combining cognitive architectures with proactive, goal directed behavior, qualities that align directly with the demands of enterprise security environments such as Security Operations Centers, where speed, scale, and continuous adaptation are not optional but operationally essential [8].

3 Taxonomy of Applications in Enterprise Cybersecurity

3.1 Introduction and Taxonomy Explanation

The body of knowledge surrounding use of AI-based agents in enterprise cybersecurity is growing rapidly. However, this body of knowledge lacks both a common language/vocabulary and an overarching framework to organize and integrate the knowledge. Recent summaries of LLM-based agents have described the rapid development of this area of study, with most indicating that there will be some form of consensus/consolidation into a more unified body of work. Broader reviews of AI indicate that AI has had significant impacts upon organizations and society as a whole across various industries. Additionally, fragmentation is evident in cybersecurity due to the lack of standardization across vendors offering similar capabilities. It is further reinforced by research prototypes being evaluated only on specific tasks, limiting comparisons between systems. As such, enterprise purchasers and security teams face a challenging market where seemingly different products may solve similar problems, while similarly named products can differ greatly in scope and level of autonomy. Therefore, a structured taxonomy is needed for purposes beyond simply providing a superficial way to categorize products. Taxonomy can provide a means to classify products according to stable dimensions which are relevant when considering product deployment, governance, and oversight.

This section proposes a taxonomy of AI agent applications in enterprise cybersecurity using the iterative method of Nickerson, Varshney, and Muntermann (2013), which combines conceptual and empirical refinement. This methodology states that a well-defined taxonomy must meet three foundational criteria. First, it must be comprehensive (exhaustive); the taxonomy must include all possible classifications within the defined scope so that no observed solution requires creation of new categories. Second, it must be mutually exclusive; a system receives only one clear designation per dimension. Lastly, it must be meaningful; the distinctions made by the taxonomy reflect real differences that are recognized as important by researchers and practitioners; it cannot simply represent differences in terms used or marketing spin [13].

In this section, the taxonomy is applied at the platform level. Enterprise decision makers generally purchase and govern integrated platforms as opposed to individual agent modules.

A conceptual note needs to be made prior to describing the dimensions. It is apparent that satisfying all three criteria simultaneously presents greater difficulty than appears at first glance. Modern enterprise security platforms are becoming increasingly multi-functional; they combine detection, response, intelligence and governance capabilities into a single product offering. Early versions of this taxonomy attempted to place each platform into a single Functional Role category. These early attempts immediately generated classification conflict that is consistent with observations from the broader literature stating that enterprise AI security solutions are rapidly converging and obscuring historical functional boundaries [6, 7]. Compelling a single functional identifier onto a platform like Microsoft Security Copilot or Darktrace, which are designed to support multiple security functions, resulted in assignments that were either inaccurate or meaningless.

This issue was resolved through a structural revision.

Rather than treating Functional Role as a single flat dimension that requires one assignment per platform, it was revised into two hierarchical sub-dimensions: Operational Orientation and Primary Security Function. The Functional Role area was treated more cautiously: Operational Orientation is single assignment, whereas Primary Security Function is assigned as a capability coverage profile within the Defensive orientation. That choice preserves the rigid logic of a taxonomy while still documenting that many modern enterprise security platforms support more than one defensive function. Simply stated, multi-functionality is documented but it does not allow for breaking the rules of a taxonomy's single assignment requirement.

The final three dimensions: Level of Autonomy, Architectural Design, and Security Lifecycle Position remain single-assignment classification. Each of these attributes can be identified clearly and objectively at the platform level.

3.2 Taxonomy Dimensions

Dimension One: Functional Role.

The Functional Role Dimension provides an explanation of the type of security activity that the AI agent platform was created to accomplish. As outlined in Section 3.1, there are two additional sub-Dimensions of this Dimension; Operational Orientation and Primary Security Function.

Operational Orientation. It represents the broadest of the two Sub-Dimensions and uses single-assignment classification. Platforms are classified into three categories according to their relationship to the security mission. These categories are mutually exclusive by definition and not by convention. No platform will operate in both a defensive and an offensive manner within the same deployment context.

Platforms that were developed to protect the organization, to monitor the organizations environment, to detect potential threats, and to respond to actual security breaches fall under the Defensive category. Due to the breadth and commercial applicability of this category it is currently the largest category in terms of commercially available platforms within the Enterprise environment. [5, 6]

Those platforms that were developed to emulate the actions of adversaries through simulations to assist with the identification of vulnerabilities prior to when they may be

exploited by a real attacker fall under the Offensive category. Examples include autonomous penetration testing agents and breach-and-attack simulation systems. Although the ultimate goal of these tools is to provide a defensive function to the organization, their operational mode is inherently offensive. Rather than protecting systems, they seek out and test system vulnerabilities. Because of this operational mode they are categorized separately. [14, 15]

Those platforms that have been developed to focus on ensuring compliance with organizational security policies, regulatory requirements, and auditing obligations fall under the Governance, Risk, and Compliance (GRC) category. In contrast to directly providing evidence regarding whether specific organizational activities complied with established policies and regulations, GRC platforms focus on determining whether or not those organizations are compliant. As regulatory pressure on enterprise deployments of AI continues to increase, so too does the relevance of GRC-capable AI Systems. [7, 16].

Primary Security Functions. Within the Defensive Category, there are four primary security functions. The first is Monitoring and Detection. Platforms that fall under this category continuously analyze their environments using log data, network traffic, endpoint telemetry, and user behavior to identify anomalies that indicate potential malicious activity and then generate actionable alerting messages. This application area is one of the most advanced and commonly used types of AI agent platforms in use in enterprise Security Operations Centers today [5].

The second primary security function is Incident Response. Those platforms that were designed to act upon confirmed or suspected threats fall under this category. Once a potential threat has been identified, incident response platforms automatically isolate the affected systems, revoke any compromised credentials, execute containment measures against the potential threats, and support any necessary post-response recovery activities. Since incident response requires significant interaction with all elements of an organization's technical infrastructure, it carries the greatest amount of risk should an organization suffer an error resulting from an incorrect autonomous decision made during the execution of incident response functions [17].

The third primary security function is Threat Intelligence. Those platforms that collect, correlate, synthesize, and present information related to adversary tactics, techniques, and procedures via both internal logs and external feeds fall under this category. Unlike other categories where the primary output is either direct protective action or some form of reactive response, the primary output of intelligence platforms is contextual knowledge that supports human and/or automated decision making processes [6].

The fourth primary security function is Vulnerability and Exposure Management. Those platforms that utilize asset discovery and vulnerability assessment scans to identify weaknesses in an organization's infrastructure and applications fall under this category. After discovering potential vulnerabilities and exposures, these platforms prioritize the findings based on exploitability and associated business risk, track remediation status over time, and operate primarily in a pre-incident posture reducing an organization's overall risk exposure before such events occur rather than reacting to them once they do occur [7].

Table 1. Structure of the Functional Role Dimension

Sub-dimension	Type	Categories
Operational Orientation	Single assignment	Defensive / Offensive and GRC
Primary Security Function (Defensive Only)	Capability Coverage	Monitoring and Detection / Incident Response / Threat Intelligence / Vulnerability and Exposure Management

Dimension Two: Level of Autonomy.

The Level of Autonomy refers to the extent that a particular AI agent platform makes decisions about how to perform its assigned tasks on behalf of the organization as opposed to making those decisions based upon direct instructions from humans. As stated above, for the purpose of defining this taxonomic level, the Level of Autonomy refers to whether an AI platform has the ability to recognize situations; process information; make decisions relative to the processing information; and take appropriate actions without the need for human intervention in the moment of action. Feng, McDonald, and Zhang (2025), in their research paper on the various levels of AI agent Autonomy, state that autonomy should be viewed as a specific design choice rather than being an inherent result of the capabilities of the AI agent [18].

This taxonomy uses a four-level scale. The rationale for four levels rather than three or five is practical: three levels collapse too many meaningfully different operating modes into a single category, while five levels introduce distinctions that current commercial platforms do not yet support cleanly. It aligns broadly with the graduated autonomy models described in recent agentic AI literature [6, 7].

Level 1, Copilot and Assistant. The platform provides responses to questions posed by humans and produces output in the form of completely advisory recommendations. The platform can generate summaries of alert messages; create natural language descriptions of security events; recommend possible courses of action; and prepare drafts of reports. However, the platform cannot autonomously begin or execute any type of action in the environment. In other words, all decisions and all keystrokes that cause some kind of environmental change will always be performed by a human.

Level 2, Semi-Autonomous. The platform identifies necessary actions and prepares them for execution prior to placing them into a queue. Prior to these queued actions taking place, they require approval by a human. The platform performs analysis and preparation to enable a faster response, however, the human retains control over what ultimately occurs in systems and data using deliberate approval for each resulting consequence.

Level 3 — Supervised Autonomous. The platform acts within a defined policy scope without seeking approval for each individual action. It executes, then notifies. Human oversight does not disappear at this level. It shifts from pre-action approval of individual steps to post-action review and ongoing policy governance. This level is where most

of the organizational governance tension lives, because it requires organizations to trust that the system's policy scope is correctly defined before acting [17].

Level 4 — Fully Autonomous. The platform analyzes, decides, and acts in real-time without needing approval from humans. The platform only escalates to human operators when circumstances occur that clearly exceed the predetermined operating boundaries of the system. Therefore, involvement by humans at this level is generally limited to establishing initial policy configurations; monitoring performance; and handling exceptions [18].

Table 2. Level of Autonomy

Level	Label	Human role	Decision Authority
1	Copilot / Assistant	Active direction of every query and action	Fully human
2	Semi - Autonomous	Approval before each consequential action	Shared
3	Supervised Autonomous	Post-action review and policy governance	Primarily agent
4	Fully Autonomous	Policy definition and exception handling only	Fully Agent

Dimension Three: Architectural Design.

Architecture defines how a problem can be divided into smaller pieces (tasks) and how information will flow between those pieces; architecture also defines how error propagates through a problem space, how a system will scale when additional requests are made upon it, how a system will recover from unplanned events, and finally how that system will integrate into the technology and people that currently exist as part of an organization's overall Security Operations Center (SOC). Four Architectural Design types are categorized within this taxonomy. These were created by using concepts found in multi-agent systems research and recent studies that explored the use of cyber-agents in the field of cybersecurity [19].

Single-Agent Architecture. A single agent is assigned the responsibility for all three phases of its assigned task: Perception (data acquisition), Reasoning (analysis), and Action (response). Since there is no need to coordinate activities, no communications occur between agents, and no opportunity exists for one agent to undermine another agent's efforts.

Multi-Agent Collaborative Architecture. This type replaces the single-agent model with a group of peers who operate at the same hierarchical level. Agents in the collaborative architecture are functionally specialized and each has their own role or strategy. However, they are not hierarchically superior to other agents [19].

Multi-Agent Hierarchical Architecture. Introduces an orchestrator agent above the specialized agents. The orchestrator receives a high-level objective from the human

operator, decomposes it into a set of sub-tasks, delegates each task to a specialized subordinate agent, monitors progress, and synthesizes results into a coherent output.
Human-Agent Hybrid Architecture. It is distinctive in that it views the Human Analyst as a participatory element in the overall process rather than simply as an observer.

Dimension Four: Security Lifecycle Position.
This taxonomy anchors the Lifecycle Position dimension to the NIST Cybersecurity Framework 2.0 [20] and its AI-specific profile [16].

Table 3. Security Lifecycle Position

Lifecycle Position	Corresponding NIST CSF 2.0 Functions	Defining Operational Characteristic
Pre-Incident	Govern, Identify, Protect	Proactive risk reduction
During Incident	Detect, Respond	Realtime, time critical response
Post-Incident	Recover	Analysis, restoration and learning
Cross-Phase	All functions	Persistent and adaptive operation

4 Review of Solutions

This section applies the taxonomy developed in Section 3 to seven representative AI-based cybersecurity solutions, encompassing five commercial platforms and two research prototypes. Each solution is classified across all four taxonomic dimensions using publicly available documentation.
The seven platforms reviewed in this section were selected according to three criteria: market presence, documentation availability, and architectural diversity. Five commercial platforms: Microsoft Security Copilot, Darktrace, CrowdStrike Charlotte AI, SentinelOne Purple AI, and Google Security Operations, were included based on their widespread enterprise adoption and the availability of sufficient technical documentation for classification purposes. Two research grade systems, PentestGPT and VulnBot, were selected to represent the multi-agent and offensive paradigms that are currently documented primarily in academic literature. GRC-oriented platforms were considered but excluded because no system of sufficient technical maturity or documentation depth was identified at the time of writing. This selection was intended to be representative rather than exhaustive and was designed to illustrate the primary architectural and functional configurations observed across the field.

Table 4. Classification matrix of selected AI-based cybersecurity solutions

Solution	Type	Operational Orientation	Primary Security Function	Level of Autonomy	Architectural Design	Security Lifecycle Position
Microsoft Security Copilot	C	Defensive	Threat intelligence; Incident response	Lvl 1, Copilot	Human-Agent Hybrid	Cross-Phase
Darktrace	C	Defensive	Monitoring and detection	Lvl 3, Supervised Autonomous	Single-Agent	During Incident
CrowdStrike Charlotte AI	C	Defensive	Monitoring and detection; Incident response	Lvl 2, Semi-Autonomous	Human-Agent Hybrid	Cross-Phase
SentinelOne Purple AI	C	Defensive	Monitoring and detection	Lvl 2, Semi-Autonomous	Human-Agent Hybrid	During Incident
PentestGPT	R	Offensive	Penetration testing	Lvl 2, Semi-Autonomous	Multi-Agent Hierarchical	Pre-Incident
VulnBot	R	Offensive	Penetration testing	Lvl 3, Supervised Autonomous	Multi-Agent Collaborative	Pre-Incident
Google Security Operations	C	Defensive	Monitoring and detection; Threat intelligence	Lvl 2, Semi-Autonomous	Single-Agent	Cross-Phase

The classification matrix (Table 4) is based on vendor documentation and prior research on AI agents in cybersecurity [5, 6, 8], and data collected from the following sources: Microsoft Security Copilot [21], Darktrace [5, 22], CrowdStrike Charlotte AI [23], SentinelOne Purple AI [24], PentestGPT [14], VulnBot [15], and Google Security Operations [25]. Due to space constraints in the table, "C" is used to denote Commercial solutions, while "R" denotes Research solutions.

The classification matrix revealed several structural patterns across the reviewed solutions. Commercial platforms predominantly adopted single-agent or human-agent hybrid architectures, while multi-agent designs remained confined to research prototypes such as PentestGPT and VulnBot [14, 15], indicating that the multi-agent paradigm had not yet entered mainstream enterprise deployment [19]. All commercial solutions operated in the Defensive orientation, whereas both offensive systems were research-grade penetration testing tools, with no GRC-oriented platform of sufficient maturity identified [7]. In terms of behavioral mode, during incident platforms (Darktrace, SentinelOne Purple AI) behaved reactively, cross-phase platforms (Microsoft Security Copilot, CrowdStrike Charlotte AI, Google Security Operations) combined reactive and proactive capabilities, and offensive prototypes were exclusively proactive [6, 8].

Finally, Level 2 (semi-autonomous) was the dominant autonomy configuration among commercial products, reflecting enterprise risk tolerance and human-in-the-loop governance requirements [17, 18], while both research prototypes reached Level 3, highlighting a notable autonomy gap between academic and commercial deployments.

5 Challenges and Future Research

The single-assignment logic that defines the taxonomy has come under increasing pressure from the growing multifunctionality of commercial platforms. The single-dimensional sub-dimensions we proposed here provide an intermediate step toward providing a taxonomic structure that can capture the effects of platform convergence. Several areas of the taxonomy remain empirically thin: no GRC-related commercial platforms were found, no solutions achieved level four autonomy, and the used list of platforms contains multi-agents only as research prototypes. Thus, while these gaps represent the existing boundaries of this research, they also represent limitations in how fully validated the taxonomy is.

The gap between the autonomy of commercial deployments and research prototypes represents perhaps the most important finding reported here. This gap does not reflect a lack of technology but rather an issue of governance. Liability exposure and an organization's willingness to take institutional risks determine the degree of autonomy an organization will deploy. A primary area for further research involves identifying the governance context in which higher degrees of autonomy become acceptable in enterprise environments. Additionally, since each platform classification relies upon vendor documentation, which by its very nature is promotional, independent empirical validation using controlled testbed experiments is required as the next step in validating these classifications.

Finally, with the emergence of evolving regulatory structures, primarily the EU AI Act (2024), there are requirements imposed upon autonomous security systems whose operational implications have not been systematically assessed [26].

6 Conclusion

This article presents two important additions to the body of literature on the use of artificial intelligence (AI) agents in enterprise cybersecurity. First, this article offers a four-dimensional taxonomy for classifying AI agents in enterprise cybersecurity. Second, using the proposed taxonomy, we apply it to seven exemplary systems to identify relationships among these systems that would otherwise be unrecognizable if each system were viewed independently. The most significant finding is one of institutional conservatism; as such, commercial products generally maintain human control, tend to act defensively, and are clustered below an autonomous capability level which has been demonstrated by prototype systems to be technologically feasible. These differences constitute a major frontiers area of the field and point directly to areas where both technology and organization must come together with regulation to allow for development and deployment of the next generation of enterprise AI based security technologies.

Several limitations of this study should be acknowledged. First, the platform sample of seven systems is small and may not fully represent the breadth of AI agent deployments currently active in enterprise cybersecurity. Second, all taxonomic classifications rely on publicly available vendor documentation, which is inherently promotional in nature and may not accurately reflect actual system behavior or the true scope of autonomous capabilities in operational conditions. These limitations do not invalidate the findings, but they do establish the current work as a foundational step that requires empirical follow-up.

Acknowledgments. This research was supported by the University of Belgrade – Faculty of Organizational Sciences and, in part, by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia through institutional funding (grant number: 200151).
The authors acknowledge the use of AI-assisted language tools solely for grammatical proofreading, language refinement, and stylistic improvements of the manuscript. These tools were not used for the generation of research content, data analysis, interpretation of results, or formulation of scientific conclusions. All scientific content, findings, and conclusions remain the sole responsibility of the authors.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Makridakis, S.: The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms. Futures 90, 46–60 (2017)
2. Wang, L., Ma, C., Feng, X., et al.: A survey on large language model based autonomous agents. Frontiers of Computer Science 18, 186345 (2024)
3. Xi, Z., Chen, W., Guo, X., et al.: The rise and potential of large language model based agents: A survey. arXiv preprint arXiv:2309.07864 (2023)
4. ISC2: ISC2 Cybersecurity Workforce Study 2024. ISC2, Alexandria, VA (2024). https://www.isc2.org/Insights/2024/10/ISC2-2024-Cybersecurity-Workforce-Study, last accessed 2026/04/02
5. Srinivas, S., Kirk, B., Zendejas, J., Espino, M., Boskovich, M., Bari, A., Dajani, K., Alzahrani, N.: AI-augmented SOC: A survey of LLMs and agents for security automation. Journal of Cybersecurity and Privacy 5(4), Article 95 (2025)
6. Lazer, S.J., Aryal, K., Gupta, M., Bertino, E.: A survey of agentic AI and cybersecurity: Challenges, opportunities and use-case prototypes. arXiv preprint arXiv:2601.05293 (2026)
7. Malatji, M.: A cybersecurity AI agent selection and decision support framework. arXiv preprint arXiv:2510.01751 (2024)
8. Adabara, I., Olaniyi, S.B., Nuhu, S.A., Ibarahim, D.Y., Venkateswarlu, M.: A Review of Agentic AI in Cybersecurity: Cognitive Autonomy, Ethical Governance, and Quantum-Resilient Defense. F1000Research 14, 843 (2025)
9. Russell, S., Norvig, P.: Artificial Intelligence: A Modern Approach. 4th edn. Pearson, Harlow (2021)
10. Devedžić, V.: Identity of AI. Discover Artificial Intelligence 2(1), 1–19 (2022)
11. Kaplan, A., Haenlein, M.: Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. Business Horizons 62(1), 15–25 (2019)
12. Von Solms, R., Van Niekerk, J.: From information security to cyber security. Computers and Security 38, 97–102 (2013)
13. Nickerson, R.C., Varshney, U., Muntermann, J.: A method for taxonomy development and its application in information systems. European Journal of Information Systems 22(3), 336–359 (2013)
14. Deng, G., Liu, Y., Mayoral-Vilches, V., et al.: PentestGPT: Evaluating and harnessing large language models for automated penetration testing. In: Proceedings of the 33rd USENIX Security Symposium, pp. 847–864. USENIX Association, Philadelphia, PA (2024)
15. Kong, H., Hu, D., Ge, J., Li, L., Li, T., Wu, B.: VulnBot: Autonomous penetration testing for a multi-agent collaborative framework. arXiv preprint arXiv:2501.13411 (2025)
16. NIST: NIST IR 8596 (iprd): Cybersecurity Framework Profile for Artificial Intelligence. National Institute of Standards and Technology, Gaithersburg, MD (2025). https://nvlpubs.nist.gov/nistpubs/ir/2025/NIST.IR.8596.iprd.pdf, last accessed 2026/04/10
17. Oesch, S., Austria, P., Chaulagain, A., Weber, B., Watson, C., Dixson, M., Sadovnik, A.: The path to autonomous cyberdefense. IEEE Security and Privacy 23(1), 38–46 (2024)
18. Feng, K.J.K., McDonald, D.W., Zhang, A.X.: Levels of autonomy for AI agents. arXiv preprint arXiv:2506.12469 (2025)
19. Han, S., Zhang, Q., Jin, W., Xu, Z.: LLM multi-agent systems: Challenges and open problems. arXiv preprint arXiv:2402.03578 (2024)
20. Pascoe, C., Quinn, S., Scarfone, K.: The NIST Cybersecurity Framework (CSF) 2.0. National Institute of Standards and Technology, Gaithersburg, MD (2024)

21. Microsoft: Microsoft Security Copilot: Generally Available. Microsoft Security Blog (2024). https://www.microsoft.com/en-us/security/blog/2024/03/13/microsoft-copilot-for-security-is-generally-available-on-april-1-2024-with-new-capabilities/, last accessed 2026/04/12
22. Darktrace Ltd.: Cyber AI: Augment your security team and stop novel threats. Darktrace (2024). https://www.darktrace.com/cyber-ai, last accessed 2026/04/12
23. CrowdStrike Inc.: Charlotte AI: The brain of the agentic SOC (2024). https://www.crowdstrike.com/en-us/platform/charlotte-ai/, last accessed 2026/04/12
24. SentinelOne Inc.: Purple AI: AI Security Analyst for Autonomous SecOps. SentinelOne (2024). https://www.sentinelone.com/platform/purple, last accessed 2026/04/12
25. Google LLC.: Google Security Operations (Chronicle SIEM). Google Cloud (2024). https://cloud.google.com/security/products/security-operations, last accessed 2026/04/12
26. European Parliament and Council of the European Union: Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union, L series (2024).

Original Research Article *https://doi.org/10.5281/zenodo.23474612*

A comparison of Flectra, Dolibarr and Odoo ERP systems based on HR module functionalities

Srđa Bjeladinović[1][0000-0001-6404-3497] and Marko Asanović[2][0009-0002-0427-9291]

[1] University of Belgrade, Faculty of Organizational Sciences, Jove Ilića 154, Belgrade, Serbia
[2] Adriatic University Bar, Faculty of traffic, communication and logistics / Municipality of Budva, Budva, Montenegro
srdja.bjeladinovic@fon.bg.ac.rs
asanovicmarko@live.com

Abstract. The business of modern companies, and governmental institutions, is almost unimaginable without information systems or ready-made, integrated software solutions, especially ERP systems. For businesses that cannot afford the high licensing costs of renowned ERP manufacturers, open-source ERP systems offer valuable alternatives. The popularisation of open-source ERP systems and the need to choose an adequate solution for a company's business (typically for a small or medium enterprise) have motivated research on comparative analysis, typically including the most popular solutions such as Odoo and Dolibarr. Flectra represents an emerging, fast-growing open-source ERP, and according to the literature available to us, has not been compared with the top three representatives of open-source ERPs. In this paper, exactly Odoo, Dolibarr and Flectra were compared in order to determine whether and to what extent Flectra can compete with the two mentioned and well-established systems, with a special focus on the functionality of the HR module. Through the analysis of available research in the field, available documentation, but also through practical testing, the criteria were defined, their weighting coefficients were determined, alternatives were evaluated, and the achieved ranks were analysed using the AHP (Analytic Hierarchy Process) method. The sensitivity analysis showed that although Odoo takes the dominant first place in achieved rankings, Flectra not only matches the Dolibarr system, but furthermore leads compared to Dolibarr in the overall ranking.

Keywords: ERP, Open-source, Human Resources, Comparative analysis.

1 Introduction

Due to rapid technological development and changes in consumer demand, modern businesses require continuous monitoring of many factors: from competitors and suppliers to employees and end users. To effectively manage all the factors and changes the dynamic market brings, it is necessary to integrate all sectors of the company and

provide the prerequisites for their smooth operation and cooperation through appropriate information resources.

With overall technological expansion and companies' desire to automate their processes, the market for ERP (Enterprise Resource Planning) systems is among the fastest-growing segments in software development. According to 2022 statistics, almost 47% of companies use an ERP system [1].

The proliferation of open-source ERP systems on the market for decision-makers in local government and private companies, who choose a solution based on their own capabilities and needs, significantly challenges the decision-making process. During this process, it is necessary to have appropriate quantitative data to support mentioned process. A company that chooses an ERP system must first conduct an internal analysis of processes, finances, and technical limitations, followed by an external analysis, to evaluate alternatives and define comparison criteria.

One of the important criteria for choosing an ERP system is the implementation cost. Open-source ERP solutions offer the benefits of working with an ERP system at a fraction of the initial cost of a vendor's ERP license. The absence of license fees and lower maintenance costs make open-source ERPs an attractive solution for companies (typically SMEs) that cannot afford significant financial investments in software. However, not all open-source ERP systems have been on the market for the same length of time; they do not have the same popularity, nor do they offer the same capabilities. For several years consecutively, Odoo and Dolibarr have been among the top three open-source ERP solutions [2], complementary to ERPNext. The authors [3] compared these three solutions, with special reference to the convenience of their application for managing the HR function of the company (one of the most important functions of each company), and concluded that there is a not-small gap between Odoo (which was ranked as the first) and Dolibarr (which was ranked as the last). In the meantime, the Flectra system emerged as a fast-growing ERP solution, a simplified Odoo branch that improves user interface usability and is targeted at small SMEs, reducing system complexity compared to Odoo. After analysing the literature available to us, we were unable to find a multi-criteria comparison of Flectra, Dolibarr and Odoo. Further, this paper was motivated to answer the following research questions:

1. Whether and to what extent Flectra can compete with established Odoo and Dolibarr systems, with a special focus on the functionality of the HR module?

2. Does the change in the significance of the selected comparison criteria affect the ranking of the Flectra system in relation to the market-leading systems Odoo and Dolibarr, and in what way if it does?

To answer the stated research questions, criteria frequently used by domain experts (papers analysed in more detail in following sections) were selected, and weight coefficients were determined with a special focus on HR functionality. After that, the AHP method and sensitivity analysis were conducted. Based on the obtained ranks, guidelines were provided to decision-makers when choosing among the three analysed systems.

The rest of the paper is organised as follows. In Chapter 2, related work is given. The basic concepts of the ERP system are given in Chapter 3, together with the display

of the selected systems. In Chapter 4, the criteria are defined; the steps of the implemented AHP method are presented in Chapter 5, while Chapter 6 contains the sensitivity analysis and a discussion of the results. Concluding remarks are given in Chapter 7.

2 Related Work

Modern ERP systems are based on the foundations set by EOQ (Economic Order Quantity), MRP (Material Requirements Planning) and MRPII (Manufacturing Resource Planning), which were further developed and improved by introducing the management process of all company resources and data, to automate and optimize the company's business processes and achieving the business goals of the company [4][5][6]. ERP systems are based on the principle of modularity, according to which the activities of related business processes, i.e. functions, are combined into individual but interconnected subsystems that can be changed separately without affecting the whole. Alongside a unique logical database [4][7], modularity has become one of the most recognisable characteristics of the ERP system. ERP systems consist of different modules, each representing a specific function or department in the company, and the inclusion or removal of a module does not disrupt the system's operation [7]. Today, ERP systems typically contain a large number of modules, but some modules stand out for their importance, such as financial management, material flow/service management and human resource management.

Open-source ERP systems are business process management software developed and distributed under open-source licenses. The term open-source software is defined as "any program or software whose source code is available for use and modification thanks to a license that meets the certification of the Open Source Initiative (OSI - Open Source Initiative) or is in the public domain" [8]. Open-source ERP systems were analysed by authors[9][10]. Some popular open-source ERP systems include Odoo, Dolibarr, Flectra, ERPNext, Openbravo, Adempiere, and Compiere. Odoo systems were dealt with by the authors [8][9][10][11], Flectra was analysed by [8][9][10], while Dolibarr was investigated by [8][9][10]. All the authors mentioned, as well as [12], conducted comparative analyses with other MCDM (Multi-criteria decision-making) methods. Authors [13] laid the foundation for a methodological approach to using the AHP in the decision-making process and established a framework for using this MCDM method for ERP system selection. However, the paper doesn’t compare ERP systems or conduct a genuine selection process. Paper [14] compares three multi-criteria decision-making approaches for the ERP selection process, comparing AHP, Fuzzy AHP (FAHP) and Neutrosophic AHP. The authors concluded that fuzzy approaches provide advantages in vague and uncertain situations, and that the classical AHP method is considered sufficient, transparent, and replicable.

Following this research and considering the applicability of AHP and the availability of dedicated software solutions, such as SuperDecisions, we have selected the AHP MCDM method for its ease of application and suitability for academic use in student research papers.

3 Open-source ERP systems

The three open-source systems studied in this paper are Odoo, Dolibarr, and Flectra.
Odoo ERP. Odoo is an open-source ERP system that integrates CRM and numerous other modules into a single platform [15]. A specific feature is that Odoo organises modules by application, so multiple applications often provide the functionality of traditionally defined modules. The described approach enables greater adaptability because you do not always need to install the entire module; you can install only the specific applications within the desired module, facilitating easy adaptation to the individual needs of SMEs. Odoo ERP is identified as a flexible, modular, open-source ERP system with high customizability, a modular architecture, and strong support. However, system effectiveness is highly dependent on implementation quality, the need for customisation, and training [16].
Flectra ERP. The Flectra open-source system was created as a "fork" of the Odoo ERP system, that is, based on the existing code and architecture of the Odoo ERP system, a new system was created with the idea of improving the existing shortcomings and developing a completely new independent project to provide users with additional flexibility and improve the user experience. The solution is recommended by the authors for SMEs with limited financial and technical resources, enabling the automation of core organisational processes such as sales, inventory management, and procurement. Flectra has a small ecosystem and a limited number of partners, which increases implementation risk [17].
Dolibarr ERP. The Dolibarr open-source system was created to offer a free solution to anyone who can provide hosting, and above all, to freelancers, regardless of the language they use. Today, Dolibarr is one of the most popular open-source ERP systems among SMEs, giving them the impression of working with a single module, even though the core system connects to and integrates additional modules as needed. Studies indicate that Dolibarr provides functional support for core business processes at a low implementation cost; however, studies also show that the system has limited scalability [17].

4 Defining criteria

To conduct a comparative analysis of the three selected open-source systems, it is necessary to define the comparison criteria, which is what this chapter addresses.

4.1 Criteria selection process

In conducting a comparative analysis of ERP systems, one key step towards making objective and informed decisions is to carefully define criteria and sub-criteria. This phase serves as the foundation for the comparison between the Odoo, Flectra, and Dolibarr ERP systems.

By defining criteria and sub-criteria, a structure is created to evaluate each system, respecting the specific research-relevant aspects. This process requires precision and attention to detail so that all relevant factors are taken into account.

This chapter discusses how the key criteria and sub-criteria were identified. Also, it explains why these criteria were chosen and how their definition helps in understanding the characteristics of each ERP system. Figure 1 shows the selected criteria and sub-criteria.

Fig. 1. Defined criteria and sub-criteria of the comparison

Several Authors [13][18][19][20] researched the ERP system selection process, emphasising the main reasons for choosing the correct system. Authors [13] define functional requirements as the most important criteria in ERP selection. Authors [18][19] state that financial considerations play a significant role in the adoption of systems, even for open-source solutions. Authors [20] place high significance on support, as the availability of technical support, maintenance, and updates affects system sustainability. Research conducted by the authors [21] shows that User interface usability and intuitiveness strongly affect ERP system effectiveness, as effectiveness depends on both the system's functionality and users' ability to operate it efficiently. Based on these research papers, criteria and sub-criteria are chosen. Finally, integration is important when an organisation uses external, non-specific ERP solutions from other sectors [13].

Their ranking and weight coefficient will be further elaborated in Section 5.

4.2 Description of selected criteria

HR functionalities criterion. The HR functionalities criterion is based on the supported functionalities of the HR module (of each of the analysed systems), which have been practically tested. Based on interviews with domain experts and existing experience with the ERP system, six use cases were identified. Each of them is categorised as a sub-criterion. Selected use cases are listed in Figure 1.

License cost criterion. When choosing the right software, cost plays a vital role. "They include acquisition, implementation and later maintenance costs" [11]. License costs are part of acquisition costs, and the analysis in this paper focuses on them (hence the absence of sub-criteria).

All three systems analysed in this paper are available in a completely free version (Community version), which is limited in functionality. All three systems also include additional functionalities in the commercial (Enterprise) version.

Within the commercial, i.e. Enterprise version, Odoo offers three packages: free application (One App Free), Standard, and customised. The free app package includes one app for unlimited users. The difference from the Community version is that this package includes only one application. In contrast, the Community includes a limited set of features, and some features found in the Enterprise package are not available in the Community version [21]. The standard package includes all applications and differs from the customised package in hosting: it includes only cloud hosting.

Flectra offers three packages within the Enterprise version: free, standard and custom. Within the free package included with the Enterprise version, all applications are available to a single company and can be used by up to five users. The level of support is low and relies on the community of experts and users of this system. The standard package includes professional add-ons and a support contract for just one company. Support is unlimited via email and ticket services. The package also includes training as needed, upon request. The Custom package includes the standard package with 100 GB of memory, plus the option to customise code, an external API, and training capabilities [22].

All functional modules are available within the Dolibarr Basic package. Support for migrations, technical support via tickets, and functional support via community forums are provided [23].

Support criterion. The criterion of support is reflected in the provision of instructions and advice during the system's implementation and maintenance, assistance with various problems, organisation of training and webinars, and the creation of user materials [10]. Within the support criteria, the quality of assistance provided by suppliers and support to users in solving problems or answering questions was considered, including responsiveness, expertise, courtesy, and effectiveness.

UI usability and intuitiveness. The author [9] states that the application's ease of use, i.e. usability, is one of the key aspects to consider when developing software. The author defines usability as the end user's perception of how easy the application is to use, how much control they have over it, how easily they can navigate it, and how much they enjoy it.

A user interface is considered intuitive when users understand its behaviour and effects without thinking, experimenting, help, or special training [24]. For an intuitive user interface, it is important to consider the application's adaptability to the user [11].

Integration criterion. Within the integration criteria, two sub-criteria were processed: Integration with other systems and the import and export of documents in different formats.

Software integration involves connecting software with other important applications, creating a unified technological ecosystem [25]. Within the sub-criterion of the same name, it was analysed whether the given system is predisposed for integration with other specialised systems through the presence of an external API (Application Programming Interface). The evaluation of alternatives under this sub-criterion was

conducted using relevant data on available integrations for each system from the ERP system comparison platform [26], along with survey responses from users.

Since the manipulation of large amounts of data is an indispensable part of business processes, the import and export of documents are essential to the system's functioning. The corresponding sub-criterion assessed whether the system supports common file types (.csv and .xlsx) for import and export.

5 Implementing the AHP method

After identifying the criteria and sub-criteria, the AHP method was implemented. The key steps in implementing the AHP method are shown in Figure 2. The AHP method was implemented through the SuperDecisions expert decision-making tool.

Fig. 2. Realisation of the AHP method[27]

5.1 Specifying a hierarchical model for the problem

Below is a hierarchical model of the problem, in which the goal, alternatives, criteria, and sub-criteria are defined.

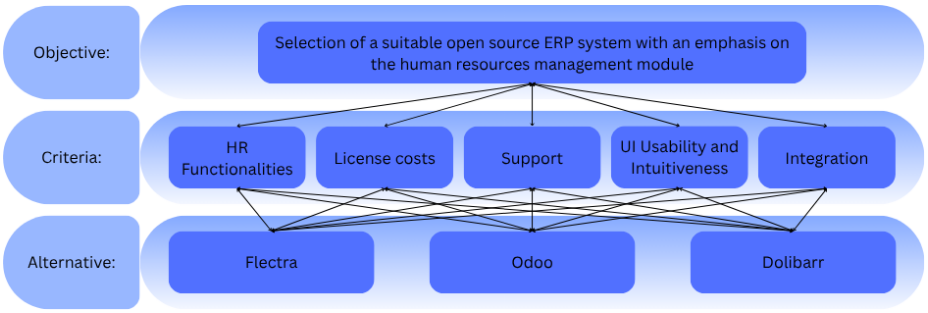


Fig. 3. AHP Hierarchical model

5.2 Ranking criteria and sub-criteria

Papers by established authors were analysed to define and rank criteria and sub-criteria. Based on frequency of occurrence and importance of criterion across studies [11][13][18][19][20], ranks were assigned to each criterion to reflect their relative importance in the decision-making process. Authors [13] show that criteria exhibit a hierarchical structure with dominant factors, where Functionality is suggested as the dominant factor by the majority of Authors. A geometric weighting scheme is used to emphasise the importance of each criterion relative to the others. Rank 1 is the most important criterion for choosing an ERP system, and rank 5 is the least important criterion for choosing an ERP system, as presented in Table 1. Furthermore, at the level of individual criteria, ranks were assigned to sub-criteria indicating their importance within each criterion.

Table 1. Ranking criteria.

Criterion	Rank
HR functionalities	1
License costs	2
Support	5
UI usability and intuitiveness	4
Integration	3

Applying the criteria, their ranks, and the geometric weight scheme to the SuperDecisions software returned the weights shown in Figure 4.

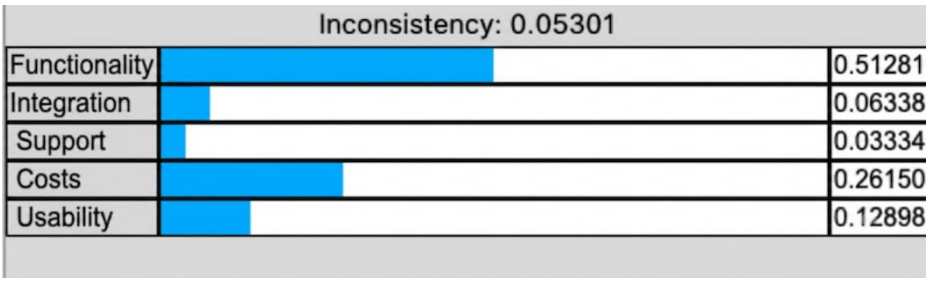


Fig. 4. The ranking of the criteria in relation to their importance for the goal achievement in the expert tool SuperDecisions.

The upper limit of inconsistency is 0.1; any index exceeding this value indicates an inconsistent assessment, in which case an inconsistency analysis is required and the assessment corrected. The achieved inconsistency index for the compared criteria is 0.05, which falls within the allowed range.

Ranking sub-criteria for the HR functionalities criterion. In this sub-chapter, a comparative analysis of the sub-criteria within the HR functionality criteria was conducted.

Table 2. Ranking sub-criteria for HR functionalities.

Criterion	Rank
Employees administration	1
Managing leaves	2
Recruitment	3
Managing costs	4
Questionnaire	5
Contract administration	6

Based on the sources [28] and [29], the sub-criteria are ranked in Table 2, where rank 1 represents the most important sub-criterion and rank 6 the least important sub-criterion within the functionality criterion. The inconsistency index is 0.0123.

Ranking sub-criteria for the Integration criterion. According to the source [25], the sub-criteria are ranked in Table 3, with rank 1 indicating the most important sub-criterion within the integration criterion and rank 2 indicating the least important.

Table 3. Ranking sub-criteria for Integration.

Criterion	Rank
Integration with other systems	1
Importing and exporting documents in different formats	2

5.3 **Evaluation process for each criterion.**

Based on data from practical tests, manufacturer websites, analysed literature, and comparison platforms (listed in the subsections above), assessments of each alternative for each criterion or sub-criterion (for HR functionalities and Integration criteria) were conducted. The collected data was entered into the SuperDecisions tool, which allows discrete values to be entered. The procedure mentioned is illustrated using the example of monthly license costs per user (Figure 5). According to the specified criteria, specific USD ($) values were entered into the tool for each of the three systems (11.8 for Odoo, 19 for Flectra, and 14 for Dolibarr). Afterwards, the tool compared the values for each pair of alternatives and generated a table. For example, dividing the value of 19 (Flectra) by 11.8 (Odoo) yielded a quotient of 1.6101; i.e., Odoo is 1.6101 times more significant than Flectra. Given that the invert option is enabled (lower values are preferred), the arrow to the displayed recommendation points to the Odoo solution (Figure

3). According to the same principle, it was determined that Odoo is 1.864 times more preferable than Dolibarr and that Dolibarr is 1.3571 times more preferable than Flectra. The same principle was used in evaluating other criteria, i.e., sub-criteria.

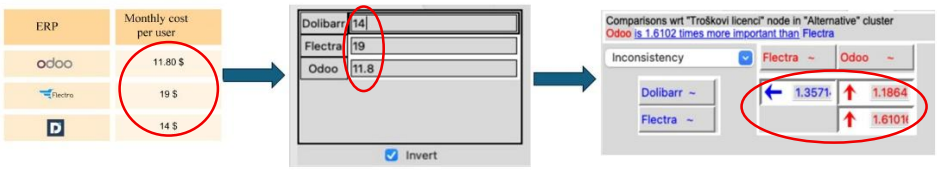


Fig. 5. Example of the evaluation process for the selected criterion

5.4 Evaluation of alternatives

After ranking the criteria and sub-criteria within each criterion, the process of evaluating alternatives for each criterion and sub-criterion is undertaken. The output of this step is a matrix showing each system's ratings for each sub-criterion on the Saaty scale. Based on the grades, the system rankings were established for each criterion and sub-criterion.

For each of the five defined criteria, three analysed systems were evaluated based on early data sources.

Evaluation of alternatives for HR functionalities criterion. The evaluation of alternatives based on sub-criteria within the HR functionality criteria was conducted using six practical use cases of using the human resources management module, as listed in Figure 1. Figure 6 shows pairwise comparison matrices from SuperDecisions software for the HR functionalities subcriteria, previously identified as the most dominant factor.

Employee administration. The Odoo and Flectra systems share the same basic functionality for maintaining employee records. At the same time, Odoo offers additional options, such as the ability to enter an employee's CV, which can be updated manually.

Managing leaves. The Odoo and Flectra systems distinguish between two types of leave requests (allocation request and leave request), whereas Dolibarr has only one type of leave request (Leave request).

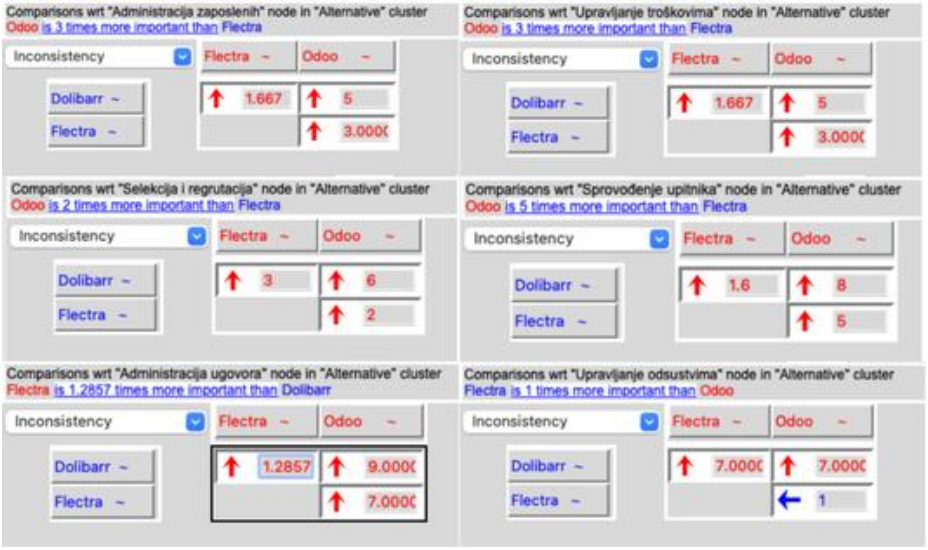


Fig. 6. Pairwise matrices for HR Functionalities subcriteria

Recruitment. Odoo and Flectra, in addition to tracking candidate applications for vacant positions, significantly streamline the process with features such as scheduling interview meetings and attaching documents. The two systems have a high degree of functional overlap, but Odoo allows you to create applicant biographies, which is important for the recruitment process. Dolibarr has fewer features than Odoo and Flectra.
Managing costs. Odoo can convert currencies, which sets it apart from the other two systems. Dolibarr differs from Odoo and Flectra in that it requires manual account entry to generate reports.
Questionnaire. Functionalities for conducting questionnaires and analysing results are available in both Odoo and Flectra, but Odoo stands out for the ability to create a live session to conduct the survey. Dolibarr has limited functionality: answers are predefined as affirmative or negative, which reduces flexibility when creating a survey.
Contract administration. Odoo deviates significantly from the Flectra system because it can track employee history, i.e., Recording of promotions and salary increases. Dolibarr lacks a contract management function, which is a significant drawback.
Evaluation of alternatives for the License costs criterion. The evaluation of alternatives based on license costs uses the standard package prices for all three systems, taken from the official websites for each [21][22][23].
Evaluation of alternatives for the Support criterion. The evaluation of alternatives in this work is based on data from the open-source ERP system comparison platform [16] and surveys of platform users.
Evaluation of alternatives for UI usability and intuitiveness criterion. The evaluations of the alternatives for usability and user interface intuitiveness were assigned using data from the ERP system comparison platform [16] and survey results.

Evaluation of alternatives for the Integration criterion. The evaluation was based on data on supported document import and export formats in the official documentation for the ERP systems considered [30][31][32].

5.5 Ranking of alternatives

This chapter presents the final step in implementing the AHP method, through which the ranks of the considered ERP systems were obtained with respect to the defined criteria and sub-criteria. Finally, the rankings of the alternatives relative to the target were obtained. It is important to emphasise that the obtained ranks represent only one possible ranking outcome among the three analysed systems, based on the defined criteria and with a special focus on the functionality of the HR module. Changing the criteria or their relative importance can affect the outcome, as analysed in the subsection on sensitivity analysis. Nevertheless, based on the defined criteria and the results achieved, it is possible to derive guidelines and recommendations for selecting one of the analysed open-source ERP systems.

Given that the first selected criterion (HR functionalities) comprises five sub-criteria, a ranking was assigned to each. The ranking based on the specified criteria is shown in Figure 7.

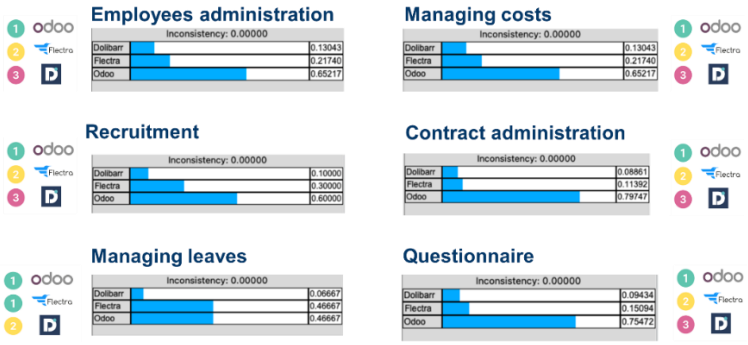


Fig. 7. Ranking alternatives for sub-criteria of HR functionalities (1st criterion)

The ranking of alternatives according to the criteria License costs, Support, and UI usability and intuitiveness is shown in Figure 8.

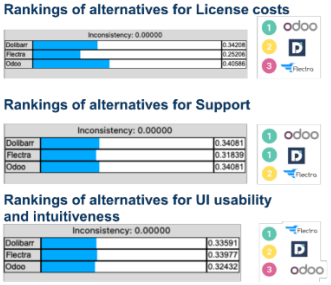


Fig. 8. Ranking alternatives for License costs, Support and UI usability and intuitiveness

The ranking of alternatives by sub-criteria within the Integration criterion is shown in Figure 9. For the import and export sub-criteria, all three analysed systems are ranked first. The explanation lies in the fact that all three systems support both tested data import and export formats, and therefore all achieve the same performance and ranking.

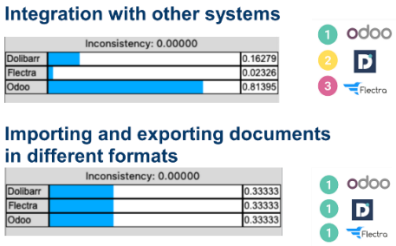


Fig. 9. Ranking alternatives for sub-criteria of Integration (5th criterion)

The last step of the AHP method is to calculate the priorities of the criteria, sub-criteria, and alternatives to obtain the ranking of the alternatives with respect to the objective. Figure 10 shows the order of the alternatives Odoo, Flectra, and Dolibarr in relation to the target, i.e., the selection of a suitable open-source ERP system.

It can be seen that Odoo is the most preferred alternative, followed by Flectra, with Dolibarr the least preferred. It is important to point out once again that this order is not universal and resulted from the previously defined criteria and sub-criteria established in this paper, along with their ordering. Depending on the selected criteria and their order according to the decision-maker's requirements, a different order of the same alternatives can be obtained.

Name	Graphic	Ideals	Normals	Raw
Dolibarr		0.411623	0.211564	0.082123
Flectra		0.534000	0.274462	0.106538
Odoo		1.000000	0.513974	0.199510

Fig. 10. The order of alternatives in relation to the goal

6 Sensitivity analysis and discussion of the results

In this chapter, a sensitivity analysis is performed, and the results obtained are discussed.

6.1 Sensitivity analysis

The sensitivity analysis is used to graphically display the influence of the criteria's weighting coefficients on the overall ranking, i.e., the ranking of alternatives relative to the goal. Figure 11 presents the sensitivity analysis for each established criterion.

It is immediately apparent that Odoo is the top-ranked solution across all criteria. According to the first four criteria, Odoo remains the top-ranked solution regardless of changes in the weight coefficients of individual criteria. Only for the last criterion (UI usability and intuitiveness) is it possible to change the outcome, but this requires a significant increase in the weight coefficient.

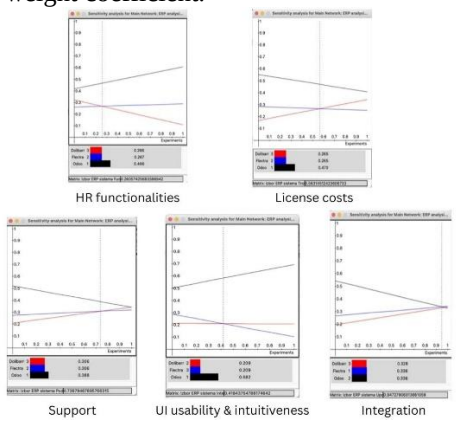


Fig. 11. Sensitivity analysis according to five established criteria

If ranks 2 and 3 are observed under individual criteria, the result is less certain. Namely, for the criteria of license costs, support, integration, and UI usability and intuitiveness, the weight coefficients need to be increased above 0.56, 0.737, 0.418, and 0.96, respectively, for Dolibarr to overtake Flectra. Regarding the HR functionality criteria, the situation is the opposite, and the weight coefficient needs to be reduced below 0.26; Dolibarr becomes a better alternative than the Flectra system.

7 Conclusion

In this paper, an approach to selecting an ERP system for human resources management is presented. As alternatives, Odoo, Flectra and Dolibarr systems are defined. First of all, the order of the selected systems was obtained in relation to each of the criteria, the sub-criteria as presented in Figure 1, and the order in relation to the goal.

Finally, the order of the selected systems relative to the goal, after analysing in SuperDecisions, is Odoo, Flectra, and Dolibarr, with Odoo the most preferred alternative for the defined set of criteria and Dolibarr the least preferred. An important note is that this order results from previously defined criteria and sub-criteria and varies depending on the specific criteria and the decision maker's priorities, as shown by the sensitivity analysis conducted. However, for the analysed systems and categories, even significant changes in the weight coefficients of the selected criteria do not affect the outcome of the first-ranked system.

Comparing the achieved results with the previous research [3], it is interesting that the ranks of the two systems, which are present in both analyses, are the same: Odoo is

ranked first, while Dolibarr is ranked third. The significance of the correlation between this result and the previous research is enhanced by the fact that the selection criteria differ between the two papers and are of varying significance. In the analysis conducted in this paper, Flectra was ranked as the second solution, with better performance than the Dolibarr system according to the defined goal and criteria. Based on the above, we can conclude that the Flectra system can be treated as an adequate alternative to leading open-source ERP systems. Although Odoo continues to take precedence, based on defined criteria and weights, Flectra represents an alternative that can be recommended to decision-makers compared to Dolibarr, especially if the focus is on HR functionality criteria. Significant emphasis needs to be given to other criteria, such as license cost and integration, with functionality having significantly lower importance for Dolibarr to overtake Flectra. The comparison shows that Flectra, even though relatively new, can compete with top-tier open-source ERP solutions, especially since some of the lower marks Flectra achieved are due to low system maturity, such as in the support criterion. It is also important to note that Flectra scored highest in UI usability and intuitiveness, even outperforming Odoo.

The goal of this paper was to demonstrate that the AHP method can be applied to the selection process for an open-source ERP system and to provide decision-makers with guidelines on which system has priority, depending on the criteria prioritised within the company, based on the sensitivity analysis conducted. The continuation of the research would involve applying additional multi-criteria decision-making methods, such as TOPSIS, VIKOR, ELECTRE, or PROMETHEE, and comparing the results.

Disclaimer. This paper includes results obtained during Anđelka's master's thesis research at the Faculty of Organizational Sciences, University of Belgrade, Belgrade, Serbia.

Acknowledgments. This research was supported by the University of Belgrade – Faculty of Organizational Sciences and, in part, by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia through institutional funding (grant number: 200151).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Strategic Market Research: ERP Software Market Report (2022), https://www.strategicmarketresearch.com/market-report/ERP-software-market, last accessed 2026/04/19.
2. ERP Softnic: Best open-source ERP software 2026. https://erpsoftnic.com/erp-software/open-source-erp-software, last accessed 2026/05/25
3. Masal, M., Bjeladinović, S.: A comparative analysis of the human resources module in open-source ERP systems. In: Proceedings of the XVIII International Symposium SymOrg 2022, pp. 31–33. University of Belgrade, Faculty of Organizational Sciences, Belgrade (2022)
4. Bjeladinović, S.: Materijali sa predavanja iz Integrisanih softverskih rešenja. Fakultet organizacionih nauka, Univerzitet u Beogradu, Beograd, Srbija (2018)

5. Rejman Petrović, D.: ERP sistemi u funkciji unapređenja kvaliteta poslovanja. Nacionalna konferencija o kvalitetu, University of Kragujevac, pp. A15–A22 (2009)
6. Jovičić, B., Vlajić, S.: Evolucija ERP sistema. InfoM – Časopis za informacione tehnologije i multimedijalne sisteme 22, 18–22 (2007)
7. Olaoye, F., Potter, K.: Enterprise Resource Planning (ERP) Systems. EasyChair Preprint No. 12516 (2024).
8. Fougatsaro, V. G.: A study of open source ERP systems. Master's thesis, Blekinge Institute of Technology (2009)
9. Pantelić, O., Maljković, D.: Komparativna analiza ERP sistema otvorenog koda. Časopis o društvenom i tehnološkom razvoju 11, 9–18 (2019)
10. Herzog, T.: A Comparison of Open Source ERP Systems. Master's thesis, Vienna University of Technology, Vienna, Austria (2006)
11. Huber, S.: ERP software system comparison between Odoo and Microsoft Dynamics NAV. Bachelor's thesis, Hame University of Applied Sciences (HAMK), Hameenlinna, Finland (2021)
12. Gómez-Llanez, C. Y., Díaz-Leal, N. R., Angarita-Sanguino, C. R.: A comparative analysis of the ERP tools, Odoo and Openbravo, for business management. AiBi Revista de Investigación, Administración e Ingeniería 8(3), 145–153 (2020). https://doi.org/10.15649/2346030X.789
13. Wei, C. C., Chien, C. F., Wang, M. J. J.: An AHP-based approach to ERP system selection. International Journal of Production Economics 96(1), 47–62 (2005). https://doi.org/10.1016/j.ijpe.2004.03.004
14. Slamaa, A. A., El-Ghareeb, H. A., Aboelfetouh, A.: Comparative analysis of AHP, FAHP and Neutrosophic-AHP based on multi-criteria for adopting ERPS. Neutrosophic Sets and Systems 41, 64–88 (2021). https://doi.org/10.5281/zenodo.4625668
15. Captivea: Odoo ERP – Powerful Business Software, https://www.captivea.com/odoo, last accessed 2026/04
16. Mortiga, A. S. S., Segura, F., Rombaoa, J., Dela Cruz, M.: Evaluating ERP implementation feasibility: A case study of Odoo adoption in a Philippine distribution firm using the TOE framework. International Journal of Research and Innovation in Social Science 9(12), 3828–3832 (2025). https://doi.org/10.47772/IJRISS.2025.91200297
17. Šimović, V., Varga, M., Soleša, D.: Analysis of possible advantages and constraints of ERP systems. Ekonomija: teorija i praksa 13(4), 41–56 (2020). https://doi.org/10.5937/etp2004041S
18. Ratkevičius, D., Ratkevičius, Č., Skyrius, R.: ERP selection criteria: theoretical and practical views. Ekonomika 91(2), 97–116 (2012). https://doi.org/10.15388/Ekon.2012.0.893
19. Molnár, B., Szabó, G., Benczúr, A.: Selection process of ERP systems. Business Systems Research 4(1), 36–48 (2013). https://doi.org/10.2478/bsrj-2013-0004
20. Raza, A., Capretz, L. F., Ahmed, F.: An empirical study of open source software usability: The industrial perspective. arXiv preprint arXiv:1511.08844 (2015). https://arxiv.org/abs/1511.08844
21. Odoo S.A.: Pricing Plan, https://www.odoo.com/pricing, last accessed 2026/04
22. FlectraHQ Inc.: Pricing, https://flectrahq.com/pricing, last accessed 2026/04
23. Dolibarr ERP/CRM Project: Pricing (DoliCloud), https://www.dolicloud.com/en-faq-pricing.php, last accessed 2026/04
24. Bhalla, M.: Intuitive UI and how it can help improve UX. Hongkiat, https://www.hongkiat.com/blog/defining-designing-intuitive-ui/, last accessed 2026/04

25. Robocoder Corporation: Key Features and Characteristics of ERP Software, Medium (1ERP blog), https://medium.com/1erp/key-features-and-characteristics-of-erp-software-5453c8f58f74, last accessed 2026/04
26. G2.com: Compare Dolibarr ERP CRM, Flectra, and Odoo ERP, https://www.g2.com/compare/dolibarr-erp-crm-vs-odoo-odoo-erp-vs-flectra, last accessed 2024/08
27. FasterCapital: Decision Making: Navigating Choices – The Role of Analytic Hierarchy Process in Decision Making, https://fastercapital.com/content/Decision-Making--Navigating-Choices--The-Role-of-Analytic-Hierarchy-Process-in-Decision-Making.html, last accessed 2026/04
28. Sherrer, K.: 15 HR system features to look for. TechnologyAdvice, https://technologyadvice.com/blog/human-resources/hr-system-features/#centralized-database, last accessed 2024/08
29. OpenHRMS: Top ten features of human resource management system, https://www.openhrms.com/blog/top-ten-features-of-human-resource-management-system/, last accessed 2026/04
30. FlectraHQ Inc: Export Data from Flectra, https://doc.flectrahq.com/2.0/applications/general/flectra_basics/export-data.html, last accessed 2023/01/29
31. Odoo S.A.: Export and import data, https://www.odoo.com/documentation/17.0/applications/essentials/export_import_data.html, last accessed 2023/01/29
32. Dolibarr ERP/CRM Project: Import and export, https://wiki.dolibarr.org/index.php/Imports/Exports, last accessed 2023/01/29

THEMATIC SECTION

Challenges of Green and Digital Transformation

Resilient and Intelligent Organizations in Times of Permanent Disruption

SYMORG 2026 • CONFERENCE PROCEEDINGS

Original Research Article *https://doi.org/10.5281/zenodo.23474660*

The Influence of ESG Performance on Attracting Foreign Direct Investments

Luka Petrović[1][0009-0005-4503-1980] and Marko Ćirović[2][0000-0002-5073-3902]

[1,2] Faculty of Organizational Science, Jove Ilića 154, Belgrade, Serbia

Abstract. The growing importance of sustainability considerations in international investment decisions has intensified interest in the relationship between sustainability performance and foreign direct investment (FDI) inflows. This paper examines the relationship between sustainability performance and FDI dynamics through a comparative analysis of Germany, Poland, and Serbia, representing economies at different levels of institutional and economic development. The study is based on a descriptive analysis of secondary data obtained from the Sustainable Development Report and the World Bank database. Sustainability performance is measured using the SDG Index, while FDI inflows are used as an indicator of investment attractiveness. The results show that all three countries improved their SDG Index values over the observed period, although with different levels of progress and institutional maturity. Germany maintained the highest sustainability performance, Poland recorded steady improvement, while Serbia achieved the most pronounced relative progress. At the same time, FDI inflows displayed considerable fluctuations across all countries, indicating that sustainability performance alone cannot explain investment dynamics. The findings suggest that higher sustainability performance may contribute to investment attractiveness; however, FDI inflows are also influenced by broader macroeconomic, institutional, and regulatory factors. The study contributes to the existing literature by providing a comparative perspective on the relationship between sustainability indicators and FDI in developed, transition, and EU candidate economies. The results highlight the importance of considering sustainability performance together with institutional quality and economic conditions when assessing national investment attractiveness. Due to the descriptive nature of the analysis, the findings should be interpreted as indicative associations rather than evidence of causal relationships.

Keywords: ESG performance, foreign investments, economic development.

1 Introduction

In recent decades, the global economy has been facing challenges that cannot be adequately explained through traditional economic indicators alone. Climate change, growing social inequalities, resource scarcity, and increasing regulatory requirements have intensified interest in sustainable development among governments, international

organizations, businesses, and investors [1]. Sustainable development is commonly defined as development that meets the needs of the present without compromising the ability of future generations to meet their own needs [2]

In the business environment, sustainability is increasingly operationalized through the Environmental, Social and Governance (ESG) framework. ESG principles provide a structured approach for evaluating environmental responsibility, social performance, and corporate governance practices. As sustainability considerations become more important in economic decision-making, ESG factors are increasingly incorporated into investment strategies and business operations.

At the global level, sustainability objectives are reflected through the United Nations Sustainable Development Goals (SDGs), adopted in 2015 as part of the 2030 Agenda. The SDGs provide a comprehensive framework for measuring progress toward economic, social, and environmental sustainability. Although ESG and SDG concepts are closely related, they represent different analytical frameworks. ESG indicators are primarily used to assess sustainability performance at the company level, while SDG indicators measure the progress of countries in achieving sustainable development objectives.

At the same time, the structure of global foreign direct investment (FDI) flows has been changing. Alongside traditional determinants such as market size, labor costs, and institutional quality, sustainability-related factors are receiving increasing attention from investors. According to the Global Sustainable Investment Review [3], sustainable investment assets continue to grow globally, while UNCTAD [4] highlights the increasing importance of investments related to renewable energy, sustainable infrastructure, and green transition. In addition, OECD initiatives promoting the quality of FDI emphasize the importance of aligning investment policies with sustainable development objectives [5].

Against this background, understanding the relationship between sustainability and foreign direct investment has become increasingly relevant for both policymakers and researchers. This study focuses on Germany, Poland, and Serbia, three countries characterized by different levels of economic development, institutional maturity, and integration into European sustainability frameworks. Germany represents a highly developed economy with a long tradition of sustainability policies, Poland represents a rapidly developing European Union member state, while Serbia represents a transition economy and candidate country for EU membership.

The main objective of this paper is to examine the relationship between sustainability and foreign direct investment by comparing SDG Index trends and FDI inflows in Germany, Poland, and Serbia during the period 2010–2024.

2 Literautre review

Sustainability and its effect on foreign direct investments (FDI) have gained increasing attention in academic research over the last decade. Traditional studies on the determinants of FDI have primarily emphasized market size, labour costs, institutional quality and macroeconomic stability as the main factors influencing investment

decisions. However, recent research suggests that sustainability-related factors are becoming increasingly important in explaining international capital flows.

A substantial body of empirical evidence indicates that sustainability performance is often associated with improved economic and financial outcomes. Friede, Busch and Bassen [6], through a meta-analysis of more than 2,000 empirical studies, found that ESG performance is positively related to corporate financial performance in most cases. However, this relationship is not uniform across all ESG dimensions. Khan, Serafeim and Yoon [7] argue that value creation depends primarily on sustainability issues that are financially material to a company's core business activities.

The growing relevance of sustainability has also transformed investment practices. ESG investing has evolved from a niche strategy into a mainstream investment approach, although significant challenges remain regarding data quality, comparability and standardization [8]. At the same time, disagreement among ESG rating agencies creates additional uncertainty for investors when assessing sustainability performance [9]. Pedersen, Fitzgibbons and Pomorski [10] suggest that integrating sustainability considerations into investment decisions can reshape the traditional risk–return framework without necessarily reducing financial performance.

The literature also emphasizes the importance of institutional factors in shaping sustainability outcomes. Corporate sustainability practices are strongly influenced by national legal systems, regulatory frameworks and cultural norms [11]. Effective ESG integration requires incorporating sustainability considerations into traditional financial analysis rather than treating them solely as ethical screening mechanisms [12]. Furthermore, sustainability considerations are becoming increasingly important in long-term capital allocation decisions [13].

More recent studies have extended the discussion from the firm level to the national level. Caetano, Marques and Afonso [14] examined the relationship between sustainable development and FDI in European Union countries and found evidence that sustainability-related indicators are associated with both inward and outward FDI flows, although the relationship is not always linear and differs across countries. Similarly, Bannour and Abdelkawy [15] highlight the role of sovereign ESG performance in attracting foreign investment and argue that trade openness can strengthen the relationship between sustainability and FDI.

Recent literature further suggests that sustainability should not be analyzed independently from broader economic and institutional conditions. Studies indicate that environmental performance, renewable energy development and governance quality may influence investment attractiveness, but their effects often depend on market characteristics and institutional capacity [16]. In addition, several authors emphasize that the relationship between sustainability and FDI may be nonlinear and influenced by external factors such as economic crises, geopolitical uncertainty and structural market changes.

Despite the growing body of literature examining sustainability and foreign direct investment, comparative studies involving countries at different stages of economic and institutional development remain limited. In particular, there is a lack of research comparing highly developed economies, emerging European Union member states and EU candidate countries within a common analytical framework. Therefore, this study

contributes to the existing literature by comparing Germany, Poland and Serbia and examining the relationship between SDG Index performance and foreign direct investment inflows.

3 Metodology

The methodology of this study is based on a comparative and descriptive analysis aimed at examining the relationship between sustainability and foreign direct investment inflows in three countries with different levels of institutional and economic development: Germany, Poland, and Serbia. These countries were selected due to their differing levels of economic development, the degree of alignment of their regulatory frameworks with European Union standards, and their positions in international sustainability rankings. Measuring sustainability at the country level is complicated by significant methodological differences among various assessment systems and available indicators [17].

The study uses secondary data collected from relevant international sources, including UNCTAD, OECD, Eurostat, and the World Bank. In addition, data from the Sustainable Development Report related to the achievement of the Sustainable Development Goals (SDGs) are used. It is important to emphasize that the SDG Index does not represent a direct measure of ESG performance. ESG and SDG are related but methodologically distinct concepts. While ESG is most commonly used to assess the sustainability performance of companies and investment portfolios, the SDG Index measures countries' progress toward achieving the United Nations Sustainable Development Goals. In this study, ESG serves as the theoretical framework, while the SDG Index is used as an indicator of sustainability at the country level. The observed period covers the years 2010–2023, providing insight into long-term trends in the movement of the observed variables.

The analytical procedure involves the application of descriptive statistics, trend analysis, and comparative assessment of the selected countries. Particular attention is paid to identifying patterns of association between sustainability indicators and foreign direct investment inflows, accompanied by a qualitative interpretation of the institutional and regulatory context. The comparative approach enables an examination of the differences between developed and transition economies, while descriptive statistical indicators provide a clear presentation of trends in the observed variables.

The limitations of the study include the fact that the SDG Index represents an indirect rather than a direct measure of ESG performance, the absence of a unified methodology for assessing sustainability at the country level, and potential differences in the statistical coverage of foreign direct investment data. Furthermore, due to the descriptive nature of the research, the results indicate possible patterns of association between sustainability and foreign direct investment inflows but do not allow conclusions regarding causal relationships. Nevertheless, the use of multiple data sources and data triangulation significantly reduces methodological bias and supports the validity of the findings.

4 Results and analysis

Firms and countries exhibiting stronger environmental responsibility tend to face lower perceived risk, which is reflected in reduced financing costs and improved investor confidence [18]. Climate-related risks have become financially material, as investors increasingly react to unexpected exposure to environmental shocks and transition risks [19]. Sustainable industrial development increasingly depends on aligning investment strategies with environmental and social objectives in order to ensure long-term competitiveness and resilience [20].

The Sustainable Development Goals (SDGs) represent a universal set of 17 goals and 169 sub-goals that all member states of the United Nations adopted in 2015 as part of the 2030 Agenda. These goals were created as successors to the Millennium Development Goals (MDGs) and have a broader and integrated approach, including not only development and social dimensions, but also economic and environmental aspects of global development. The special importance of the SDGs is reflected in the integration of three dimensions of sustainable development, economic, social and environmental. For example, economic goals (such as goal 8: decent work and economic growth) are closely linked to social development goals (eg goal 10: reducing inequality), but also to environmental goals (eg goal 13: climate action).

Such interconnectedness allows development not to be viewed individually, but as a complete system. The implementation of sustainable development goals depends on the political will, institutional capacities and economic resources of individual countries. Every year, the UN publishes progress reports, from which it is evident that no country is still fully on its way to achieving all the goals by 2030. Nevertheless, the SDGs have become a key instrument for shaping public policies, corporate strategies, international development programs and academic research.

According to the Sustainable Development Indicators (SDG), Germany, Poland and Serbia are at different stages of achieving the Sustainable Development Goals, with each of these countries showing specific patterns of progress and challenges. Germany generally records good results in most areas, but with certain negative trends in social indicators (Figure 1). The greatest progress is seen in the areas of affordable and clean energy, decent work and economic growth, industry and innovation, and climate action. These sectors are experiencing stable and moderate growth, which shows that Germany continues to strengthen its economy through sustainable technologies and green policies. However, it is worrying that the areas of poverty reduction and inequality are deteriorating. Also, the quality of education, health and gender equality are stagnating, which indicates that Germany, although a developed country, has problems in maintaining social cohesion and equal opportunities for all. Overall, Germany is making steady but uneven progress.

Fig. 1. Trends in the Sustainable Development Goals in Germany

Poland, on the other hand, shows steady but moderate progress (Figure 2). The country is making improvements in poverty eradication, climate action and affordable energy, while significant progress has been made in the area of protecting life on land, where targets have already been met. These results speak of Poland's success in strengthening environmental and energy policies, as well as in reducing poverty. However, Poland faces challenges in sustainable production and consumption, which is the area with the greatest problems. Also, the areas of health, industry and innovation are stagnating, which shows that the transition to a high-tech and environmentally sustainable economy is going slowly. Although Poland has made steady progress in most areas, its development is still uneven. Success has been achieved in social and environmental policies, while economic transformation towards sustainable models has yet to be strengthened.

Fig. 2. Trends in Sustainable Development Goals in Poland

Serbia is in a much more difficult situation compared to Germany and Poland. Most of the sustainable development goals show significant or major challenges (Figure 3). However, there is visible progress in some areas, primarily in eradicating poverty, reducing inequality, developing affordable energy and sustainable cities. These developments indicate that Serbia is making gradual progress in socio-economic development and infrastructure modernization. However, it is worrying that the areas

of gender equality and climate action recorded a deterioration, while living on land and strengthening institutions were marked as areas of major challenges. Also, data on life under water is missing, which shows a weaker system of monitoring sustainable policies in some segments. Other goals, such as health, education, innovation and partnerships, are in the zone of moderate challenges, with slow progress. Overall, Serbia is on the path of gradual development, but with numerous structural problems that slow down the achievement of sustainable development goals, especially in the areas of ecology, institutional stability and equality.

Fig. 3. Trends of sustainable development goals in Serbia

Germany is at the highest level of the SDG index, with values ranging between 82 and 84 points. The line is almost flat, which indicates a stable and constant level of development. Germany has already largely achieved the goals of sustainable development and is in the phase of maintaining the current situation, with occasional minor improvements. Such stability indicates effective public policies, developed institutions and a high degree of coordination of social, economic and environmental development. **Poland** occupies an intermediate position between Germany and Serbia and shows stable growth throughout the entire period. From around 80 points in 2015, Poland is gradually approaching 82 points by 2024. This trend shows that the country is making progress in aligning itself with European sustainability standards, particularly through improvements in energy, infrastructure and environmental protection. The progress is stable, but not too fast, which indicates gradual, planned development and continuous implementation of reforms. **Serbia** shows the lowest initial level of the SDG index, but also the most pronounced growth during the observed period. From around 75 points in 2015, Serbia recorded growth almost every year, reaching around 79 points by 2022. After that, there is a slight decline in 2023 and 2024, which may indicate delays in the implementation of sustainable development policies or external factors that influenced the results. Despite this, Serbia has made visible progress and reduced the difference compared to Poland and Germany, which shows that the processes of reform and investment in sustainable development still pay off in the long term.

Taken as a whole, the graph (Figure 4) clearly shows three different levels of development and stages of progress towards achieving the goals of sustainable development.

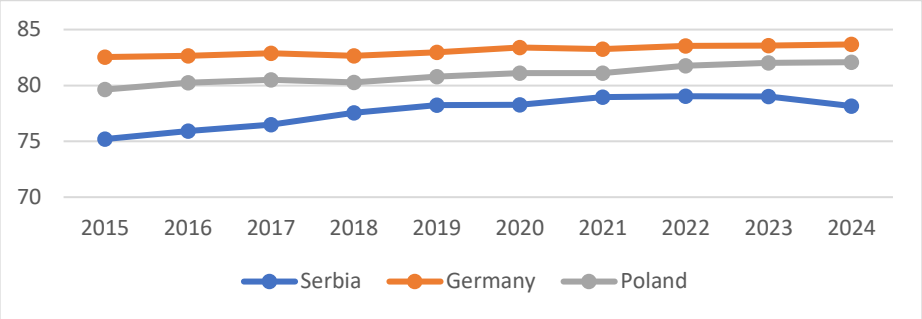


Fig. 4. Comparison of SDG Index Scores of Serbia, Germany and Poland

In the period from 2010 to 2024, Germany recorded a stable increase in the SDG Index. The index value increased from 80.6 in 2010 to 83.67 in 2024.. The increase in the SDG Index indicates progress in sustainable development at the national level, including economic, social and environmental dimensions. The relatively high and stable SDG Index values suggest that Germany maintained a high level of sustainable development throughout the observed period. Trends in foreign direct investment show unevenness from year to year. In 2010, the inflow of FDI in Germany amounted to 86.04 billion dollars. The minimum investment in Germany in the researched period was 17.36 billion dollars in 2014. which represents a drop of 79.82%. The following period was marked by the growth of investments, in which the maximum amount of 176.78 billion dollars was reached in the year of the Corona virus pandemic (2020). The latest data for 2024 show that the inflow of FDI amounted to only 47.6 billion. The decrease in the level of direct foreign investments can be explained by the action of several different factors. Despite an increase in SDG Index scores during the period under consideration, FDI flows proved to be relatively unstable with a decrease in the recent periods. It can thus be assumed that an improvement in sustainability scores on a national scale is not the sole driver of FDI. Variations in FDI inflows could also be influenced by other changes taking place in the economies and markets, such as the occurrence of financial crises and changing international trading relationships. In addition, it is necessary to further improve incentive measures and fiscal incentives in order to increase the attractiveness of the market and maintain the interest of investors in the conditions of growing global competition.

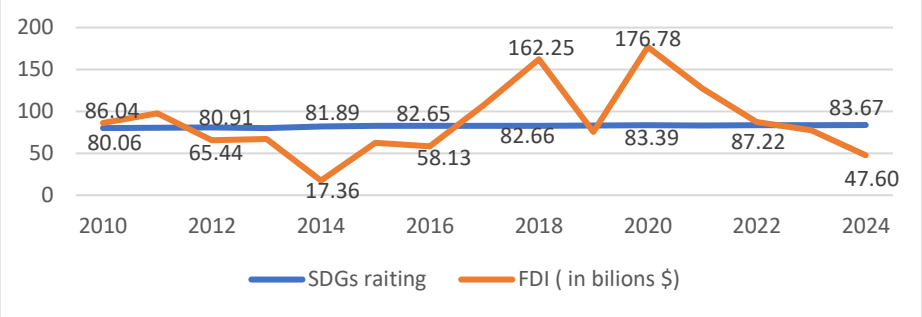


Fig. 5. SDGs rating and foreign direct investment in Germany from 2010 to 2024.

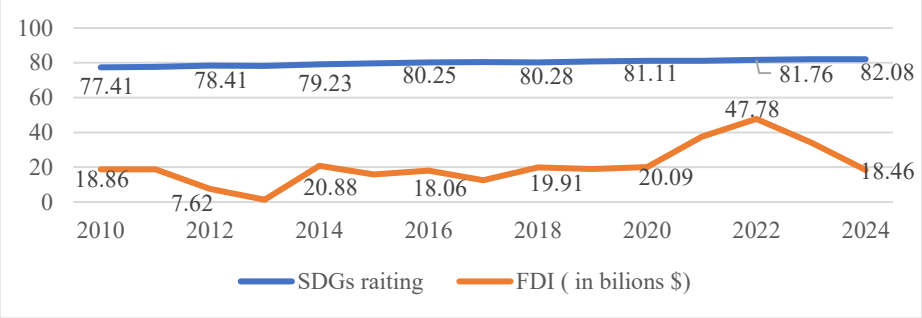


Fig. 6. SDGs rating and foreign direct investment in Poland from 2010 to 2024.

In the last 15 years, Poland has recorded an increase in SDGs Index scores from the starting 77.41 to 82.08. This increase is not only a quantitative improvement, but also a qualitative shift in the way sustainable business principles are understood and applied in the country. Progress in this area is the result of a combination of regulatory measures, economic incentives and the increasingly pronounced willingness of companies to invest in responsible business models for the long term. Investments in energy transition , development of low-carbon technologies , as well as the improvement of corporate governance and ethical standards, contributed to the growth of overall trust in the Polish business environment. When it comes to foreign direct investments, it is characterized by an unstable inflow of foreign capital. In 2010, 18.86 billion US dollars flowed into Poland, after which already in 2013 that amount was at a minimum value of only 1.34. billions. Poland is one of the few countries that did not experience a drop in investment during the pandemic, and this enabled it to achieve an inflow of 47.78 billion dollars in 2022. The latest data for 2024 show a drop of almost 62% compared to 2022. In the period from 2010 to 2024, it can be observed that the application of ESG criteria in Poland has a positive but indirect impact on foreign direct investment (FDI) flows. Although investment inflows in that period were variable and influenced by global economic trends, the growth from 77.41 to 82.08 indicates a gradual strengthening of the institutional framework and business stability in the country. The sharp decline in FDI in 2024 shows that ESG standards alone are not

enough to guarantee a stable inflow of foreign capital without the support of macroeconomic stability, fiscal policy and favorable global conditions.

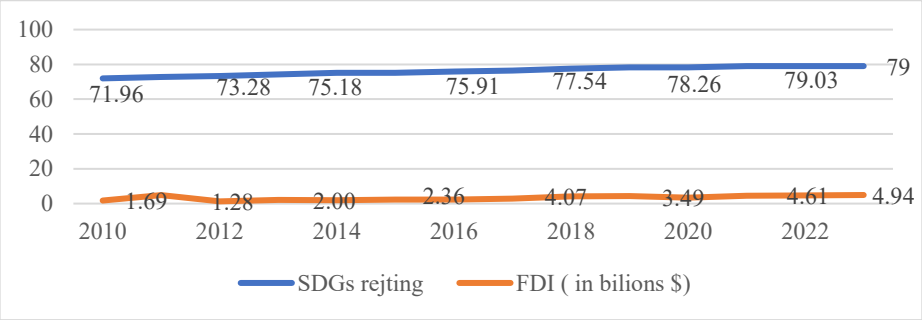


Fig. 7. SDGs rating and foreign direct investments in Serbia from 2010 to 2023.

SDGs rating values, as in Germany and Poland, have grown exponentially from 71.96 in 2010 to 79 in 2023. It can be seen that Serbia has made the most progress of all three examined countries in this category. The increase in ESG rating in Serbia was primarily due to accelerated structural reforms and adaptation to international business standards. Over the last decade, increasing attention has been paid to more efficient use of resources, greater transparency and inclusion of sustainability in development policies. In addition, the growing importance of sustainability in investment decisions encouraged the adoption of practices aligned with sustainable development principles, which coincided with the increase in SDG Index values in Serbia.. In the observed period, FDI flows in Serbia showed stable growth, with certain oscillations that are in line with global economic trends and domestic social crises. In 2010, the total amount of investments amounted to 1.69 billion dollars, while at the end of 2023 it amounted to almost 5 billion dollars, which represents a growth of 196%. The growth of FDI in Serbia is the result of a combination of a favorable geographical location, competitive business costs and an active policy of attracting investors implemented by the state through subsidies, tax breaks and infrastructure projects. However, it is important to emphasize that the structure and motives of investments show certain differences compared to more developed economies. The largest part of foreign investments in Serbia is still concentrated in the manufacturing industry, automotive production, energy and telecommunications, while a smaller part is directed towards high-tech sectors and research and development. In the case of Serbia, the visible increase of indicators indicates a gradual change in the way the country builds investment attractiveness. Although the growth of FDI in previous years was stable, the slowdown in the last period shows that sustainability must be supported by macroeconomic stability, more efficient administration and consistent application of ESG principles in practice.

5 Discussion

The results suggest that sustainability, measured through the SDG Index, may represent one of the factors associated with the attractiveness of countries for foreign direct investment. However, the findings do not support the existence of a simple or linear relationship between sustainability performance and FDI inflows. Although Germany and Poland recorded improvements in SDG Index values, both countries experienced significant fluctuations and declines in FDI inflows in certain years. This indicates that sustainability performance alone cannot explain investment dynamics. These findings are consistent with recent studies which emphasize that the relationship between sustainability and FDI depends on a broader set of economic, institutional and geopolitical factors. The results support the view that sustainability may contribute to investment attractiveness, but that macroeconomic stability, institutional quality and global market conditions remain important determinants of foreign investment decisions. The comparative analysis also highlights differences between the observed countries. Germany maintained the highest SDG Index values throughout the period, Poland recorded steady progress, while Serbia achieved the largest relative improvement. At the same time, FDI trends varied considerably across all three countries, suggesting that improvements in sustainability indicators do not automatically translate into higher investment inflows. From a policy perspective, the findings indicate that sustainability-oriented reforms should be accompanied by broader measures aimed at strengthening institutional capacity, regulatory predictability and economic stability. For countries such as Serbia, continued alignment with European sustainability standards may contribute to improving the investment environment, but sustainable development indicators should be viewed as one component of a wider set of factors influencing FDI.

6 Conclusion

This paper examined the relationship between sustainability performance and foreign direct investment inflows through a comparative analysis of Germany, Poland, and Serbia. Using SDG Index data as an indicator of national sustainability performance and FDI inflow data for the observed period, the study explored whether improvements in sustainability-related indicators were accompanied by changes in investment attractiveness. The results indicate that countries with higher levels of sustainability performance generally demonstrate stronger investment attractiveness; however, the relationship is not uniform across all countries and time periods. While Germany and Poland achieved relatively high and stable SDG Index values, both countries experienced significant fluctuations in FDI inflows. Serbia recorded the largest improvement in sustainability indicators and a substantial increase in FDI inflows, although investment dynamics were also influenced by broader economic and institutional factors. These findings suggest that sustainability performance may represent one of the factors considered by investors, but it should not be viewed as the sole determinant of foreign direct investment. The analysis further indicates that the

relationship between sustainability and FDI is more complex than a simple linear association. Changes in investment flows may also reflect macroeconomic conditions, geopolitical developments, regulatory changes, market size, labor costs, and institutional quality. Therefore, improvements in sustainability indicators alone do not automatically guarantee higher levels of foreign investment. This study contributes to the existing literature by providing a comparative perspective on sustainability performance and FDI dynamics in countries with different levels of economic and institutional development. The findings highlight the importance of considering sustainability indicators together with broader economic and institutional conditions when evaluating national investment attractiveness. Several limitations should be acknowledged. The research is based primarily on descriptive analysis and therefore does not allow for the establishment of causal relationships between sustainability performance and FDI inflows. In addition, the study relies on SDG Index data, which measure progress toward sustainable development goals and should not be interpreted as direct ESG ratings. Future research should employ econometric approaches, such as panel data models or fixed-effects analysis, in order to control for macroeconomic variables and examine the individual contribution of sustainability-related factors to foreign investment decisions more precisely. Overall, the findings suggest that sustainability performance is becoming an increasingly relevant component of the broader investment environment. However, its effects on foreign direct investment should be interpreted within the context of wider economic, institutional, and regulatory conditions.

References

1. OECD: OECD Business and Finance Outlook 2020: Sustainable and Resilient Finance. OECD Publishing, Paris (2020)
2. WCED: Our Common Future. World Commission on Environment and Development (1987)
3. Global Sustainable Investment Alliance: Global Sustainable Investment Review 2022 (2022)
4. UNCTAD: World Investment Report 2024: Investment Facilitation and Digital Government. United Nations Conference on Trade and Development, Geneva (2024)
5. OECD: FDI Qualities Indicators: Measuring the Sustainable Development Impacts of Investment. OECD Publishing, Paris (2019)
6. Friede, G., Busch, T., Bassen, A.: ESG and financial performance: Aggregated evidence from more than 2000 empirical studies. Journal of Sustainable Finance & Investment 5(4), 210–233 (2015). https://doi.org/10.1080/20430795.2015.1118917
7. Khan, M., Serafeim, G., Yoon, A.: Corporate sustainability: First evidence on materiality. The Accounting Review 91(6), 1697–1724 (2016). https://doi.org/10.2308/accr-51383
8. Boffo, R., Patalano, R.: ESG Investing: Practices, Progress and Challenges. OECD Working Papers on Finance, Insurance and Private Pensions, No. 44 (2020)
9. Gibson Brandon, R., Krueger, P., Schmidt, P.S.: ESG rating disagreement and stock returns. Financial Analysts Journal 77(4), 104–127 (2021). https://doi.org/10.1080/0015198X.2021.1963186

10. Pedersen, L.H., Fitzgibbons, S., Pomorski, L.: Responsible investing: The ESG-efficient frontier. Journal of Financial Economics 142(2), 572–597 (2021). https://doi.org/10.1016/j.jfineco.2020.11.001

11. Ioannou, I., Serafeim, G.: What drives corporate social performance? International evidence. Journal of International Business Studies 43(9), 834–864 (2012). https://doi.org/10.1057/jibs.2012.26

12. Kotsantonis, S., Pinney, C., Serafeim, G.: ESG integration in investment management. Journal of Applied Corporate Finance 28(2), 10–16 (2016). https://doi.org/10.1111/jacf.12169

13. Eccles, R.G., Klimenko, S.: The investor revolution. Harvard Business Review 97(3), 106–116 (2019)

14. Caetano, R.V., Marques, A.C.: Nonlinear relationships between foreign direct investment decisions and environmental degradation in high- and middle-income countries. Journal of Industrial and Business Economics 51(1), 135–188 (2024). https://doi.org/10.1007/s40812-023-00282-0

15. Bannour, S., Abdelkawy, N.A.: Sovereign ESG and foreign direct investment in the GCC: The amplifying role of trade openness in economic diversification. Sustainability 16(21), 9326 (2024). https://doi.org/10.3390/su16219326

16. Sidana, N., Gogoi, P.: A macroeconomic analysis of foreign direct investment and the environmental, social, governance, and market potential pillars in BIMSTEC nations. Model Assisted Statistics and Applications 20(2), 146–169 (2025). https://doi.org/10.1177/15741699251334332

17. Berg, F., Kölbel, J.F., Rigobon, R.: Aggregate confusion: The divergence of ESG ratings. Review of Finance 26(6), 1315–1344 (2022). https://doi.org/10.1093/rof/rfac033

18. El Ghoul, S., Guedhami, O., Kim, Y., Park, K.: Corporate environmental responsibility and the cost of capital. Journal of Banking & Finance 86, 130–148 (2018). https://doi.org/10.1016/j.jbankfin.2017.09.012

19. Pankratz, S., Bauer, R., Derwall, J.: Climate change, firm performance, and investor surprises. Journal of Finance 76(4), 1719–1774 (2021)

20. UNIDO: Global Industrial Development Report 2022: Transforming Industries for a Sustainable Future. United Nations Industrial Development Organization, Vienna (2022)

Original Research Article *https://doi.org/10.5281/zenodo.23474684*

Impact of extreme temperatures on cardiovascular morbidity and mortality

Marjan Senegačnik[1][0000-0002-0013-2762]
and Davorin Žnidarič[2,3][0009-0000-8522-5767]

[1] University Maribor- Faculty of Organisation science Kranj
[2] University of Ljubljana- Faculty of Law; [3.]Institute for Sustainability Development KUM

Abstract. Global warming and climate change are very topical issue. Among the consequences of climate change are also negative impacts on health. Climate change is associated with more frequent and intense heat waves. On the other hand, despite global warming also periods of very cold weather in some regions are expected in the future. Both, heat waves as well as periods of very cold weather pose significantly increased health risk. People with chronic diseases, such as cardiovascular disease, are particularly at risk. The paper addresses the problem of the impact of extreme temperature periods (heat waves and very cold periods) on cardiovascular morbidity and mortality and is based on the review of literature. Risk factors for cardiovascular disease are presented, with an emphasis on environmental risk factors which also include periods of extreme temperatures. In recent decades, scientists have observed more frequent and intense heat waves. On the other hand, despite global warming, periods of very cold weather remain quite common in some regions. An overview of the findings of some of the most comprehensive studies on the impact of extreme temperatures on cardiovascular morbidity and mortality is presented. Results on the impact of heat waves on mortality in Slovenia are also briefly presented. Finally, a framework of measures is briefly described which will help to improve the preparedness of healthcare system, increase its resilience and mitigate the harmful effects of extreme temperatures on health.

Keywords: Heat waves, Cold spells, Cardiovascular diseases, Improved resilience of the healthcare system to climate change, Temperature impact on health.

1 Introduction

Greenhouse effect and global warming have been one of the most outstanding topics for many years. One of the important consequences of global warming and the resulting temperature changes is the potential adverse impact on health. Global warming and temperature changes in general can cause the emergence of many diseases or worsening of the existing health conditions. This is case for both communicable and chronic non-communicable diseases. This paper is limited on the influence of outdoor air temperatures on cardiovascular morbidity and mortality. Cardiovascular diseases are the

leading cause of mortality in the developed countries. Exposure to extreme temperature is also one of the reasons for cardiovascular complications which can lead to the necessity of hospitalization or can even cause death. Scientists are concerned that rising temperatures will also increase the frequency of heat waves. Periods of extreme high temperatures lasting several days pose considerable stress to the body. Although this fact does not pose a significant risk for young or healthy people, the situation is significantly different for sensitive people (elderly, people with cardiovascular or lung diseases, diabetes etc). On the other hand, low temperatures also pose risks, particularly for people suffering from cardiovascular diseases. Many experts believe, that despite expected global warming, periods of extreme cold (cold waves or cold spells) cannot be ruled out in future.

Increasing resilience to periods of extreme temperatures and ensuring a reduction in the associated life-threatening health complications presents a complex challenge, not only from a medical perspective but also from organizational, administrative and urban planning perspectives. According to scientists' predictions hospitals, emergency medical services, public health institutions, municipalities and urban infrastructure can be expected to face increasing health risks associated with climate change in the future. The purpose of this paper is to analyse the impact of extreme temperatures on morbidity and mortality from cardiovascular diseases and to develop a conceptual framework that links health issues related to climate change with the preparedness of health systems, organizational resilience and adaption strategies.

The central research question of the paper is:

How can healthcare systems, public institutions, and municipal administrative structures improve their municipal resilience and preparedness for cardiovascular risks associated with extreme temperatures?

2 Literature review methodology

The paper is based on a qualitative literature review, which focused on the links between extreme temperatures, cardiovascular morbidity and mortality, organizational resilience, health system preparedness, climate change adaptation and urban governance.

Scientific literature was obtained from the following databases:
- PubMed
- Web of Science
- ResearchGate.

The following key words or searching strings were used:
- "heat waves AND cardiovascular mortality",
- "cold spells AND cardiovascular morbidity",
- "climate change AND cardiovascular risks",
- "organizational resilience AND climate adaptation",

- “healthcare preparedness AND extreme temperatures”,
- “urban heat islands AND public health”,
- “extreme temperatures AND mortality”,
- “heat-health action plans”.

The literature review focused on studies from the period between 2000 and 2026. The main focus was on studies published after 2015, and older publications were included only if they were conceptually relevant.

Inclusion criteria were:
- Peer-reviewed scientific articles,
- International institutional reports,
- Studies on the impact of extreme temperatures on cardiovascular health
- Studies dealing with the preparedness of health systems
- Studies on climate change adaptation and resilient urban systems.

Exclusion criteria were:
- Publications not related to the impact of extreme temperatures
- Publications not related to the cardiovascular disease
- Publications not related to climate change adaption or organizational preparedness.

The selection process included:
- Initial identification based on the keywords
- Abstract review
- Thematic configuration
- Number of citations
- Synthesis of findings.

3 Cardiovascular disease

Cardiovascular diseases are the leading cause of mortality, at least in developed countries. In Europe, approximately 6 million new cases of cardiovascular disease and more than 1,7 million of deaths from cardiovascular disease are recorded each year (37 % of all deaths) [1-3].

3.1 Risk factors for cardiovascular disease

There are many factors that contribute to the development of cardiovascular disease, but they can be divided into two main groups: modifiable and non-modifiable. Non-modifiable factors are factors that cannot be influenced: age, sex, race, ethnicity and genetics. Modifiable factors are those that can, in principle, be influenced: psychosocial, behavioural, environmental and partly also clinical factors. The individual categories of risk factors are partially interdependent, as shown in Fig.1 [1].

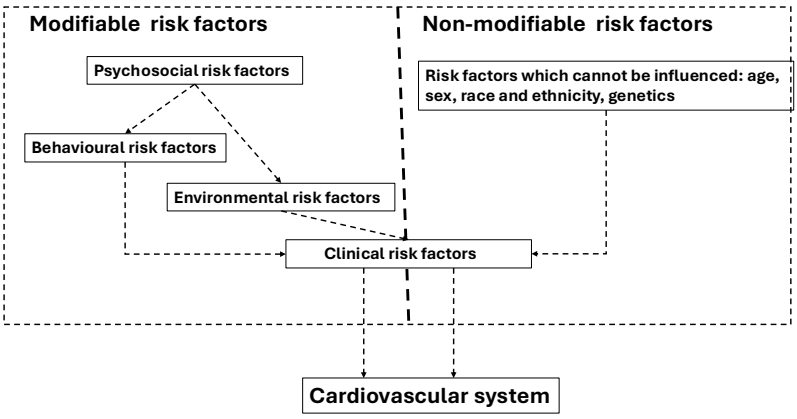


Fig. 1. Schematic representation of modifiable and non-modifiable risk factors for cardiovascular disease and their mutual relations [1].

Psychosocial risk factors depend on an individual's socioeconomic status, education, access to health care and so on. To a certain extent, psychosocial risk factors can influence both behavioural and environmental risk factors [1].

Behavioural risk factors include smoking, excessive consumption of alcohol, unhealthy diet, lack of physical activity etc. Environmental risk factors include the influence of both the external and internal environment and may also depend on the individual's occupation [1].

Clinical risk factors include parameters such as are high blood pressure, elevated blood lipids and glucose concentrations, obesity etc. These factors largely depend on the individual's behaviour and environment, both of which can be influenced by everyone. On the other hand, clinical risk factors also depend on genetics, age etc., which the individual cannot change [1].

3.2 Environmental risk factors

Environmental risk factors include both internal and external environment. The internal environment is largely specific and depends on the layout of living or working space. Internal environment is not the topic of the present article.

The impact of the external environment depends on various factors, among which air quality and temperature are particularly important. The consequences of polluted air on health are well known. According to the estimate of Global Burden Disease, in 2019, polluted air was responsible for 6,7 million deaths worldwide (95 % confidence interval) [4].

The main air pollutants are carbon monoxide (CO), hydrocarbons (HC) (particularly benzene), nitrogen oxides (NO_x), tropospheric ozone (O_3) and particulate matter (PM). PM particles are especially problematic. For decades, the harmful effects particulate

matter on the respiratory system have been highlighted; however, numerous studies conducted since 1990 indicate that these particles pose an even greater risk for cardiovascular disease.

Another environmental issue that is attracting even much more attention than polluted air is greenhouse effect and global warming. The present paper is concerned with influence of temperature and temperature changes on cardiovascular morbidity and mortality. Temperature extremes, both very high and very low temperatures, present a great stress on human organism. Such stress is particularly problematic for sensitive people with enhanced risk factors and can lead to severe health complications and even life-threating situations.

4 Climate change and its impact on extreme temperatures

According to scientists, global warming and climate change are linked to more frequent periods of extreme temperatures. Although periods of exceptionally high temperatures (heat waves) are primarily expected, periods of very cold weather cannot be ruled out. Both types of temperature extremes pose a health risk.

4.1 Heat waves

The scientific literature contains numerous warnings that heat waves will become more frequent and intense in future due to global warming. This could lead to an increase in cardiovascular complications caused by heat stress.

Analysis of heat waves in the United States have shown that in the period from 1960 to 2020 there systematically increase the frequency of heat waves (average number of heat waves per year), duration (the average number of days a heat wave lasts), season of heat waves (duration of a period in which heat waves can occur) and intensity (average temperature above the local threshold during heat waves). For example, the frequency of heat waves in the period 1960-69 was 2,17 per year while in the period 2010-19 it was 5,96 per year. Similarly, the average duration of heat waves has also increased (3,00 days in 1960s and 4,05 days in 2010s), as well as the season of heat waves (23,75 days in 1960s and 68,52 days in 2010s) and the intensity of heat waves (average temperature above the local threshold during heat waves) – 1,95 in 1960s and 2,36 in 2010s. More detailed presentation about heat waves characteristics in the United States in the period 1960-2020 is shown in Fig. 2 [5].

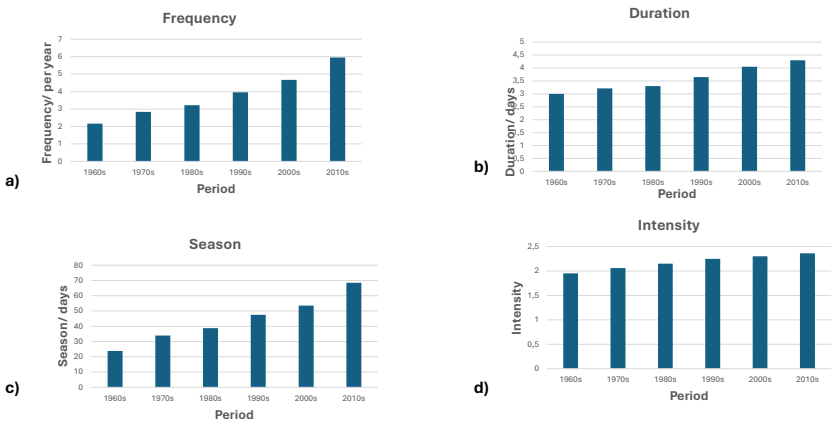


Fig. 2. Trends of heat waves in the United States for period between 1960 and 2020: a) Frequency of heat waves per year, b) Average duration of heat wave/ days, c) Season of occurring heat waves/days, d) Intensity of heat waves [5].

Rousi et al. [6] emphasize that the problem of heat waves is particularly pronounced in Europe, where the upward trends over the past four decades have been significantly stronger than in other mid-latitude regions of the Northern hemisphere. This is particularly evident in the western part of the continent.

Between 2000 and 2020, the mortality rate among persons with age 65 or older due to high temperatures nearly doubled in Europe. In 2018, more than 300.000 people died as a result. People suffering from cardiovascular, respiratory, kidney or endocrine diseases, as well as metabolic disorders, are particularly at risk. In addition to climate change, two other factors are urbanization and aging population. The effects of high temperatures can be direct (heatstroke, dehydration) or indirect (deterioration of cardiovascular, respiratory or renal health, electrolyte imbalances etc.). Scientists warn that a scenario of rapid global warming could result in an additional 100.000 deaths per year. However, if the temperature rise could be limited to 2 ºC, this could significantly reduce the increase in mortality [7].

4.2 Periods of extreme cold

Another case of extreme temperatures, which also place a significant strain on the body, are periods of very cold weather (cold spells). Some researchers claim that the frequency of cold spells will decrease, which could be consistent with global warming. However, some other scientists argue that while the frequency of cold spells is indeed decreasing in certain regions, they are occurring even more frequently in others [8].

Tomczyk et al. [8] studied the frequency of cold spells in Germany and Poland. They observed a slight decrease in frequency of cold spells, consistent with global warming.

However, they state that in the most regions the changes are not statistically significant. Some other researchers also report a lack of statistically significant changes.

Jarzyna and Krzyżewska [9] studied the frequency of cold spells in the period between 1981 and 2018. They observed that cold spells are rare or almost absent in Western Europe, while they are quite frequent and intensive in Eastern Europe. They have not observed statistically significant differences. Cold spells in Central and Eastern Europe are expected to continue to pose significant health risks. Although the danger of heat waves is often emphasized in the literature, they state that cold spells pose an even greater health risk than heat waves [9].

There are also some other researchers who argue that climate change is linked to more frequent extreme weather events and that this could lead to more frequent and intense cold spells in some regions, in addition to more frequent heat waves [10].

5 Influence of extreme temperatures on cardiovascular morbidity and mortality

One of the most comprehensive studies concerned with the impact of extreme temperatures on cardiovascular mortality presents the study of Alahmad et al. [11] which is based on the extensive international data base. Many countries from Africa, Asia, North, Central and South America and Europe are included. These countries are located in different climatic zones and differ in socio-economic and demographic characteristics. The study covers 567 cities from 27 countries. The researchers collected data of 32,15 million cardiovascular deaths, of which 11,75 million were due to ischemic heart disease, 9,35 million were due to stroke, 3,67 million were due to heart failure and 0,67 million were due to arrhythmias [11].

Because the countries and cities included in the study are located in different climate zones, temperature differences were also significant. Therefore, it is more convenient to express the temperature in percentiles (the interval between the highest and lowest temperature is divided into 100 parts) rather than in temperature units (°C). The range of extremely low temperatures is between the 0^{th} and $2{,}5^{th}$ percentiles, and the range of extremely high temperatures is between the $97{,}5^{th}$ and 100^{th} percentiles. The relative risk (RR) dependence on temperature percentile is nonlinear for both total cardiovascular mortality and mortality from individual diseases. The RR value is high at very low temperature percentiles and then decreases with increasing temperature. The lowest value of RR is reached at the point of minimal mortality temperature (MMT). The position of MMT varies considerably across countries, ranging from the 78^{th} percentile in Taiwan to the 93^{rd} percentile in Estonia. At temperature percentiles higher than MMT, the RR dependence curve begins to increase steeply. A simplified representation of the dependence of the RR value for total cardiovascular mortality on the temperature percentile is shown in Fig. 3 [11].

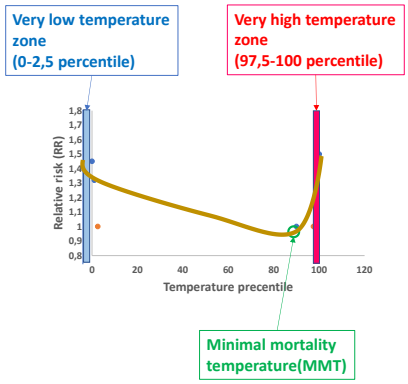


Fig. 3. Dependence of relative risk (RR) for total cardiovascular mortality on temperature percentile [11].

The relative mortality risk (RR) at a given temperature is expressed according to the minimal mortality rate (MMT point in Fig.3). The temperature dependence of the relative risk (RR) for mortality from ischemic heart disease, stroke, and also for heart failure is very similar to the temperature dependence for total cardiovascular mortality, while in the case of arrythmias there is more uncertainty and no clear conclusions can be drawn. For the illustration, in Table 1 are shown RR values for various cardiovascular diseases for extremely low temperatures (1stpercentile) and extremely high temperatures (99th percentile). For every 1000 cardiovascular deaths there were found 2,2 excess deaths due to very high temperatures (over 97,5th percentile) and 9,1excess deaths due to very low temperatures (under 2,5th percentile) [11].

Table 1. Relative risk (RR) for various cardiovascular diseases due to very low temperatures (under 1st percentile) and due to very high temperatures (over 99th percentile) (95 % CI). RR values are expressed according to MMT value [11].

Cardiovascular disease	Very low temperatures (1st percentile)	Very high temperatures (99th percentile)
Ischemic heart disease	1,33 (1,26 – 1,41)	1,07 (1,04 - 1,10)
Stroke	1,32 (1,26 – 1,38)	1,10 (1,06 – 1,15)
Heart failure	1,37 (1,28 – 1,47)	1,12
Arrhythmias	1,19 (1,07 – 1,33)	1,05

Liu et al. [12] conducted a comprehensive review of studies examining the impact of high temperatures and heat waves on cardiovascular morbidity and mortality. They reviewed 282 papers from the period 1990-2022. Despite large differences in environmental conditions and population dynamics, they confirmed a positive association between morbidity and mortality from cardiovascular diseases and increasing temperature. A 1 ºC increase in temperature was associated with 2,1 % increased risk of mortality from cardiovascular diseases (95 % CI). The increase was most pronounced in the case of stroke and ischemic heart disease, but was also noticeable in the case of arrythmias and cardiac arrest. An increased risk of morbidity and mortality of 11,7 % was observed for heat waves – 6,7 % for low-intensity waves and 18,9 % for high-intensity waves [12].

5.1 Characteristics of heat waves impact on morbidity and mortality in Slovenia

For illustration purposes, the situation regarding cardiovascular morbidity and mortalty due to heat waves in Slovenia for the period 2000-2020 is briefly presented. In this context, two studies should be highlighted: study of Šelb-Šemrl and Tomšič [13] and study of Perčič et al. [14].

A somewhat older study of Šelb-Šemrl and Tomšič [13] analyses the impact of the intense heat wave of August 2003 on morbidity and mortality in Slovenia. They found that the number of deaths in August 2003 was higher than the average number for August in the period 1997-2006. During the heat wave period between August 2 and 15, mortality increased by 13 % compared to the same period between 1997 and 2002. Mortality among people over 74 years of age has increased, particularly due to influenza, pneumonia, diabetes and chronic lower respiratory diseases [13].

Perčič et al. [14] studied the impact of heat waves on mortality in Slovenia in the decade 2006-2015, and especially in 2015. An increase in overall population mortality was observed during the 2015 heat waves, but this increase was not statistically significant. However, the increase in overall population mortality from cardiovascular disease during the 2015 heat waves was statistically significant. In the population over 75 years of age, both total mortality and cardiovascular disease mortality were statistically significant.

Similarly, over the entire decade 2006-2015, total population mortality during heat waves increased, but this increase was not statistically significant. In the population over 75 years of age, the increase in mortality was statistically significant for both overall mortality and cardiovascular mortality [14].

According to the data from the National Institute of Public Health (NIJZ), no statistically significant increase in mortality during the heat waves was observed in Slovenia in the period 2015-2020. This is also thanks to timely warnings and preventive health measures [15].

6 Proposals of solutions

As is well known, much effort is being invested in reducing greenhouse gas emissions, which is important for limiting temperature increases. This may also lead to less frequent and intense heat waves.
However, there are some other measures that can help mitigate the effects of extreme temperatures. This includes timely meteorological information, medical advice, accommodation arrangements and appropriate urban planning which can help prevent the formation of heat islands. Such measures can be systematically united in the so-called heat-health action plan (HHAP). In the year 2020 16 member countries of European Environmental Agency have HHAP on the national level and some other countries have similar plans prepared on regional or local level [7,16,17].

Martinez et al. [16] and Vanderplanken et al. [17] emphasize that properly developed action plans significantly improve the institutional preparedness and reduce health risks associated with climate change.

The framework of a system for adaptation to changing environmental conditions is a comprehensive interdisciplinary model of interconnected environmental, health, organizational and management factors.

6.1 Administrative and policy measures

It is important to establish a system at the national and/or regional level which should include:

- climate change adaption policies
- public warning systems
- institutional coordination
- risk management systems and
- organizational resilience planning.

6.2 Measures at the individual level

Preventing cardiovascular complication is very important. For individuals in groups with an increased cardiovascular risk, the likelihood of complications can be reduced if appropriate measures are taken in a timely manner. These measures include:

- providing information on appropriate preventive measures
- timely warnings of upcoming heat waves or cold spells
- adequate hydration and
- monitoring of health status.

6.3 Healthcare system preparedness

The healthcare system must incorporate appropriate organizational measures:

- Action plans for heat waves or cold spells
- Emergency protocols
- Hospital capacity planning
- Preparedness of healthcare personnel and
- Coordination among healthcare facilities.

6.4 Urban planning measures

These measures are important for long-term climate change preparedness. They include:

- Green infrastructure
- Urban cooling systems
- Mitigation of effects of urban heat islands
- Resilient infrastructure and
- Sustainable urban planning.

7 Conclusion

Due to global warming and climate change, heat waves have become more frequent and intense in recent decades. Extreme temperatures pose a great strain on the human body and increase the risk of developing various diseases or deteriorating health. This poses practically no problem for young and healthy people. However, this can pose a serious problem for elderly and people with chronic illnesses and even lead to life threatening situations.

In addition to systematic efforts to reduce the human contribution to global warming, preventive measures also play an important role and can help effectively mitigate the consequences of extreme temperatures. These measures range from meteorological notifications and medical advice to appropriate urban planning.

Many articles can be found in the literature that deal with the health effects of high temperatures, and especially heat waves. On the other hand, it is interesting that there are far fewer articles devoted to the health effects of low temperatures and cold waves. Scientists warn that despite global warming, periods of very low temperatures in some regions will not be uncommon in future. Extremely low temperatures also pose a significant health risk.

The paper presents also a framework of some proposed measures to mitigate the harmful effects of extreme temperature on health. These measures include administrative and policy measures, measures which can be performed on individual level, healthcare system preparedness and urban planning measures. They range from timely and accurate meteorological forecasts, appropriate education of citizens for protective behaviour, effective health services to appropriate architectural and urban planning.

Acknowledgments: The first version of the paper was prepared without the use of any AI tool. During the revision of the paper AI was used to search for certain information. However, the text of the paper was not generated by AI, the DeepL tool was only used to check the grammatical correctness of the text.

Disclosure of Interests. Authors have no competing interests.

References

1. EEA, Beating cardiovascular disease – the role of Europe's environment. EEA – European Environment Agency: Web report. An official website of the European Union. online: Homepage, https://www.eea.europa.eu/publications/beating-cardiovascular-disease, last accessed 2026/04/14
2. Timmis, A., Vardas, P., Townsend, N., …, Achenbach, S.: European Society of Cardiology: cardiovascular disease statistics. Eur Heart J., **43**(8), 716-799 (2022). doi: 10.1093/eurheartj/ehab892.
3. WHO. The top 10 causes of death. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death, last accessed 2026/06/02
4. GBD 2019 Risk Factors Collaborators: Global burden of 87 risk factors in 204 countries and territories, 1990-2019: A systematic analysis for the Global Burden of Disease Study 2019. Lancet **396** (10258), 1135-1159 (2020). doi: 10.1016/S0140-6736(20)30752-2
5. EPA Climate Change Indicators in the United States: Heat Wave characteristics. Homepage: www.epa.gov/climate-indicators heat wave characteristics, last accessed 2026/04/14
6. Rousi, E., Kornhuber, K., Beobide Arsuaga, G., Luo, F. in Coumou, D.: Accelerated western European heatwave trends linked to more-persistent double jets over Eurasia. Nat Commun, **13**, 3851 (2022). https://doi.org/10.1038/s41467-022-31432-y
7. European Climate and Health Observatory. Climate ADAPT – Sharing Adaptation Knowledge for a Climate-resilient Europe. Homepage https://climate-adapt.eea.europa.eu/en/observatory/evidence/health-effects/heat-and-health, last accessed 2026/04/14
8. Tomczyk, A.M., Bednorz, E. and Sulikowska, A.: Cold spells in Poland and Germany and their circulation conditions. Int J Climatol.39 (2019) 40032-4014. DOI: 10.1002/joc.6054
9. Jarzyna, K and Krzyżewska, A.: Cold spell variability in Europe in relation to the degree of climate continentalism in 1981-2018 period. Weather, **76**(4), 122-128 (2021) DOI: 10.1002/wea.3937
10. Monge-Sanz, B.: Global warming may be behind an increase in the frequency and intensity of cold spells. The Conversation. Homepage https://theconversation.com/global-warming-may-be-behind-an-increase-in-the-frequency-and-intensity-of-cold-spells-223153, last accessed 2026/04/14
11. Alahmad, B., Kraishah, H., Royé, D. …, Koutrakis, P.: Associations Between Extreme Temperatures and Cardiovascular Cause-Specific Mortality: Results From 27 Countries. Circulation, **147**, 35-46 (2023). DOI: 10.1161/CIRCULATIONAHA.122.061832
12. Liu, J., Varghese, B.M., Hansen, A., …; Bi, P.: Heat exposure and cardiovascular health outcomes: a systematic review and meta-analysis. Lancet Planet Health, 6, e484-e495 (2022). doi: 10.1016/S2542-5196(22)00117-6.

13. Šelb-Šemerl, J. and Tomšič, S. (2008).: Vpliv vročinskih valov na umrljivost – Vročinski val avgusta 2003 v Sloveniji. Inštitut za varovanje zdravja Republike Slovenije. (In Slovene) Homepage: https://ec.europa.eu/health/archive/ph_information/dissemination/unepected/docs/vpliv.pdf, last accessed 2026/04/15
14. Perčič, S.; Kukec, A., Cegnar, T., Hojs, A. (2017).: Vplivi vročinskih valov na umrljivost v Sloveniji. Vetrnica – glasilo Slovenskega meteorološkega društva. (In Slovene). Homepage https://meteo.arso.gov.si/uploads/probase/www/fproduct/media/sl/podcast/ep015/percic.pdf, last accessed 2026/04/15
15. NIJZ (2022). Svetovni dan zdravja 2022 – Podnebne spremembe: Vročinski valovi in zdravje ljudi. Nacionalni inštitut za javno zdravje, 7.4. 2022. Prepared by S. Perčič, A. Hojs, M. Pohar, L. Lampič and K. Bitenc. (World Health Day 2022 - Climate Change: Heat waves and Public Health - In Slovene) Homepage https://nijz.si/moje-okolje/podnebne-spremembe/svetovni-dan-zdravja-2022-podnebne-spremembe-vrocinski-valovi-in-zdravje-ljudi/, last accessed 2026/04/15
16. Martinez, G. S., et al.: Heat-health action planning in the WHO European Region: Status and policy implications, Environmental Research 214, 113709 (2022). DOI: 10.1016/j.envres.2022.113709
17. Vanderplanken, K., et al.: Governing heatwaves in Europe: comparing health policy and practices to better understand roles, responsibilities and collaboration', Health Research Policy and Systems 19(1), 20, (2021). DOI: 10.1186/s12961-020-00645-2

Original Research Article *https://doi.org/10.5281/zenodo.23474727*

Assessing Sustainability in Surface Water Treatment: A Circular Economy Perspective with Data-Driven Decision Support

Nevena Purić[1] [0009-0001-9797-2791], Jelena Andreja Radaković[2] [0000-0001-8771-8215], Vladimir Mandić[3] [0009-0009-5112-8382], Maja Đolić[4] [0000-0002-4125-6497]

[1] University of Belgrade – Institute for Multidisciplinary Research, National Institute of the Republic of Serbia, Kneza Višeslava 1, 11000 Belgrade, Serbia
[2] University of Belgrade – Faculty of Organizational Sciences, Jove Ilića 154, 11000 Belgrade, Serbia
[3] Public Utility Company Waterworks and Sewerage, Obrenovački put bb, 11000 Belgrade, Serbia
[4] University of Belgrade – Faculty of Technology and Metallurgy, Karnegijeva 4, 11000, Belgrade, Serbia
nevena.puric@imsi.bg.ac.rs

Abstract. Rising pressure on freshwater resources, the decreasing quality of source water, and the need to achieve sustainability have led to innovation in drinking water treatment plants. In this respect, machine learning and artificial intelligence have become useful tools to optimize chemical dosing and support operational decision-making, especially for clarification processes, where dosing decisions directly affect further treatment. Implementing artificial intelligence in this sector can contribute significantly towards reaching the goals of both environmental sustainability and economic viability. This paper addresses the gap between machine learning approaches for chemical dosing prediction in drinking water treatment and possibility of deployment, examined through the lens of resource efficiency, waste minimization, and sustainable operational practice. The research findings from literature are illustrated through a preliminary case study at drinking water treatment plant Makiš, operated by the Public Utility Company Belgrade Waterworks and Sewerage. Random Forest regression model was developed using operational laboratory data collected over the past four years, without additional infrastructure investment. An analysis of operational consistency showed that the random forest model reproduces the dosing decisions of experienced operators 63.4% of the time for coagulants and 71.1% for flocculants, with dangerous deviations occurring in fewer than 2.1% of predictions. The study's findings indicate preliminary technical feasibility of machine learning-based models for chemical dosing prediction, through deployment on existing operational data, without additional infrastructure investment. However, implementation remains constrained by sociocultural and regulatory conditions that require deliberate organizational investment.

Keywords: drinking water treatment, sustainable approach, machine learning, random forest, chemical dosing, water utility management.

1 Introduction

Freshwater availability has become one of the major concerns of recent decades, intensified by population growth, climate change, pollution, and rising demands from future generations. Access to safe drinking water is a fundamental prerequisite for sustainability and for achieving sustainable development goals (SDGs), since it directly impacts human health, food production, ecosystems, economic development, and climate resilience. At the same time, clean drinking water remains inaccessible to billions of people worldwide [1–4]. Freshwater resources are increasingly stressed due to overexploitation, the quality of the source water deteriorating, and climate change which is causing hydrological variability [1, 2]. These pressures in turn affect the possibility of achieving SDGs, especially SDG 6 (Clean water and sanitation) as well as SDGs related to health, food security, energy, cities, ecosystems, and poverty reduction [4].

With water quality depending greatly on human activities, physical, chemical, and biological properties of source water reflect higher concentrations of nutrients, organic load and, pathogens, which calls for well-implemented monitoring and water treatment systems [5–7]. Drinking water treatment plants (DWTPs) must therefire maintain high drinking water quality while using less energy, fewer chemicals, and producing fewer harmful byproducts. Therefore, the main goals of sustainable water management are efficiency and resilience [3, 8]. The failure of such DWTPs poses a concern and health risk, positioning it as a societal question as well as technical. Source protection, optimized treatment, pipe integrity and well managed distribution systems should be improved to provide for urban public health [9–11]. In order for drinking water treatment (DWT) to be sustainable it needs to meet the quality standards of drinking water and reduce the energy intensity, as well as optimizing chemical dosing, minimizing sludge, and other treatment residuals, while lowering the broader environmental footprint of the treatment processes [12, 13]. In accordance with this, the enhancement of DWT performance is increasingly acknowledged as both a technical imperative and a critical step toward enduring water security and environmental sustainability. This is driving the focus of DWTPs toward becoming more closely aligned with circular economy (CE) principles, which are designed to enhance operational sustainability, limit waste generation, and reduce resource consumption. As defined by the Ellen MacArthur Foundation, CE is based on the following three principles: the elimination of waste and pollution, the circulation of products and materials, and regenerating nature [14].

In this context, CE within the water sector seeks to move beyond conventional linear models of abstraction, treatment, utilization, and disposal, by fostering resource efficiency, waste reduction, and more sustainable system management [15–17]. This approach prioritizes the reduction of sludge and other residuals, the minimization of water, energy, and treatment chemical losses within water systems, and the enhancement of treatment infrastructure's operational efficiency through improved monitoring and control [3, 8]. Consequently, there is rising interest in new methodologies, tools and analyses, that incorporate sustainability and CE principles into water treatment

systems, with a focus on resource efficiency, waste minimization, and data-driven optimization of the treatment processes [18, 19].

A DWTP encompasses multiple units within the process line, with clarification as the initial and most fundamental stage. This design creates a system in which chemical dosing decisions directly affect the performance of all downstream units [20]. Consisting of coagulation, flocculation and, sedimentation, clarification presents the greatest potential for optimizing chemical dosing, minimizing waste production and energy consumption [21]. Primary chemical inputs subjected to operators' adjustment are coagulant aluminum sulfate and flocculant polyacrylamide. Coagulant and flocculant dosage relies on operators' empirical knowledge and experience in alignment with the jar test. Given the continuous fluctuation of water quality, using jar test can be seen as unsuitable for real time operation [21, 22]. The operators' dependence on prior knowledge implies that expertise cannot be systematized, which makes personnel change a complicating factor. Personnel changes can lead to loss of technical knowledge, which cannot be fully documented through records within facility's protocols [23].

In this context, using machine learning (ML) and artificial intelligence (AI) in DWT processes offers promising opportunities for improvement. In contrast to conventional methods, ML models can identify complex and non-linear relationships between water quality parameters and chemical dosing [24–26]. The ability to develop and embed these learned relationships into Decision Support System (DSS) allows for consistent dosing recommendations, thereby reducing inconsistencies by creating a well-structured electronic knowledge base [22]. ML-based DSS with its tools alongside historical data and operators' hands-on knowledge, can transform tacit dosing expertise into a transparent, data-driven framework [27, 28]. The shift in perspective from being operator-dependent to data-driven decisions means that dosing optimization through ML-based modelling should be assessed not only for technical accuracy but also from a technological, organizational, and societal point of view.

This paper, positioned at the intersection of machine learning-based decision support and CE-oriented water utility management, explores their combined implications for chemical dosing optimization in DWTPs. It is complemented by preliminary empirical insights from a case study conducted at DWTP Makiš, operated by the Public Utility Company Belgrade Waterworks and Sewerage. The paper is set around three research objectives: (1) to analyze the state of the art of the various ML and AI approaches being used in decision support regarding chemical dosing in the DWTP; (2) to examine CE principles and their implementation conditions within the DWTP context; and (3) to demonstrate the potential applicability of the Random Forest-based model at DWTP Makiš.

2 Conceptual Background

2.1 Circular Economy Dimension of Chemical Dosing

Morseletto et al. [29] explain the concept of the CE of water to be a system approach that encompasses reduction, preservation, and optimization of the consumption of water by minimizing waste, efficient usage, and conservation of its quality by protecting the environment. Likewise, Smol et al. [30] explain that moving towards circularity in the water industry will require more than just technological changes, necessitating a shift in the organization and society as well, using a model that consists of the following: reduction, reclamation, reuse, recycling, recovery, and rethinking. While many CE discussions related to the water sector have concentrated on the recycling of wastewater and resource recovery, in DWT poor efficiency may lead to excessive use of chemicals, increased amounts of sludge and energy consumption [13, 22, 30–33]. Bearing that in mind, sustainability of DWT requires ensuring both the quality of water and health protection of populations, as well as improving efficiency, reducing waste, and minimizing environmental impacts associated with DWT. Furthermore, providing safe drinking water for everyone necessitates closer coordination between DWT technologies, monitoring and evaluation, risk assessment, and decision making amidst challenges posed by pollution, water scarcity, and fluctuating water flows [34, 35].

Regarding DWTPs, the idea of sustainability is tied with the capability of producing high-quality potable water and decreasing the materials and energy consumption associated with the process. This means that DWTPs must guarantee high effectiveness in treating the water while consuming fewer chemicals, using less energy, producing less sludge, and managing environmental risks. Water can be contaminated by physical, chemical, biological, and radiological contaminants, which means that the treatment process needs to eliminate them efficiently while including risk management and operational flexibility [9, 35].

Another sustainability viewpoint is that DWT needs to be considered in terms of the resilience of water systems, variations in source-water quality, and the organizational prerequisites for the circular transition. Boltz et al. [36] stress that the notion of resilience is a critical prerequisite for sustainable development (SD) because of the variation in water availability and its influence on the parameters for SD. In terms of the quality of the source water, it can be said that both environmental factors as well as anthropogenic influences play an important role in affecting the quality of streams and rivers in different ways [37]. Simultaneously, there are technological barriers to the shift towards circular water management practices, as well as substantial energy demands, insufficient regulations, and poor social acceptance, suggesting that sustainable adoption relies just as much on governance and coordination as technology [15].

In light of circular and sustainability principles, DWTPs are critical infrastructures whose operation is evaluated not just by whether they comply with regulatory standards, but also based on how resilient, efficient, and effective they are in reducing resource waste [30, 38, 39]. In such circumstances, enhancing monitoring and dosing

control within DWTPs can be seen as a key method for achieving the goals of public health, sustainability, and CE.

Chemical dosing is a crucial factor in DWTPs since it affects how well the treatment works, how much energy is used, how much sludge is produced, the overall cost of treatment, and the quality of the treated water [22, 35]. Coagulants and flocculants are especially important for the clarifiers in the treatment process. Dosing that is too low may negatively affect the quality of the treated water while overdosing would lead to increased chemical consumption and sludge formation, which could result in an increased impact on the environment. Optimization of chemical dosing can be considered as an opportunity to achieve both sustainability and circularity goals for DWTPs, especially in case of variable raw water quality since a more accurate dosage will allow for better water treatment efficiency and cost-effectiveness. Coagulant dosing optimization with the help of big data proves that by improving the dosage process, not only can coagulants' consumption be significantly reduced but their effectiveness be increased as well [22, 28, 35].

2.2 Artificial Intelligence and Machine Learning Approaches for Chemical Dosing

The implementation of ML-based models into DWTPs has demonstrated benefits for enhancing operational reliability while also supporting transitioning to smart and resource-efficient drinking water systems [40]. Advanced machine learning approaches including Artificial Neural Networks (ANNs), Tree-Based and Ensemble Methods, and Deep Learning Architectures allow transition from rough approximations to a more accurate prediction of optimal chemical doses within drinking water facilities.

ANNs, primarily multilayer perceptrons (MLP) and generalized regression neural networks (GRNN), were among the first approaches for designing a data-driven model, establishing the technical viability of using learning-based models for predicting optimal chemical dosing and improving the overall efficiency of the water treatment [40–42]. Mälzer and Strugholtz [42] demonstrated that MLP could be installed at a full-scale plant for online coagulation control, while Dadebo et al. [27] reported strong predictive capability in coagulant dosage (R^2 = 0.94) which allows data informed dosing and annual cost savings (USD 8,036) through resource efficiency. Merging fuzzy logic principles with neural learning, improved adaptability under variable influent conditions, which resulted in neuro-fuzzy systems to extend traditional ANNs' capabilities for providing feasible solutions [43]. Despite these advantages, challenges remain related to poor data integration and generalizability, model interpretability and handling time-dependent process dynamics [25, 44]. Additional issues concern logistical and regulatory issues associated with successful deployment [27].

Tree-Based and Ensemble Methods, especially Random Forest (RF) model, have become highly effective tools when it comes to optimization of chemical dosing in DWTPs, as they demonstrate remarkable predictive accuracy, robustness and feasibility as operational tool [45, 46]. Application of RF and gradient boosting ensembles show strong performance on nonlinear water quality dataset, with studies reporting

quantifiable reduction in chemical usage in full scale settings [28, 46]. Random Forest is especially well-suited for environmental and water treatment applications due to its robustness to multicollinearity common in water quality datasets, and ability to effectively capture interactions between process variables without necessitating adherence to rigorous data structure assumptions [47, 48]. The balance of predictive accuracy, interpretability and low computational requirements makes RF well-suited for real time support of operational decisions, including those of low-resource utilities [49]. Preference for RF-based models in water utilities can be understood through organizational justification beyond technical performance, because they can be explained to non-technical staff, rather than state-of-the-art architecture [50]. Moussouni et al. [51] showcased that RF-Extra tree hybrid model can emerge as promising tool in full scale at real time, achieving $R^2 = 0.985$, with jar testing used for calibration and validation. While RF handles large datasets with many variables effectively, its performance can degrade when it comes to small datasets due to limited training time [44, 52, 53].

Deep learning models, with their multi-layered architecture, can progressively extract high-level feature representation from raw data input. Several deep learning architectures including Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU), and Graph Attention-based Multivariate Time Series Forecasting Model (GAMTF) remain at the laboratory stage of research, but with promising predictive performance [25, 54, 55]. Using large datasets gathered continuously, these models enable adaptive control of treatment processes to provide real-time monitoring and optimal dosing, because they utilize specialized gating mechanisms to capture temporal dependencies in time-series water quality data [22, 41, 44]. Automated feature extraction and selection offered with deep learning enables the elimination of manual dimensionality reduction, which are typically necessary when using traditional ANNs [41, 44, 56]. In particular, integrating operator knowledge during training of the models will provide more accurate predictions as well as create better optimization strategies [22]. Lin et al. 2023 [41] reported that predictions of using GRU and GAMTF models were reasonably aligned with operational chemical dose value, showcasing its advantage over MLP approach. However, such prediction performance has as a result models that require high-frequency data sets for training and are more prone to overfitting. The effectiveness of these technologies is highly dependent on the presence of adequate, good quality and long-term data sets that are used to develop these models [22, 41, 57]. In some instances, deep learning models are not as good at generalizing when they include temporal engineering as a separate layer with fewer variables, which makes then not as suitable for optimal chemical dosing as classic ML models [55].

The transition of ML-based models from research validation to active deployment within DWTPs remains significant challenge for researchers in this field [50]. Yu et al. [44] proposed three-stage integration strategy to ensure user trust and process safety through gradual adoption. Shown in Fig 1, the framework illustrates a gradual transition from passive recommendation systems toward operational integration.

Fig 1. Graduated three-phase deployment framework for ML-based DSS in drinking water treatment utilities.

3 Case Study: Random Forest-Based Decision Support System at DWTP Makiš

The case study was conducted at DWTP Makiš, the primary water supply facility for the City of Belgrade, operated by the Public Utility Company Belgrade Waterworks and Sewerage. Within the plant, the treatment process consists of multiple interdependent stages, meaning that chemical dosing in the clarification stage has impact on downstream units [58]. Coagulant aluminum sulfate $Al_2(SO_4)_3$ (hereafter referred as Al-S) and flocculant polyacrylamide (PAA) are the main chemical inputs daily managed by the operator.

The model is based entirely on existing laboratory monitoring records, spanning four years (2021 – 2024) and eight physicochemical parameters alongside operator-determined doses of Al-S and PAA. Season was encoded as a categorical variable and in order to preserve seasonal representativeness across training and test sets, a stratified sampling strategy was applied, according to Xu et al. [46] recommendation. A RF regressor was used as the primary model, trained on 100 estimators with min-max normalization and fixed random state. Two sequential models were developed for predicting operational doses of Al-S and PAA based on influent water quality, evaluating the capacity of data-driven models to replicate full-scale chemical dosing patterns. The analysis of operational concordance was conducted to match predicted doses within discrete classes used in plant practice and to determine the frequency with which there would be alignment between RF-based model and actual operating doses.

RF-based model predictions for Al-S deviated by two or more dose classes in only 2.1% of test cases and 1.4% for PAA, with all remaining errors limited to adjacent dosing classes (±10 mg/L for Al-S and ±0.1 mg/L for PAA). Overall concordance rates are calculated as the sum of diagonal cells divided by total test cases, Fig 2.

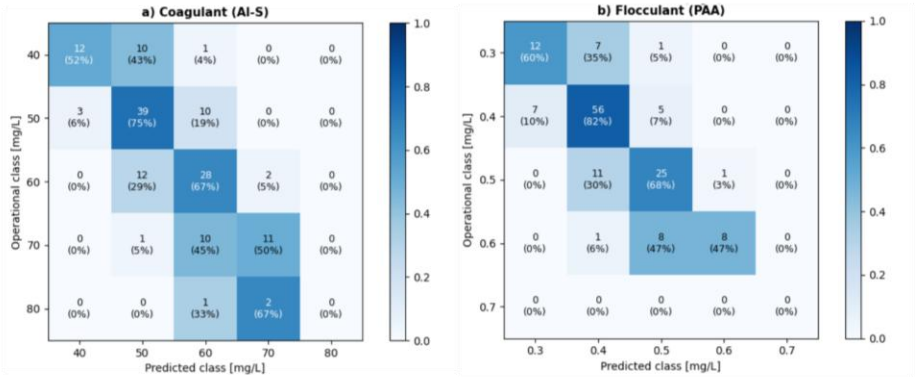


Fig 2. Operational dose class concordance matrices for the RF model: (a) coagulant Al-S and (b) flocculant PAA.

At the exact class level, the model reproduced the operator's dose decision in 63.4% of AL-S cases and 71.1% of PAA cases, as illustrated in the confusion matrices. Since this model is developed using routine laboratory data only (no extra infrastructure investment), these concordance ratios may be seen as a reflection of deployment readiness for Phase 1 Shadow Mode deployment [44]. Results show that there was an agreement with experienced operator judgement on approximately 2/3 of coagulant decisions and 3/4 of flocculant decisions, which provides a valid operational benchmark. There are two main reasons for this judgment. The first reason is that the Al-S concordance was at its weakest at the extremes of the doses; 52% of the time for 40 mg/L and 50% of the time for 70 mg/L shows that the model has a tendency to fall between those two limits for an unlikely raw water source. Therefore, operators should use more careful control and quality check of ML-based results at times when conditions require applying doses that deviate from the standard range.

4 Discussion

In order to evaluate preliminary findings from case study against existing literature, the implementation of ML-based solutions in DWTPs should be considered within four conditions: technological, economic, sociocultural and regulatory.

Adopting AI-based DSS in DWTPs is a subject to a variety of interconnected technology-related barriers in addition to those posed by infrastructure needs. Primary barrier lies in need to acquire new sensor networks and integrate Supervisory Control and Data Acquisition (SCADA) systems, with the operational reliability of DWTPs being an equally important concern [50]. Sudden changes in plant operating conditions or influent water quality may lead to substantial accuracy degradation of models developed on historical data. Determining if that is a non-trivial management challenge and whether such degradation remains within an acceptable operating range is of utmost importance [59]. To address these multiple barriers, both the literature and practitioner guidance emphasize two principles: choosing simple and interpretable

models to start with and conducting pilot studies to validate the performance of chosen model before scaling it up [50].

From an economic perspective, the use of AI and ML in DSS for chemical dosing requires a careful evaluation of the costs involved in implementing these systems compared to the financial benefits they are expected to bring in the future. This may include costs on software development or acquisition, establishment of the data infrastructure required for modeling, integrating sensors, maintaining models, and training staff, all of which can pose a substantial obstacle for those organizations that have little technical and financial capabilities [27, 53]. On the other hand, the economic advantages associated with the application of intelligent decision-support tools include higher dosing accuracy and consequently reduced expenses for the excessive consumption of coagulants and flocculants, decreased sludge generation and its subsequent removal, and greater stability of water treatment facility operations in the face of variability of the water supply [27, 60]. Economic efficiency, therefore, goes beyond cost savings in chemical dosing agents. It encompasses an enhancement of operational effectiveness such as minimizing the burden on operators, decreasing the frequency of necessary interventions, and optimizing the utilization of existing treatment capacities.

AI-based DSS success in DWTPs lies in accurate predictions, as well as in social aspects, such as operators' trust, training, organizational acceptance, and the ability of the water utility organization to implement necessary changes [50, 61]. Moreover, there is a need to train people in the technical aspects of digital decision-making as well as understanding data and how information models work for them to be able to trust new systems [62, 63]. Also, adoption of AI-based DSS may imply certain organizational changes that need to be introduced in the process, meaning that sociocultural conditions are an important aspect of the problem that needs to be considered when evaluating the potential for using artificial intelligence [64, 65].

At the same time, regulatory factors play a significant role in the implementation of the AI-based DSS in DWTPs due to the very nature of such systems. Contrary to purely technical optimization algorithms, AI dosing systems in DWTPs should perform under the pressure of stringent standards and regulations. The processes using AI in water systems need to be transparent, primarily in cases when the model output can lead to changes with health implications. In this regard, some machine-learning methods might affect their adoption negatively unless their recommendations are interpretable, verifiable, and auditable according to current guidelines and laws [61, 66, 67]. The issue of trustworthiness complicates adoption even further: the European Union’s framework for trustworthy AI states that such system should be lawful, ethical and robust, which is something that many black-box architecture struggle to achieve in regulated utility environments [68]. The use of AI in water management has raised wider questions about governance, such as the need to clarify roles, maintain human supervision, and establish appropriate organizational mechanisms to hold people accountable for decisions made using digital technologies [69, 70].

The findings from the case study at DWTP Makiš demonstrate that the technological condition for Phase 1 shadow mode deployment can be met without capital investment to improve the existing infrastructure. Considered alongside the economic,

sociocultural and, regulatory conditions examined in this discussion, models developed using existing operational data are both technically and economically viable, with the literature indicating a relatively low implementation cost associated with this approach. The RF-based model provides a data-informed reference that acts as supporting tool for preventing excess chemical inputs at the clarification stage, reducing both the disposal burden associated with the sludge as well as the environmental footprint of the entire treatment process. The sociocultural and regulatory conditions require more deliberate organizational investment to achieve success through building operator trust, developing institutional AI competence and establishing governance frameworks appropriate in terms of the water utility management.

5 Conclusion

The pressure resulting from environmental degradation, climate change, and resource scarcity, is creating a demand for a transition towards circularity within the water industry. In this regard, the concept of DWTP must be considered beyond the technical aspect of meeting certain regulatory standards and made more efficient and resilient. For that, CE becomes essential, as operational inefficiency can lead to increased chemical usage, sludge generation, energy usage, and overall environmental load.

Current literature implies that the application of ML and AI could bring substantial benefits for helping determine chemical dosages under varying conditions of raw water quality. Based on connections between characteristics of the source water and requirements for its processing, ML-based DSS might assist in converting a subjective approach to dosing operations into a more systematic and dynamic process. In this regard, the Random Forest technique utilized in the DWTP Makiš analysis can be regarded as an encouraging example of leveraging historical information from a plant's operation for making decisions during the operational phase.

Additionally, the adoption of AI-based DSS in DWTPs is not limited to considerations related to forecasting effectiveness alone. It also requires technological maturity, financial viability, social acceptability, and regulatory alignment. Effective data management systems, model clarity, employee training, operational confidence, organizational flexibility, and responsibility are all essential requirements for effective implementation. Hence, moving toward AI-assisted circular water treatment can be regarded as both a technological revolution and a management issue.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. UNESCO: The United Nations World Water Development Report 2024: Water for prosperity and peace. UNESCO (2024). https://www.unesco.org/reports/wwdr/2024/en, last accessed 2026/04/03

2. WHO/UNICEF: Progress on Household Drinking Water, Sanitation and Hygiene 2000–2022: Special Focus on Gender. WHO/UNICEF Joint Monitoring Programme (2023). https://data.unicef.org/resources/jmp-report-2023/, last accessed 2026/04/10
3. Liu, W., Fu, Z., Van Vliet, M.T.H., Davis, K.F., Ciais, P., Bao, Y., Bai, Y., Du, T., Kang, S., Yin, Z., Fang, Y., Wada, Y.: Global overlooked multidimensional water scarcity. Proc. Natl. Acad. Sci. 122, e2413541122 (2025). https://doi.org/10.1073/pnas.2413541122
4. United Nations: The Sustainable Development Goals Report 2025. United Nations, New York (2025). https://unstats.un.org/sdgs/report/2025/The-Sustainable-Development-Goals-Report-2025.pdf, last accessed 2026/04/03
5. Bojarczuk, A., Jelonkiewicz, Ł., Lenart-Boroń, A.: The effect of anthropogenic and natural factors on the prevalence of physicochemical parameters of water and bacterial water quality indicators along the river Białka, southern Poland. Environ. Sci. Pollut. Res. 25, 10102–10114 (2018). https://doi.org/10.1007/s11356-018-1212-2
6. Jinbing, X., Long, W., Jie, Z., Tianlun, Z.: Hydrochemical variation characteristics and driving factors of surface water in arid areas: a case study of Beichuan River in Northwest China. Front. Environ. Sci. 12, 1493390 (2024). https://doi.org/10.3389/fenvs.2024.1493390
7. Zhang, X., Yu, B., Xin, Z., Cong, M., Zhang, C.: Spatial–temporal variations of river water quality under human-induced land use changes in large river basins. Sci. Rep. 15, 36955 (2025). https://doi.org/10.1038/s41598-025-20876-z
8. United Nations Department of Economic and Social Affairs (UN DESA): The Sustainable Development Goals Report 2023: Special Edition. United Nations, New York (2023). https://unstats.un.org/sdgs/report/2023/, last accessed 2026/04/10
9. Ejiohuo, O., Onyeaka, H., Akinsemolu, A., Nwabor, O.F., Siyanbola, K.F., Tamasiga, P., Al-Sharify, Z.T.: Ensuring water purity: mitigating environmental risks and safeguarding human health. Water Biol. Secur. 4, 100341 (2025). https://doi.org/10.1016/j.watbs.2024.100341
10. Chalchisa, D., Megersa, M., Beyene, A.: Assessment of the quality of drinking water in storage tanks and its implication on the safety of urban water supply in developing countries. Environ. Syst. Res. 6, 12 (2018). https://doi.org/10.1186/s40068-017-0089-2
11. Imran, U., Khan, M., Jamal, R., Sahulka, S.Q., Goel, R., Mahar, R., Weidhaas, J.: Probabilistic risk assessment of water distribution system in Hyderabad, Pakistan reveals unacceptable health hazards and areas for rehabilitation. Ecotoxicol. Environ. Saf. 191, 110233 (2020). https://doi.org/10.1016/j.ecoenv.2020.110233
12. Teodosiu, C., Gilca, A.-F., Barjoveanu, G., Fiore, S.: Emerging pollutants removal through advanced drinking water treatment: a review on processes and environmental performances assessment. J. Clean. Prod. 197, 1210–1221 (2018). https://doi.org/10.1016/j.jclepro.2018.06.247
13. García-Ávila, F., Cabello-Torres, R., Iglesias-Abad, S., García-Mera, G., García-Uzca, C., Valdiviezo-Gonzales, L., Donoso-Moscoso, S.: Cleaner production and drinking water: perspectives from a scientometric and systematic analysis for a sustainable performance. South Afr. J. Chem. Eng. 45, 136–148 (2023). https://doi.org/10.1016/j.sajce.2023.05.003
14. Ellen MacArthur Foundation: The Circular Economy Definition and Model Explained. Ellen MacArthur Foundation (2013). https://www.ellenmacarthurfoundation.org/topics/circular-economy-introduction/overview, last accessed 2026/04/10
15. Farfán Chilicaus, G.C., Cruz Salinas, L.E., Silva León, P.M., Lizarzaburu Aguinaga, D.A., Vera Zelada, P., Vera Zelada, L.A., Luque Luque, E.O., Licapa Redolfo, R., Ramos Farroñán, E.V.: Circular economy and water sustainability: systematic review of water man-

agement technologies and strategies (2018–2024). Sustainability 17, 6544 (2025). https://doi.org/10.3390/su17146544
16. Korhonen, J., Nuur, C., Feldmann, A., Birkie, S.E.: Circular economy as an essentially contested concept. J. Clean. Prod. 175, 544–552 (2018). https://doi.org/10.1016/j.jclepro.2017.12.111
17. Hogeboom, R.J., Kamphuis, I., Hoekstra, A.Y.: Water sustainability of investors: development and application of an assessment framework. J. Clean. Prod. 202, 642–648 (2018). https://doi.org/10.1016/j.jclepro.2018.08.142
18. Intergovernmental Panel on Climate Change (IPCC): Climate Change 2022: Impacts, Adaptation and Vulnerability. Cambridge University Press, Cambridge (2022). https://doi.org/10.1017/9781009325844
19. European Commission: A New Circular Economy Action Plan: For a Cleaner and More Competitive Europe. European Commission, Brussels (2020). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2020%3A98%3AFIN, last accessed 2026/04/10
20. Gitis, V., Hankins, N.: Water treatment chemicals: trends and challenges. J. Water Process Eng. 25, 34–38 (2018). https://doi.org/10.1016/j.jwpe.2018.06.003
21. Achite, M., Farzin, S., Elshaboury, N., Valikhan Anaraki, M., Amamra, M., Toubal, A.K.: Modeling the optimal dosage of coagulants in water treatment plants using various machine learning models. Environ. Dev. Sustain. 26, 3395–3421 (2024). https://doi.org/10.1007/s10668-022-02835-0
22. Kim, J., Hua, C., Kim, K., Lin, S., Oh, G., Park, M.-H., Kang, S.: Optimizing coagulant dosage using deep learning models with large-scale data. Chemosphere 350, 140989 (2024). https://doi.org/10.1016/j.chemosphere.2023.140989
23. Abdulameer, L., Al Maimuri, N.M.L., Hassan Nama, A., Lafta Rashid, F., Al-Dujaili, A.N.: The role of artificial intelligence in managing sustainable water resources: a review of smart solution implementations. Water Conserv. Manag. 9, 01–11 (2025). https://doi.org/10.26480/wcm.02.2025.281.291
24. Farhaoui, M., Derraz, M.: Review on optimization of drinking water treatment process. J. Water Resour. Prot. 8, 777–786 (2016). https://doi.org/10.4236/jwarp.2016.88063
25. Aliashrafi, A., Zhang, Y., Groenewegen, H., Peleato, N.M.: A review of data-driven modelling in drinking water treatment. Rev. Environ. Sci. Biotechnol. 20, 985–1009 (2021). https://doi.org/10.1007/s11157-021-09592-y
26. Ortiz-Lopez, C., Bouchard, C., Rodriguez, M.: Machine learning models with potential application to predict source water quality for treatment purposes: a critical review. Environ. Technol. Rev. 11, 118–147 (2022). https://doi.org/10.1080/21622515.2022.2118084
27. Dadebo, D., Obura, D., Etyang, N., Kimera, D.: Economic and social perspectives of implementing artificial intelligence in drinking water treatment systems for predicting coagulant dosage: a transition toward sustainability. Groundw. Sustain. Dev. 23, 100987 (2023). https://doi.org/10.1016/j.gsd.2023.100987
28. Feng, L., Zhang, Y., Wei, X., Wang, M., Niu, Z., Wang, C.: Interpretable prediction of coagulant dosage in drinking water treatment plant based on automated machine learning and SHAP method. J. Water Process Eng. 75, 107925 (2025). https://doi.org/10.1016/j.jwpe.2025.107925
29. Morseletto, P., Mooren, C.E., Munaretto, S.: Circular economy of water: definition, strategies and challenges. Circ. Econ. Sustain. 2, 1463–1477 (2022). https://doi.org/10.1007/s43615-022-00165-x

30. Smol, M., Adam, C., Preisner, M.: Circular economy model framework in the European water and wastewater sector. J. Mater. Cycles Waste Manag. 22, 682–697 (2020). https://doi.org/10.1007/s10163-019-00960-z
31. Silva, J.A.: Wastewater treatment and reuse for sustainable water resources management: a systematic literature review. Sustainability 15, 10940 (2023). https://doi.org/10.3390/su151410940
32. Voulvoulis, N.: Water reuse from a circular economy perspective and potential risks from an unregulated approach. Curr. Opin. Environ. Sci. Health 2, 32–45 (2018). https://doi.org/10.1016/j.coesh.2018.01.005
33. Molinos-Senante, M., Sala-Garrido, R.: Evaluation of energy performance of drinking water treatment plants: use of energy intensity and energy efficiency metrics. Appl. Energy 229, 1095–1102 (2018). https://doi.org/10.1016/j.apenergy.2018.08.102
34. Amarasinghe, P., Liu, A., Egodawatta, P., Barnes, P., McGree, J., Goonetilleke, A.: Modelling resilience of a water supply system under climate change and population growth impacts. Water Resour. Manag. 31, 2885–2898 (2017). https://doi.org/10.1007/s11269-017-1646-1
35. Kundu, D., Dutta, D., Joseph, A., Jana, A., Samanta, P., Bhakta, J.N., Alreshidi, M.A.: Safeguarding drinking water: a brief insight on characteristics, treatments and risk assessment of contamination. Environ. Monit. Assess. 196, 180 (2024). https://doi.org/10.1007/s10661-024-12311-z
36. Boltz, F., LeRoy Poff, N., Folke, C., Kete, N., Brown, C.M., St. George Freeman, S., Matthews, J.H., Martinez, A., Rockström, J.: Water is a master variable: solving for resilience in the modern era. Water Secur. 8, 100048 (2019). https://doi.org/10.1016/j.wasec.2019.100048
37. Hamid, A., Bhat, S.U., Jehangir, A.: Local determinants influencing stream water quality. Appl. Water Sci. 10, 24 (2020). https://doi.org/10.1007/s13201-019-1043-4
38. Liu, J., Shao, Z., Wang, W.: Resilience assessment and critical point identification for urban water supply systems under uncertain scenarios. Water 13, 2939 (2021). https://doi.org/10.3390/w13202939
39. Tchórzewska-Cieślak, B., Pietrucha-Urbanik, K., Kuliczkowska, E.: An approach to analysing water consumers' acceptance of risk-reduction costs. Resources 9, 132 (2020). https://doi.org/10.3390/resources9110132
40. Wang, Y.-Q., Wang, H.-C., Xiao, Z.-J., Bu, L.-J., Li, J., Feng, X.-C., Liang, B., Liu, W.-Z., Sun, F.-Y., Zhou, S.-Q., Wang, A.-J.: Machine learning strategy secures urban smart drinking water treatment plant through incremental advances. Water Res. 280, 123541 (2025). https://doi.org/10.1016/j.watres.2025.123541
41. Lin, S., Kim, J., Hua, C., Kang, S., Park, M.-H.: Comparing artificial and deep neural network models for prediction of coagulant amount and settled water turbidity: lessons learned from big data in water treatment operations. J. Water Process Eng. 54, 103949 (2023). https://doi.org/10.1016/j.jwpe.2023.103949
42. Mälzer, H.-J., Strugholtz, S.: Artificial neural networks for cost optimization of coagulation, sedimentation and filtration in drinking water treatment. Water Supply 8, 383–388 (2008). https://doi.org/10.2166/ws.2008.086
43. Narges, S., Ghorban, A., Hassan, K., Mohammad, K.: Prediction of the optimal dosage of coagulants in water treatment plants through developing models based on artificial neural network fuzzy inference system (ANFIS). J. Environ. Health Sci. Eng. 19, 1543–1553 (2021). https://doi.org/10.1007/s40201-021-00710-0
44. Yu, W., Zhou, Z., Yang, Y., Li, X., Lu, Z., Liu, Y., Ren, J., Liu, Y.: Artificial intelligence in chemical dosing for drinking water treatment: a systematic review of algorithmic appli-

cations, implementation frameworks, and future challenges. J. Environ. Chem. Eng. 13, 118936 (2025). https://doi.org/10.1016/j.jece.2025.118936
45. Shin, D., Lee, J.Y.: Machine learning regression-based prediction model for the autonomous control of coagulant dosing in smart water purification plants. Water Pract. Technol. 19, 4036–4048 (2024). https://doi.org/10.2166/wpt.2024.230
46. Xu, S., Wang, S., Yang, Y., Zheng, L., Xiao, F.: Importance analysis of key parameters for coagulation dosage prediction based on machine learning models. Process Saf. Environ. Prot. 205, 108152 (2026). https://doi.org/10.1016/j.psep.2025.108152
47. Bressane, A., Goulart, A.P.G., Melo, C.P., Gomes, I.G., Loureiro, A.I.S., Negri, R.G., Moruzzi, R., Reis, A.G.D., Formiga, J.K.S., Da Silva, G.H.R., Thomé, R.F.: A non-hybrid data-driven fuzzy inference system for coagulant dosage in drinking water treatment plant: machine-learning for accurate real-time prediction. Water 15, 1126 (2023). https://doi.org/10.3390/w15061126
48. Achite, M., Samadianfard, S., Elshaboury, N., Sharafi, M.: Modeling and optimization of coagulant dosage in water treatment plants using hybridized random forest model with genetic algorithm optimization. Environ. Dev. Sustain. 25, 11189–11207 (2023). https://doi.org/10.1007/s10668-022-02523-z
49. Gonzales Medina, R.I., Carlos Mendoza, J.A., Zuñiga Goyzueta, E.J., Morán-Silva, R.M., López-Gonzales, J.L.: Random forest model for optimizing coagulant doses in drinking water treatment: application at the Miguel de la Cuba Ibarra plant. Environments 13, 17 (2025). https://doi.org/10.3390/environments13010017
50. Vekaria, D., Sinha, S.: aiWATERS: an artificial intelligence framework for the water sector. AI Civ. Eng. 3, 6 (2024). https://doi.org/10.1007/s43503-024-00025-7
51. Moussouni, A., Zeghmar, A., Benchaiba, L., Maaliou, A., Spalevic, V., Jaufer, L., Ksibi, M., Helfer, F., Kader, S.: Ensemble learning-driven optimization of coagulant dosing for drinking water treatment plants using a scalable framework for smart and sustainable process control. Environ. Res. 288, 123229 (2026). https://doi.org/10.1016/j.envres.2025.123229
52. Randive, P., Bhagat, M.S., Bhorkar, M.P., Bhagat, R.M., Vinchurkar, S.M., Shelare, S., Sharma, S., Beemkumar, N., Hemalatha, S., Kumar, P., Kedia, A., El Sayed Massoud, E., Gupta, D., Lozanovic, J.: Adaptive optimization of natural coagulants using hybrid machine learning approach for sustainable water treatment. Sci. Rep. 15, 16096 (2025). https://doi.org/10.1038/s41598-025-96750-9
53. Ding, C., Shen, L., Liang, Q., Li, L.: Machine learning in flocculant research and application: toward smart and sustainable water treatment. Separations 12, 203 (2025). https://doi.org/10.3390/separations12080203
54. Sharafi, M., Rezaverdinejad, V., Behmanesh, J., Samadianfard, S.: Development of long short-term memory along with differential optimization and neural networks for coagulant dosage prediction in water treatment plant. J. Water Process Eng. 65, 105784 (2024). https://doi.org/10.1016/j.jwpe.2024.105784
55. Peng, S., Guo, Y., Wang, J., Wang, Y., Zhang, W., Zhou, X., Jiang, L., Lai, B.: The coagulation-precipitation turbidity prediction model for precision drug delivery system based on deep learning and machine vision. J. Environ. Chem. Eng. 12, 112211 (2024). https://doi.org/10.1016/j.jece.2024.112211
56. Li, L., Rong, S., Wang, R., Yu, S.: Recent advances in artificial intelligence and machine learning for nonlinear relationship analysis and process control in drinking water treatment: a review. Chem. Eng. J. 405, 126673 (2021). https://doi.org/10.1016/j.cej.2020.126673

57. Lin, S., Kim, J., Hua, C., Park, M.-H., Kang, S.: Coagulant dosage determination using deep learning-based graph attention multivariate time series forecasting model. Water Res. 232, 119665 (2023). https://doi.org/10.1016/j.watres.2023.119665
58. Jiang, J.-Q.: The role of coagulation in water treatment. Curr. Opin. Chem. Eng. 8, 36–44 (2015). https://doi.org/10.1016/j.coche.2015.01.008
59. Chowdhury, S., Karanfil, T.: Applications of artificial intelligence (AI) in drinking water treatment processes: possibilities. Chemosphere 356, 141958 (2024). https://doi.org/10.1016/j.chemosphere.2024.141958
60. Liu, J., Long, Y., Zhu, G., Hursthouse, A.S.: Application of artificial intelligence in the management of coagulation treatment engineering system. Processes 12, 1824 (2024). https://doi.org/10.3390/pr12091824
61. Richards, C.E., Tzachor, A., Avin, S., Fenner, R.: Rewards, risks and responsible deployment of artificial intelligence in water systems. Nat. Water 1, 422–432 (2023). https://doi.org/10.1038/s44221-023-00069-6
62. Rapp, A.H., Capener, A.M., Sowby, R.B.: Adoption of artificial intelligence in drinking water operations: a survey of progress in the United States. J. Water Resour. Plan. Manag. 149, 06023002 (2023). https://doi.org/10.1061/JWRMD5.WRENG-5870
63. Jayakumar, D., Bouhoula, A., Al-Zubari, W.K.: Unlocking the potential of artificial intelligence for sustainable water management focusing operational applications. Water 16, 3328 (2024). https://doi.org/10.3390/w16223328
64. Kinney, M., Anastasiadou, M., Naranjo-Zolotov, M., Santos, V.: Expectation management in AI: a framework for understanding stakeholder trust and acceptance of artificial intelligence systems. Heliyon 10, e28562 (2024). https://doi.org/10.1016/j.heliyon.2024.e28562
65. Gerlich, M.: Perceptions and acceptance of artificial intelligence: a multi-dimensional study. Soc. Sci. 12, 502 (2023). https://doi.org/10.3390/socsci12090502
66. Aderemi, I.A., Kehinde, T.O., Daniel Okwor, U., Ahmad, K.H., Adjei, K.Y., Cyriacus Ekechi, C.: Explainable AI for water quality monitoring: a systematic review of transparency, interpretability, and trust. IEEE Sens. Rev. 2, 419–443 (2025). https://doi.org/10.1109/SR.2025.3595500
67. Gacu, J., Monjardin, C., Mangulabnan, R., Pugat, G., Solmerin, J.: Artificial intelligence (AI) in surface water management: a comprehensive review of methods, applications, and challenges. Water 17, 1707 (2025). https://doi.org/10.3390/w17111707
68. Kaur, D., Uslu, S., Rittichier, K.J., Durresi, A.: Trustworthy artificial intelligence: a review. ACM Comput. Surv. 55, 1–38 (2023). https://doi.org/10.1145/3491209
69. Doorn, N.: Artificial intelligence in the water domain: opportunities for responsible use. Sci. Total Environ. 755, 142561 (2021). https://doi.org/10.1016/j.scitotenv.2020.142561
70. Takeda, T., Kato, J., Matsumura, T., Murakami, T., Abeynayaka, A.: Governance of artificial intelligence in water and wastewater management: the case study of Japan. Hydrology 8, 120 (2021). https://doi.org/10.3390/hydrology8030120

Original Research Article *https://doi.org/10.5281/zenodo.23474753*

Measuring Efficiency in Green and Sustainable Energy across Europe Using Data Envelopment Analysis

José Rodrigo Juárez Brenner [0009-0006-8529-0749]

University of Belgrade, Faculty of Organizational Sciences, 11010 Belgrade, Serbia
rodrigojuarezb@gmail.com

Abstract. In the transition to renewable energy sources with a lower economic, social, and environmental impact, efficient electricity generation is crucial to ensuring sustainability. As European countries advance towards climate neutrality goals while guaranteeing a continuous energy supply, understanding the relative efficiency of green energy production has become increasingly important for designing effective national and regional strategies. This study measures the technical efficiency of green and sustainable electricity generation across 34 countries in Europe over the past decade using Data Envelopment Analysis (DEA), proposing an impartial evaluation method to estimate their productivity. An output-oriented DEA model was applied to assess how efficiently countries generate electricity from green sources, which can be used to identify best practices across the region in the design of an energy mix that helps provide a sustainable and continuous electricity supply during their energy transition process.

Keywords: Green Energy, Sustainable Energy, Energy in Europe, Efficiency Measurement, Data Envelopment Analysis.

1 Conceptual framework

Energy generation is vital to economic and technological progress, as it drives modern life and enables the functioning of essential services. It is the foundation of modern infrastructure, without which societies could not develop sustainably. Since the end of the century, the energy transition to renewable sources has begun to take shape in response to the scarcity of fossil fuels and concerns about the environmental impact of electricity production [2]. However, although most primary energy sources have been known since ancient times, the inefficiency of generating electricity through these means has led to a gradual and complex transition, in which countries and industries have adopted them as technological advances have increased their efficiency.

Energy efficiency is fundamental in the analysis and evaluation of generation systems, as it helps determine the extent to which the resources can be transformed into useful energy. In this sense, measuring efficiency is not only a technical issue but also a strategic one, as it has direct economic, environmental, and social implications [16]. Every energy source, regardless of its origin, requires material, financial, and technological inputs; assessing their transformation allows for maximising the use of limited

resources to reduce costs and avoid waste, which is also essential for reducing environmental impacts and ensuring sustainability.

It is important to note that green energy, renewable energy and sustainable energy refer to different concepts, although they are related and sometimes mistakenly used interchangeably. Renewable energy is the one that comes from sources that are naturally replenished on a human timescale; thus, its focus lies on the availability of resources and proper disposal. On the other hand, green energy is the one that produces little or no environmental impact, or its production includes mitigating any negative effects on the environment. While the concept of energy renewability underlies the efficient consumption of inputs and the responsible use of resources, green energy focuses more on the output of generation, expecting the process to have zero impact or that any potential damage or waste is adequately mitigated [10].

Not all renewable energy is green, and vice versa. For instance, burning biomass to generate electricity could be considered renewable, as it utilises agricultural waste that would otherwise be discarded and plants can regrow on a relatively short timescale, but it can also release a significant amount of greenhouse gases and small particles into the environment that are harmful to health and ecosystems. Likewise, nuclear power could be considered green since it has very low or almost no greenhouse gas emissions and helps mitigate the negative effects of other energy sources, but relies on uranium availability, which is not renewable. Mismanagement of radioactive waste can also cause significant environmental damage, so proper disposal is required to prevent leaks or even a large-scale catastrophe.

Sustainable energy directly links power generation with the idea of sustainability, which is a multi-dimensional concept that encompasses several variables, both quantitative and qualitative. In a broad sense, sustainability means “the ability to meet the needs of the present without compromising the ability of future generations to meet their own needs” [1]. It implies imposing limits to mitigate the impact of human activities on resources and the ability to mitigate the imbalances caused in economic, social and environmental systems, and the promotion of fairness in the distribution of resources. Sustainable energy emphasises a long-term balance of energy use: environmentally viable, economically feasible, and socially equitable, encompassing both inputs and outputs [20].

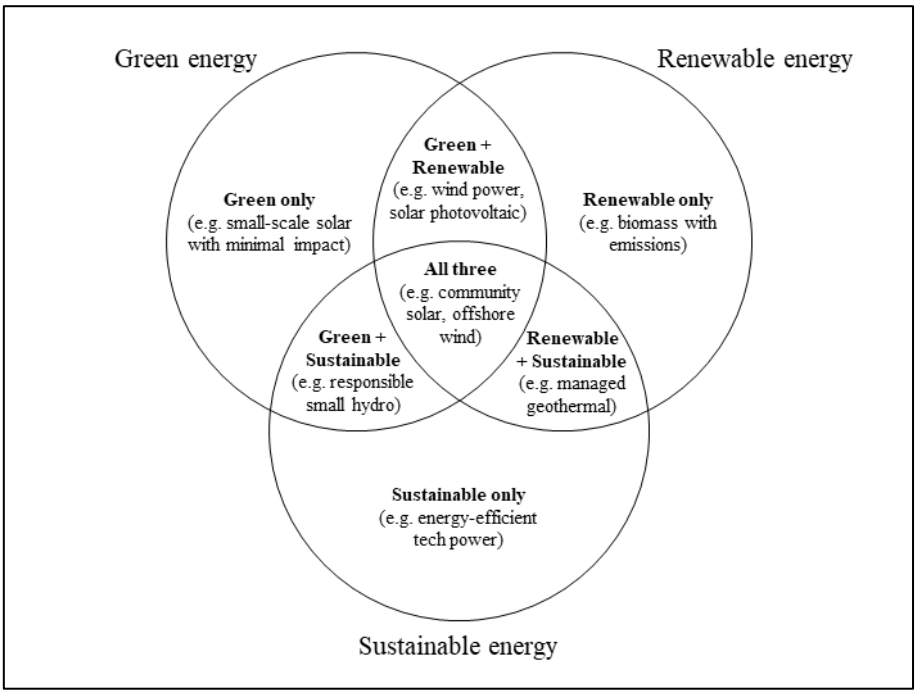


Fig. 1. Venn diagram showing how green, renewable, and sustainable concepts of energy overlap and differ.

Although it is a trendy concept, and many countries and regions are enacting laws and policies on the matter, sustainability remains an ambiguous idea open to various interpretations, which makes its quantitative measurement difficult. One methodology that is increasingly being used to measure the efficiency of sustainability is Data Envelopment Analysis (DEA), as it provides a structured framework for quantifying information that allows for determining whether the objectives are being met, since sustainable development should always reconcile technological development and efficiency [16]. This is especially valuable in the energy sector, where the production process is complex, heterogeneous, and shaped by technological and geographical factors.

According to estimates from the European Environmental Agency, the share of renewable energy has increased from 10% in 2005 to approximately 24.5% of the gross final energy consumption by 2023, achieving the initial target of a 20% share in 2020 [4]. The European Union has set the goal to achieve climate neutrality by 2050, while ensuring a steady supply of electricity at affordable prices. For this purpose, it is crucial to identify and incentivise the efficiency of green energy sources, as this can help countries optimise their energy mix, especially for those lagging behind and lacking strategic resources and structural limitations.

2 Methodology

Data Envelopment Analysis (DEA) is a non-parametric mathematical model used in operations research to measure the relative efficiency of a set of decision-making units (DMUs) that perform comparable activities and consume inputs (resources) to produce outputs (results). DEA proposes an impartial evaluation method that analyses the productivity of several entities that share common characteristics, sorting them according to their performance score and creating a best-practice frontier [20]. Based on the efficiency and benchmarking, organisations can make informed decisions to improve performance, such as reducing resources or increasing production, or determine the best practices of the most efficient entities to replicate them.

The definition of efficiency in DEA is given as the ratio between the sum of weighted outputs and the sum of weighted inputs for each decision-making unit, subject to certain constraints [21]. Considering n DMUs, every DMU_j uses m inputs defined as x_{ij} ($i = 1, \ldots, m$), and s outputs defined as y_{rj} ($r = 1, \ldots, s$), the objective function for an output-oriented model would be:

$$\max_{u,v} E_j = \frac{\sum_{r=1}^{s} u_r \cdot y_{rj}}{\sum_{i=1}^{m} v_i \cdot x_{ij}}$$

Being u_r and v_i the optimal weights, which are subject to the constraints:

$$E_j \leq 1, \qquad \forall j$$

$$u_r, v_i \geq 0$$

The linear programming problem for output-oriented cases may also be expressed as follows:

$$\varphi_j = \max_{\varphi,\lambda} \varphi$$

$$E_j = \frac{1}{\varphi_j}$$

subject to:

$$\sum_{j=1}^{n} \lambda_j \cdot y_{rj} \geq \varphi \cdot y_{r0}, \qquad \forall r$$

$$\sum_{j=1}^{n} \lambda_j \cdot x_{ij} \leq x_{i0}, \qquad \forall i$$

$$\lambda_j \geq 0, \qquad \forall j$$

The equations above express the optimisation problem as output-oriented, identifying which DMUs are efficient by maximising the outputs with the existing inputs. The best-performing DMUs are the ones with an output expansion factor (φ) of 1 or close to 1, being less efficient as the value of φ increases. The technical efficiency of the assessed DMU is given by the ratio $1/\varphi$, which shows in a range between 0 and 1 how relatively efficient it is compared to the other DMUs. A major benefit of this methodology is that it also indicates that the DMU could increase its outputs by $(\varphi - 1) \cdot 100\%$ while maintaining its inputs, providing an objective reference on the efficiency it can achieve.

This methodology is also useful for input-oriented optimisation problems, identifying which DMUs are efficient by minimising the inputs used to achieve their outputs. In this case, the objective function is an equation of minimisation, and the input factor (θ) multiplies the input constraint instead of the outputs. The problem is mathematically expressed as follows:

$$\theta_j = \min_{\theta,\lambda} \theta$$

subject to:

$$\sum_{j=1}^{n} \lambda_j \cdot y_{rj} \geq y_r, \qquad \forall r$$

$$\sum_{j=1}^{n} \lambda_j \cdot x_{ij} \leq \theta \cdot x_i, \qquad \forall i$$

$$\lambda_j \geq 0, \qquad \forall j$$

In this optimisation problem, the best performing DMU will be at the frontier with an input expansion factor (θ) close to 1. The lower this value, the lower the efficiency of the DMU.

The advantages of performing a Data Envelopment Analysis are that it does not require identifying any type of relation between inputs and outputs or any specific statistical distribution for the data. Furthermore, it provides insights on how to improve the performance of inefficient entities, and which variables are more important to improve. DEA is typically useful for measuring efficiency in economic and environmental dimensions, but it has some disadvantages when addressing qualitative issues, such as social ones. Therefore, it can be combined with other methodologies to mitigate its inherent limitations. At the same time, it measures relative, not absolute efficiency, so the results will depend on the selection of DMUs for the analysis.

3 Data description

In order to measure the relative efficiency of European countries in energy generation, an output-oriented DEA model has been conducted with the aim of identifying the countries that have generated electricity most efficiently over the last decade from green energy sources, highlighting those that can also be considered renewable or transitional, and ranking them as a benchmark among European countries, which can then evaluate and potentially replicate the best practices in energy optimisation towards a sustainable energy system.

The information required for the analysis was collected from the database managed by the European Statistical Office, which is provided directly by the member states of the European Union and other associated European countries. A sample of 34 countries was selected, consisting of those that did not show inconsistencies to ensure the reliability of the model. The analysis was performed on annual data from the most recent reported decade, corresponding to the period between 2014 and 2023, using four indicators: three inputs and a single output. These indicators were selected because they properly reflect the main process of electrical transformation, avoiding conflation with supporting activities and preserving comparability across countries and years.

Table 1. Indicators for the Data Envelopment Analysis model.

Inputs	Output	Index
• Net maximum electricity production capacity (MW) • Transformation inputs for electricity generation (GWh) • Self-consumption of electricity production (GWh)	Net electricity production from main activity producers (GWh)	Technical efficiency per country per year

The study includes information on energy generated by six green sources, five of which are considered fully renewable: hydro, geothermal, wind, solar, and marine (tidal and wave). Nuclear energy was considered a green source in the dataset, even though it is not renewable, because it is a very low-emission source and is exceptionally useful for mitigating variability in the energy transition toward sustainable generation. On the other hand, biomass energy was discarded, since it was not possible to identify which data came from renewable bioenergy sources (such as biofuel or wastewater) and which from non-renewable (such as deforestation or over-harvesting), the latter of which cannot be considered green due to pollutant emissions. Hence, sustainable energy has been treated as equivalent to green energy, since the databases contained only technical data but did not feature energy-related social variables.

The analysis considered only net electricity generation and excluded power generation for other purposes, such as heating or transport. Consequently, the data included were obtained only from production plants primarily dedicated to electricity generation,

dismissing those that produce electricity as a byproduct. Furthermore, only information on energy production was considered, omitting data on transportation, supply, and consumption, as these are not affected by the energy source (disregarding the variability in generation from certain sources). For the same reason, energy imports and exports between countries were also excluded, as energy trade is not typically affected by the origin of electricity.

4 Analysis results

Prior to conducting the Data Envelopment Analysis, the share of energy from renewable sources was assessed to determine the impact these sources have on the energy mix of each country, as well as their year-on-year growth from 2014 to 2023, based on the environmental target described above.

Fig. 2. Percentage share of renewable energy sources in electricity in 2023.

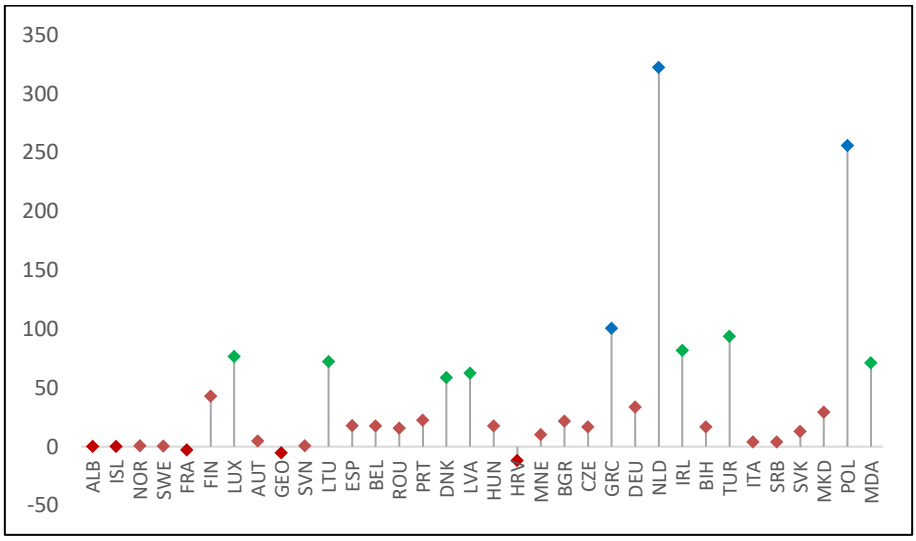


Fig. 3. Percentage growth in the share of renewable energy sources from 2014 to 2023.

An output-oriented DEA model was carried out to determine the relative technical efficiencies of the 34 selected countries according to their generation performance, since the study aimed to identify how green electricity production was obtained given existing capacities and energy-use structures. The analysis employed a classical CCR model (Charnes, Cooper and Rhodes, 1978), seeking a common technical benchmark across countries rather than a decomposition of scale effects. This model is particularly useful to provide an operational and non-parametric benchmark for the region.

The optimisation problem was therefore posed as a maximisation problem, and the constraints were established based on the three input variables and the single output variable, which were jointly weighted to form the efficiency frontier. The optimisation model was performed individually for each year to calculate a relative technical efficiency index for each country over those periods. Datasets with inconsistent or incomplete records were excluded to avoid bias and strengthen the model. An average of the ten computed indices was estimated to rank the performance of all selected countries during the decade.

Table 2. Top ten countries by estimated technical efficiency indices from 2014 to 2023.

Ranking	Country	Avg. Efficiency	Share of 2023
1	Luxembourg	1.000	18.04%
2	Norway	1.000	116.5%
3	Denmark	0.999	79.39%
4	Iceland	0.999	99.89%
5	France	0.996	29.98%
6	Sweden	0.996	87.55%
7	Czechia	0.976	16.45%
8	Slovenia	0.976	41.89%
9	Finland	0.971	52.38%
10	Serbia	0.952	31.75%

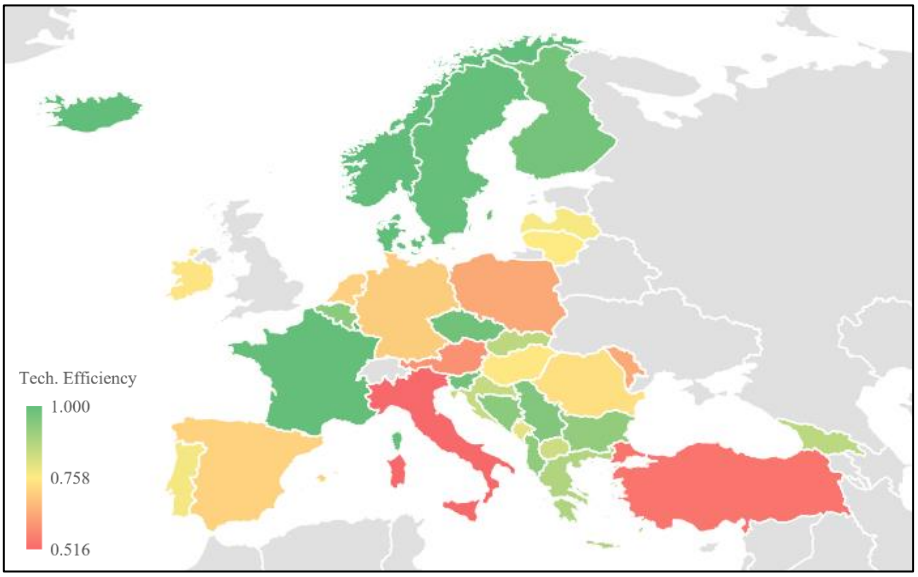


Fig. 4. Map showing the average technical efficiency index from 2014 to 2023.

5 Discussion of results

When evaluating the share of renewable energy sources in electricity, almost all the countries, except for Moldova, exhibit a green energy share above 20% by 2023, which was the target set by the European Environmental Agency. Additionally, 21 of the 34 countries have a share exceeding 50%, indicating that most of the electricity generated in these countries already comes from green sources, which is a positive outcome for the upcoming proposed goals for the continent. Regarding the growth rates of

renewable energy use, 10 of the 34 countries present a growth rate above 50% in the period between 2014 and 2023, with 3 of them exceeding 100%; and 3 countries present negative growth. These growth percentages do not show a correlation with the ranking of countries with the highest green energy share nor with the relative technical efficiency calculated for them.

After calculating the DEA model in each of the years evaluated, several countries obtained technical efficiency indices of 1.000, corresponding to countries that were on the efficiency frontier in their respective years. During the evaluation period, 2 countries achieved a perfect efficiency index for all ten years: Luxembourg and Norway, and another 5 countries remained on the efficiency frontier for more than five years: Denmark, Iceland, France, Sweden, and Czechia. These countries are reflected in the top seven positions in the ranking. Of the other 3 countries in the top ten, Slovenia and Finland were on the efficiency frontier at some point, with Serbia being the exception although it achieved high results throughout the decade. The individual performances of the study can be seen in more detail in Table 3 of the Appendix.

When evaluating the energy mix of the ten highest-ranked countries individually, they all show interesting combinations and quite diverse adoption rates of green sources. Despite its small size, Luxembourg generates almost half of its electricity from hydroelectric plants and is one of the countries that best utilises its solar parks, even without high levels of sunlight. Norway is also heavily dependent on hydropower, which accounts for 89% of its energy mix, despite being a major producer of hydrocarbons. Denmark makes extensive use of wind power, which accounts for 58% of its energy production, and is world-famous for its offshore wind farms. Similarly, Iceland is the country that best leverages the benefits of geothermal energy, being one of the countries that most incorporates this source and the leader in Europe, combining it with hydropower. France, in turn, has a long history of marine energy development and is the undisputed European leader in the use of tidal power, with some of the oldest and most successful plants, and it is also the European leader in nuclear power production, which accounts for two-thirds of its electricity generation.

While the five countries mentioned above depend very little on non-green sources, the others that make up the top ten have more diverse energy mixes. The two Nordic countries have a minimal share of non-green sources, relying mainly on nuclear, hydro, and wind power. Czechia and Slovenia also show a high share of nuclear energy in their mix. Serbia is a curious case in point, as it produces around two-thirds of its energy from non-renewable sources, primarily coal, yet demonstrates a high level of efficiency in the green energy it has adopted, with hydro power being the main source, which denotes a great potential for sustainable energy generation.

6 Conclusions

As technology improves the efficiency of electricity generation from renewable sources and mitigates its inherent variability, it becomes increasingly important for countries to design an efficient energy mix in the transition to environmentally friendly models, ensuring a sustainable and continuous electricity supply. To achieve this, governments and energy companies must make informed decisions based on objective and quantifiable information to invest in sources that maximise productivity and profitability within available resources. In this regard, Data Envelopment Analysis represents a valuable quantitative model for evaluating generation performance and providing an objective and practical benchmark among similar entities in the energy sector.

After applying a classic DEA model to evaluate the performance of 34 European countries in generating electricity from green sources between 2014 and 2023, an annual technical efficiency index was obtained. This index served as the basis for ranking individual production and determining performance during the decade. Subsequently, it was compared with the share of green energy in each country's energy mix and its adoption rate to find out whether there was any correlation, but none was found. However, correspondence was found between countries that recently increased their production from green sources and exhibit higher technical efficiency indices, while those that started their transformation process earlier but did not modernise showed lower values.

When evaluating the energy mix of the ten high-ranking countries individually, it was not possible to identify any practice or use of any energy source that boosted efficiency. Although several are leaders in some type of green energy, they do not demonstrate a favourable use of any specific source, which proposes that the ideal energy mix must be designed for each country according to its available capacities and natural conditions, and there is still considerable room for optimisation. Businesses and policymakers should look at relative performance among countries with similar characteristics to establish policies and design strategies, replicating transferable best practices rather than non-transferable advantages, which is particularly important for countries with structural and geographical constraints.

A notable finding is that five of these ten countries make extensive use of nuclear energy. Even when there is debate about whether it should be considered green, it proves useful as a transition source towards a sustainable model. Concerns about nuclear accidents, rare-earth mining, and radioactive waste disposal have prompted criticism and opposition to nuclear power, even within environmental circles; however, many advocate that it provides a constant energy supply without emitting greenhouse gases, thus contributing to energy independence. In this sense, the phase-out of nuclear power in countries like Germany and Italy would be an interesting case study, as both presented low efficiency indices despite being recognised for sustainable policies, resorting to non-renewable sources to mitigate instabilities.

This study was conducted using a classic output-oriented DEA methodology, with DMUs that have complex operations and share a set of inputs and outputs, working more as a benchmarking tool by comparing individual performances against a relative best-practice frontier, rather than a full account of national energy performances. This approach is practically useful to identify the units that achieve higher output under similar resource conditions but requires in-depth interpretation by business leaders and policymakers to develop energy plans, prioritise investments, and integrate technologies. Since inefficiencies may reflect differences in scale, future analysis could be performed using a BCC-DEA model (Banker, Charnes, and Cooper, 1984), which also comprises variable returns to scale, or with a network DEA model, which evaluates interconnected subprocesses within regional energy systems.

Although the present study focuses on technical efficiency in electricity generation, future research could complement the current DEA model by incorporating social indicators that capture a broader view of sustainable energy or by adjusting the results with social outcomes that better address all three aspects of sustainability. Such an extension would allow the analysis to move beyond mere technical efficiency and provide a more comprehensive assessment of sustainability, serving as a basis for the design of public policies where energy transition outcomes also depend on access to electricity, regional equity, and public participation. Other methodologies, such as stochastic frontier analysis, multicriteria decision-making methods, and hybrid models, could also provide different perspectives of efficiency in sustainable energy that can be compared.

References

1. Brundtland, G. (1987). Report of the World Commission on Environment and Development: Our Common Future. United Nations General Assembly Document A/42/427. https://doi.org/10.1080/07488008808408783
2. Bukowski, M., Majewski, J., & Sobolewska, A. (2023). The environmental impact of changes in the structure of electricity sources in Europe. Energies, 16(1), 501. https://doi.org/10.3390/en16010501
3. Ervural, B. Ç., Ervural, B., & Zaim, S. (2016). Energy Efficiency Evaluation of Provinces in Turkey Using Data Envelopment Analysis. Procedia - Social and Behavioral Sciences, Vol. 235, pp. 139-148. https://doi.org/10.1016/j.sbspro.2016.11.009
4. European Environmental Agency. (2024). Trends and projections in Europe 2024. Publications Office of the European Union. ISSN 1977-8449 https://doi.org/10.2800/7574066
5. European Statistical Office. (2025). Complete energy balances [Dataset]. Eurostat Database. https://doi.org/10.2908/nrg_bal_c
6. European Statistical Office. (2025). Electricity production capacities by main fuel groups and operator [Dataset]. Eurostat Database. https://doi.org/10.2908/nrg_inf_epc
7. European Statistical Office. (2025). Gross and net production of electricity and derived heat by type of plant and operator [Dataset]. Eurostat Database. https://doi.org/10.2908/nrg_ind_peh
8. European Statistical Office. (2025). Share of energy from renewable sources [Dataset]. Eurostat Database. https://doi.org/10.2908/nrg_ind_ren

9. International Energy Agency. (2025). *Sources of electricity generation in Europe since 2000* [Dataset]. IEA Energy System. https://www.iea.org/regions/europe/electricity
10. Halawa, E. (2024). *Sustainable energy: Concept and definition in the context of the energy transition — A critical review*. *Sustainability, 16*(4), 1523. https://doi.org/10.3390/su16041523
11. Halkos, G. E., & Tzeremes, N. G. (2012). Analyzing the Greek renewable energy sector: A Data Envelopment Analysis approach. Renewable and Sustainable Energy Reviews, Vol. 16 (5), pp. 2884-2893. https://doi.org/10.1016/j.rser.2012.02.003
12. Kim, K. T., Lee, D. J., Park, S. J., Zhang, Y., & Sultanov, A. (2015). Measuring the efficiency of the investment for renewable energy in Korea using data envelopment analysis. Renewable and Sustainable Energy Reviews, Vol. 47, pp. 694-702. https://doi.org/10.1016/j.rser.2015.03.034
13. Menegaki, A. N. (2013). *Growth and renewable energy in Europe: Benchmarking with data envelopment analysis*. Renewable Energy, Vol. 60, pp. 363-369. https://doi.org/10.1016/j.renene.2013.05.042
14. Taleb, M., Emrouznejad, A., Charles, V., Khalid, R., & Ramli, R. (2024). An extended-directional mix-efficiency measure: Performance evaluation of OECD countries considering NetZero. Computers & Industrial Engineering, Vol. 189, 109967. ISSN 0360-8352. https://doi.org/10.1016/j.cie.2024.109967
15. Tone, K. (2017). Advances in DEA Theory and Applications: with Extensions to Forecasting Models. John Wiley & Sons. https://doi.org/10.1002/9781118946688
16. Tsaples, G., & Papathanasiou, J. (2021). Data envelopment analysis and the concept of sustainability: A review and analysis of the literature. Renewable and Sustainable Energy Reviews, Vol. 138, 110664. https://doi.org/10.1016/j.rser.2020.110664
17. Veronese da Silva, A., de Oliveira Ribeiro, C., & Rego, E. (2024). A data envelopment analysis approach to measuring socio-economic efficiency due to renewable energy sources in Brazilian regions. Brazilian Journal of Chemical Engineering, Vol. 41, pp. 875–888. https://doi.org/10.1007/s43153-023-00356-5
18. Wang, Q., Jiang, F., & Li, R. (2022). Assessing supply chain greenness from the perspective of embodied renewable energy – A data envelopment analysis using multi-regional input-output analysis. Renewable Energy, Vol. 189, pp. 1292-1305. https://doi.org/10.1016/j.renene.2022.02.128
19. Zeng, Y., Guo, W., Wang, H., & Zhang, F. (2020). A two-stage evaluation and optimization method for renewable energy development based on data envelopment analysis. Applied Energy, Vol. 262, 114363. https://doi.org/10.1016/j.apenergy.2019.114363
20. Zhou, H., Yang, Y., Chen, Y., & Zhu, J. (2018). Data envelopment analysis application in sustainability: The origins, development and future directions. European Journal of Operational Research, Vol. 264 (1), pp. 1-16. https://doi.org/10.1016/j.ejor.2017.06.023
21. Zhu, J. (2014). Quantitative Models for Performance Evaluation and Benchmarking: Data Envelopment Analysis with Spreadsheets (3rd ed.). Springer International Publishing. ISSN 2214-7934. https://doi.org/10.1007/978-3-319-06647-9

Appendix

Table 3. Estimated technical efficiency indices for each country from 2014 to 2023.

Country	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Luxembourg	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Norway	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Denmark	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.990	1.000
Iceland	0.986	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
France	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.962	1.000
Sweden	0.975	1.000	0.984	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Czechia	1.000	1.000	0.870	0.932	0.962	1.000	1.000	1.000	1.000	1.000
Slovenia	1.000	0.896	0.986	0.957	0.987	1.000	1.000	0.995	0.943	0.997
Finland	1.000	1.000	1.000	1.000	1.000	0.981	0.937	0.961	0.893	0.939
Serbia	0.958	0.934	0.941	0.917	0.960	0.933	0.974	0.977	0.933	0.997
Belgium	0.992	0.837	1.000	0.960	0.841	0.957	0.900	1.000	0.957	0.926
Bosnia-Herz.	0.936	0.890	0.879	0.848	0.986	0.999	0.886	0.932	1.000	1.000
Albania	0.958	0.932	0.914	0.832	0.984	0.910	0.926	1.000	0.960	0.867
Bulgaria	0.826	0.820	0.896	0.900	0.904	0.957	1.000	0.968	0.980	0.985
Greece	0.952	0.615	1.000	1.000	0.483	0.896	1.000	1.000	1.000	0.851
Georgia	0.837	0.796	0.854	0.837	0.872	0.870	0.828	0.865	0.933	0.926
Slovakia	0.738	0.754	0.802	0.801	0.779	0.866	0.948	0.927	0.980	1.000
Croatia	1.000	0.901	0.815	0.797	0.855	0.729	0.790	0.870	0.823	0.899
Macedonia	0.794	0.826	0.774	0.642	0.833	0.815	0.914	0.919	0.901	0.919
Montenegro	0.706	0.658	0.707	0.615	0.848	0.860	0.821	0.923	0.888	0.941
Portugal	0.816	0.711	0.782	0.654	0.765	0.788	0.798	0.825	0.788	0.806
Latvia	0.603	0.608	0.832	0.870	0.750	0.777	0.782	0.798	0.793	0.885
Lithuania	0.886	0.771	0.781	0.829	0.734	0.765	0.776	0.697	0.568	0.775
Hungary	0.301	0.240	0.267	1.000	1.000	1.000	0.995	0.928	0.903	0.897
Ireland	0.754	0.748	0.720	0.728	0.736	0.704	0.774	0.751	0.762	0.774
Romania	0.745	0.728	0.695	0.708	0.729	0.685	0.760	0.752	0.780	0.784
Spain	0.761	0.680	0.694	0.678	0.690	0.639	0.728	0.740	0.746	0.760
Netherlands	0.812	0.756	0.738	0.784	0.726	0.744	0.709	0.585	0.599	0.626
Germany	0.741	0.707	0.702	0.722	0.681	0.670	0.717	0.720	0.692	0.672
Moldova	0.758	0.721	0.625	0.647	0.531	0.693	0.658	0.677	0.504	0.557
Poland	0.727	0.690	0.609	0.661	0.600	0.603	0.621	0.609	0.602	0.646
Austria	0.542	0.509	0.550	0.562	0.542	0.595	0.591	0.578	0.632	0.828
Türkiye	0.516	0.566	0.566	0.512	0.509	0.625	0.564	0.480	0.543	0.492
Italy	0.559	0.507	0.500	0.497	0.531	0.530	0.554	0.516	0.464	0.500

THEMATIC SECTION

Creativity, Innovation and Sustainable Management

Resilient and Intelligent Organizations
in Times of Permanent Disruption

SYMORG 2026 • CONFERENCE PROCEEDINGS

Original Research Article *https://doi.org/10.5281/zenodo.23474785*

What influences sustainability reporting in the Republic of Serbia? An Empirical Analysis

Bojana Novićević Čečević[1][0000-0002-4269-0586] , Jelena J. Stanković[0000-0002-9875-9861], Žarko Popovic [0000-0002-4347-6960]

[1] Faculty of Economics, University of Niš,
Trg kralja Aleksandra Ujedinitelja 11, 18000 Niš, Serbia[1]

Abstract. Sustainability reporting represents an important segment of contemporary corporate reporting; however, its implementation in practice is not uniform across companies. The aim of this paper is to examine whether sustainability reporting practices are associated with selected company characteristics and governance structure features. The research was conducted on a sample of 92 of the most successful companies based on operating revenues in 2023. Data were collected from publicly available sources, and the analysis was performed using descriptive statistics and the Pearson chi-square test. The results indicate that more than half of the companies in the sample prepare sustainability reports, but this practice is not equally represented among all firms. No statistically significant relationship was found between sustainability reporting and the gender of the director, or region of establishment. On the other hand, the results reveal a statistically significant relationship between the presence of foreign nationals in governance bodies and sustainability reporting. The findings suggest that the practice of sustainability reporting is more strongly associated with managerial and institutional characteristics of companies than with their regional attributes.

Keywords: sustainability reporting, regional affiliation, director gender, presence of foreign nationals in management.

[1] This paper is part of a project CROSS-REIS – HORIZON WIDERA-2023-ACCESS-03 that has received funding from the European Union's Horizon Europe research and innovation program under grant agreement No. 101136834

This paper is part of the research conducted within the international project "Advancing Data-Based Policy-Making in Urban and Regional Development and Wide-Scale Implementation for Sustainable Environments – UR WISE" (Grant Agreement No. 101187119), funded by the European Union's Horizon Europe Framework Program.

1 Introduction

Sustainability reporting has gained increasing importance in contemporary corporate reporting in recent years, as it enables stakeholders to assess not only the financial results of a company, but also the environmental, social, and governance aspects of its operations. Its importance stems from growing demands for transparency, accountability, and sustainable management, as well as from the need of investors, creditors, customers, and the wider public for more reliable information on the long-term sustainability of a company's business model. At the level of the European Union, this area has been further institutionalized through the Corporate Sustainability Reporting Directive (CSRD), which has modernized and strengthened the rules on the disclosure of social and environmental information, while companies covered by this regime report in accordance with the European Sustainability Reporting Standards (European Commission, 2024).

In the Republic of Serbia, the normative framework for this area was established by the Accounting Law, which, among the issues it regulates, also includes non-financial reporting, thereby introducing into the domestic regulatory framework an obligation to report on non-financial aspects of business operations for certain entities. In addition, the Ministry of Finance of the Republic of Serbia issued the Guidelines for Non-Financial Reporting in Serbia, with the aim of improving the understanding of reporting obligations and the quality of reporting practice (Accounting Law, 2019; Ministry of Finance of the Republic of Serbia, 2025).

However, although the regulatory framework has been formally established, practice shows that the preparation of sustainability reports is not equally developed across all companies. Many companies in Serbia still perceive non-financial reporting as a formality and face numerous challenges in its implementation. It is precisely this gap between normative obligation and actual business practice that represents an important research problem and the main motivation for conducting this study.

Building on this, the aim of the paper is to examine whether the preparation of sustainability reports is associated with certain characteristics of companies and their governance structure, namely the gender of the director, the presence of foreign nationals in governance bodies and the region in which the company was established. Accordingly, the paper consists of a theoretical part, which highlights the importance of sustainability reporting and the regulatory framework in the Republic of Serbia, and an empirical part, in which the relationship between sustainability reporting and selected company characteristics is examined using descriptive statistics and the Pearson chi-square test.

2 Literature Review

Sustainability reporting has developed alongside the expansion of the concept of sustainable business and the growing demands of stakeholders for greater transparency in corporate behavior. The theoretical foundation for understanding this practice is provided by the concept of the triple bottom line, according to which company perfor-

mance is not measured solely by economic outcomes but also by social and environmental contributions (Elkington, 1997), as well as stakeholder theory, which holds that a company must be accountable to a broader group of stakeholders, not only to capital owners (Freeman, 1984). In this context, sustainability reporting is viewed as an important instrument for communication and building trust with the external environment, as previously emphasized in studies on social and environmental disclosure (Gray et al., 1995; Patten, 1991).

In addition, legitimacy theory explains sustainability reporting as a mechanism through which companies seek to align their activities with societal expectations and maintain organizational legitimacy (Patten, 1991). From this perspective, sustainability reporting is not only an instrument of information disclosure, but also a means of strengthening corporate reputation, transparency, and stakeholder trust.

A significant body of research has focused on identifying the factors that influence whether a company engages in sustainability reporting at all. Review studies indicate that the most frequently examined determinants include firm size, profitability, institutional environment, ownership structure, and various corporate governance characteristics (Dienes et al., 2016). Hahn and Kühnen (2013) emphasize that the literature on the determinants of sustainability reporting has developed rapidly, highlighting in particular institutional, organizational, and governance-related factors. Similarly, Dienes et al. (2016) conclude that the decision to report is influenced by both external pressures and internal company characteristics, including governance structures and the strategic orientation of management.

An important stream of research focuses on the role of the institutional and regulatory context. Reverte (2009) shows that the practice and extent of CSR disclosure are associated with company characteristics and the market environment, while Ortas et al. (2015) indicate that the institutional and social context significantly influence the ESG performance of companies participating in voluntary corporate social responsibility initiatives. Khan et al. (2020) further emphasize the multi-level institutional embeddedness of CSR reporting, demonstrating that this practice is shaped by both national regulations and pressures, as well as organizational and managerial decisions. A similar conclusion is reached by Bose and Khan (2022), who find that SDG reporting is associated with country-level institutional factors. Such findings are important for interpreting differences in reporting practices among companies operating in different institutional and governance contexts.

In addition to institutional factors, the literature has also paid particular attention to characteristics of the governance structure. Studies focused on gender diversity in governance are based on the assumption that more diverse and inclusive boards and management may be more open to issues of accountability, transparency, and sustainability. Hasan et al. (2022) indicate that certain corporate governance characteristics, including gender diversity, are associated with a higher likelihood of publishing sustainability reports. Similarly, Cicchiello et al. (2021) show that greater female representation on boards may contribute to more intensive sustainability reporting, while Singhania et al. (2024) find that gender diversity may also be associated with the quality of sustainability reporting.

A particularly important aspect for this paper is the international dimension of governance, namely the presence of foreign members in governance bodies. The literature indicates that international experience and the presence of foreign board members may contribute to the adoption of more advanced corporate governance practices and a greater inclination toward sustainability disclosure. The study on the role of foreign boards and foreign ownership shows that the international component of the governance structure can positively affect the quality of sustainability disclosure, suggesting that companies with a stronger international element are often under greater pressure to develop a more formalized practice of non-financial reporting.

Recent literature has also shown a strong interest in the relationship between sustainability reporting and financial performance. Thus, Monteiro et al. (2024), Lehenchuk et al. (2023), Hussain et al. (2024), as well as Shaban and Zarnoun (2024), examine whether and to what extent sustainability reporting and ESG disclosure affect profitability, risk, and company performance. Although this group of studies is not directly focused on the question of what is related to the practice of reporting itself, it confirms that sustainability reporting has become a central topic of contemporary empirical research. A review of studies on the impact of sustainable practices on financial performance further supports this trend (Todorović et al., 2024).

3 Methodology of empirical research

The aim of this study is to examine whether there is an association between sustainability reporting practices and selected company characteristics, namely the gender of the company's director, the presence of foreign nationals in governance bodies, and the region in which the company was established.

The initial basis for the research was the list of the 100 most successful companies in 2023 according to operating revenues, published by the Serbian Business Registers Agency. However, during the process of data collection and processing, the initial sample was adjusted in accordance with the methodological requirements of the research. Companies that did not demonstrate continuity of operations during the previous five-year period, including newly established companies, were excluded from the analysis in order to ensure greater sample homogeneity and eliminate the potential effects arising from different stages of organizational development. In addition, a certain number of companies were excluded due to the unavailability of all necessary data required for conducting the analysis and calculating the observed indicators. This procedure was applied in order to ensure the completeness, comparability, and reliability of the dataset. However, it should be acknowledged that such a sample selection procedure may limit the generalizability of the findings, since the results refer only to the most successful companies with available and complete data. Consequently, the final sample consisted of 92 companies. Table 1 presents the structure of the observed sample according to whether a company prepares a sustainability report or not.

Table 1. Sustainability reporting practice in the sample

Variables	Frequency	Valid %
Prepares a sustainability report as a separate document	48	52.2
Does not prepare a sustainability report	44	47.8
Total	92	100.0

Source: Authors' calculation

Out of the total of 92 observed companies, 48 (52.2%) prepare a sustainability report, while 44 companies (47.8%) do not prepare such a report. Based on these data, it can be observed that the representation of both groups is relatively balanced, with a slight predominance of companies that prepare sustainability reports.

The obtained distribution indicates that the practice of preparing sustainability reports is present in slightly more than half of the most successful companies in the observed sample. This may suggest that sustainability reporting is becoming an increasingly important part of business practice among large and successful companies in Serbia, although it has not yet been universally adopted across all firms. At the same time, the fact that nearly half of the companies do not prepare sustainability reports indicates differences in the approach to non-financial reporting, which further justifies the need to examine whether these differences are related to the gender of the director, the presence of foreign nationals in governance bodies and the region in which the company was established.

Table 2 presents the structure of the observed sample according to the gender of the director.

Table 2. Distribution of companies by director gender

Variables	Frequency	Valid %
Male	76	82.6
Female	16	17.4
Total	92	100.0

Source: Authors' calculation

In 76 companies, the directors are male, representing 82.6% of the sample, while in 16 companies the director is female, accounting for 17.4% of the sample. This distribution indicates a pronounced dominance of men in director positions among the observed companies.

The obtained data indicate gender imbalance at the top of the governance structure of the most successful companies in Serbia, given that women are significantly underrepresented in director positions. The following table presents the structure of the sample according to whether there are individuals in the company's governance bodies who are not citizens of the Republic of Serbia, i.e., whether foreign nationals are present.

Table 3. Presence of foreign nationals in governance bodies

Variables	Frequency	Valid %
No foreign nationals in the governance structure	54	58.7
Presence of foreign nationals in the governance structure	38	41.3
Total	92	100.0

Source: Authors' calculation

In 54 companies (58.7%), there are no foreign nationals in the governance bodies, while in 38 companies (41.3%) there are individuals who are not citizens of the Republic of Serbia. This distribution shows that companies without foreign nationals in their governance structure are more numerous, but the share of companies with foreign members in governance bodies is also relatively high.

The obtained data indicate that the international dimension of the governance structure is present in a significant number of the most successful companies in Serbia. This may be the result of foreign ownership, links with international companies, or the inclusion of foreign experts in the management process. This is particularly important given the fact that companies that are part of a group are required to prepare sustainability reports, either as a standalone document or as part of the group report.

Previous studies suggest that gender diversity in governance structures may be associated with greater sensitivity to transparency, accountability, and sustainability issues. Research indicates that women in governance bodies may contribute to stronger sustainability orientation and more intensive sustainability reporting practices (Hasan et al., 2022; Cicchiello et al., 2021). In addition, gender diversity has also been linked to the quality of sustainability reporting (Singhania et al., 2024). Building on these findings, it is reasonable to examine whether director gender is associated with sustainability reporting practices in the context of Serbian companies.

H1: There is a statistically significant relationship between the gender of the director and the preparation of sustainability reports.

The international dimension of governance has increasingly attracted attention in sustainability reporting research. The presence of foreign nationals in governance bodies may indicate stronger international orientation, foreign ownership links, exposure to global standards, and more formalized corporate governance practices. Previous studies suggest that companies with international governance characteristics may be more inclined toward sustainability disclosure and transparency practices. Therefore, the presence of foreign nationals may represent a factor associated with sustainability reporting.

H2: There is a statistically significant relationship between the presence of foreign nationals in governance bodies and the preparation of sustainability reports.

Regional affiliation may reflect differences in institutional environment, concentration of economic activity, market exposure, access to professional support services, and proximity to administrative and financial centers. Previous studies emphasize the role of institutional context in shaping sustainability reporting practices (Khan et al., 2020; Bose & Khan, 2022). Consequently, it is reasonable to examine whether sustainability reporting practices differ across territorial groups within the Republic of Serbia.

H3: There is a statistically significant relationship between the region in which the company is established and the preparation of sustainability reports.

4 Research results and analysis

In order to examine whether the preparation of sustainability reports is associated with selected categorical characteristics of companies, the Pearson chi-square test of independence was applied. This test is appropriate as the study investigates the relationship between two categorical variables: a binary variable indicating whether a company prepares a sustainability report and categorical characteristics such as the gender of the director, the presence of foreign nationals in governance bodies and the region in which the company was established. In addition to statistical significance, the strength of the relationship was also assessed using the Phi coefficient and Cramer's V.

4.1 Director Gender and Sustainability Reporting

In this part of the empirical analysis, it was examined whether there is a statistically significant relationship between the gender of the company's director and the preparation of sustainability reports.

Table 4. Measures of Association between Director Gender and Sustainability Reporting

Measure	**Value**	**Approx. Sig.**
Phi	0,095	0,363
Cramer's V	0,095	0,363
N valid cases	92	

Source: Authors' calculation.

The obtained values of Phi and Cramer's V coefficients indicate a very weak association between the gender of the director and the preparation of sustainability reports. Since the p-value is greater than 0.05 ($p = 0.363$), this relationship is not statistically significant. Accordingly, based on the observed sample, it can be concluded that there are no statistically significant differences in sustainability reporting practices between companies led by male directors and those led by female directors. In addition, the assumptions for the application of the Chi-Square test are fully satisfied, as no cell has an expected frequency less than 5, and the minimum expected count is 7.65. Therefore, the results of the test can be considered statistically valid and reliable. The following table presents the distribution of companies by director gender and sustainability reporting practices.

Table 5. Association between director gender and sustainability reporting

Director Gender	Has a sustainability report	Does not have a sustainability report	Total
Male	38	38	76
Female	10	6	16
Total	**48**	**44**	**92**

Source: Authors' calculation.

Observed within the group of companies led by male directors, 38 out of 76 companies (50.0%) prepare sustainability reports, while the same number of companies

(50.0%) do not. This indicates that within this group, the relationship between companies that prepare and those that do not prepare sustainability reports is completely balanced.

Among companies led by female directors, 10 out of 16 companies (62.5%) prepare sustainability reports, while 6 companies (37.5%) do not. This means that within this group, there is a somewhat higher relative share of companies that prepare sustainability reports compared to companies led by male directors.

If the table is observed from the perspective of the total number of companies that prepare sustainability reports, it can be noted that the majority of these companies are those led by male directors. Specifically, out of the total of 48 companies that prepare reports, 38 (79.2%) have a male director, while 10 companies (20.8%) have a female director. A similar structure is observed among companies that do not prepare sustainability reports, where 86.4% are companies with male directors and 13.6% are companies with female directors.

Based on the descriptive analysis, it can be observed that there is a certain difference between the two groups, as the relative share of companies that prepare sustainability reports is somewhat higher among companies led by female directors. However, this difference is not sufficiently pronounced to conclude, based solely on the descriptive distribution, that there is a statistically significant relationship.

Although a slightly higher relative share of companies preparing sustainability reports was observed among those with female directors, the results of the statistical analysis do not provide sufficient evidence to confirm this difference as significant. Therefore, it can be concluded that the hypothesis regarding the existence of a statistically significant relationship between director gender and sustainability reporting is not supported in this study.

4.2 Presence of Foreign Nationals in Governance Bodies and Sustainability Reporting

In the next part of the empirical analysis, it was examined whether there is a statistically significant relationship between the presence of foreign nationals in companies' governance bodies and the preparation of sustainability reports.

Table 6. Measures of Association between the Presence of Foreign Nationals in Governance Bodies and Sustainability Reporting

Measure	Value	Approx. Sig.
Phi	0.229	0.028
Cramer's V	0.229	0.028
N of valid cases	92	

Source: Authors' calculation.

The obtained values of the Phi and Cramer's V coefficients indicate a weak association between the presence of foreign nationals in governance bodies and sustainability reporting. Since the p-value is less than 0.05 ($p = 0.028$), this relationship is statistically significant. This suggests that, based on the observed sample, there is a statistically significant association between the presence of foreign nationals in governance

structures and sustainability reporting practices among companies. It is important to note that the assumptions for the Pearson Chi-Square test are fully satisfied, as no cell has an expected frequency lower than 5, with the minimum expected count being 18.17. Therefore, the results can be considered methodologically reliable. Additionally, Fisher's Exact Test confirms the statistical significance of the association ($p = 0.035$), supporting the overall conclusion of a relationship between the observed variables.

Table 7. Association between foreign nationals in governance bodies and sustainability reporting

Foreign Nationals in Governance Bodies	Has a sustainability report	Does not have a sustainability report	Total
No foreign nationals	23	31	54
Presence of foreign nationals	25	13	38
Total	48	44	92

Source: Authors' calculation.

Observed within the group of companies where there are no foreign nationals in governance bodies, 23 out of 54 companies (42.59%) prepare sustainability reports, while 31 companies (57.41%) do not. In contrast, among companies whose governance bodies include foreign nationals, 25 out of 38 companies (65.79%) prepare sustainability reports, while 13 companies (34.21%) do not. This distribution indicates that the share of companies preparing sustainability reports is higher in the group of companies with foreign nationals in governance bodies.

If the table is observed from the perspective of the total number of companies that prepare sustainability reports, it can be noted that 23 out of the total of 48 such companies (47.9%) belong to the group without foreign nationals in governance bodies, while 25 companies (52.1%) belong to the group with foreign nationals. On the other hand, among companies that do not prepare sustainability reports, those without foreign nationals in governance bodies dominate, as 31 out of 44 companies without reports (70.5%) belong to this group, while only 29.5% are companies with foreign nationals.

The results indicate that companies with foreign nationals in their governance bodies more frequently prepare sustainability reports than companies without foreign members in their governance structure. Although the strength of this relationship is relatively weak, it is statistically significant, which may be connected with the conclusion that the hypothesis regarding the existence of a statistically significant relationship between the presence of foreign nationals in governance bodies and sustainability reporting is confirmed in this study.

4.3 Region of Company Establishment and Sustainability Reporting

In this part of the empirical analysis, it was examined whether there is a statistically significant relationship between the region in which a company is established and the preparation of sustainability reports. Since nominal variables are also involved in this

case, measures of association, namely Phi and Cramer's V, were applied to assess the strength of their relationship.

Table 8. Measures of Association between the Region of Company Establishment and Sustainability Reporting

Test	Value	Asymp. Sig.
Pearson Chi-Square	1.148	0.563
Likelihood Ratio	1.149	0.563
Linear-by-Linear Association	0.512	0.474
N of Valid Cases	92	

Source: Authors' calculation

The results of the Pearson Chi-Square test indicate that there is no statistically significant association between the region of company establishment and sustainability reporting ($\chi^2 = 1.148$, df = 2, p = 0.563). A similar conclusion is obtained from the Likelihood Ratio test, which also shows no statistically significant relationship between the observed variables. The Linear-by-Linear Association analysis does not indicate the presence of a statistically significant trend (p = 0.474), further confirming the absence of a systematic relationship between the examined categories. In addition, all assumptions for the application of the Chi-Square test are satisfied, as no cell has an expected frequency less than 5, with the minimum expected count being 7.65. Therefore, the results of the test are considered statistically valid. It can be concluded that, based on the observed sample (N = 92), there is no statistically significant association between the region of company establishment and sustainability reporting practices.

Table 9. Association between region and sustainability reporting

Region	Has a sustainability report	Does not have a sustainability report	Total
Belgrade	31	24	55
Vojvodina	9	12	21
Rest of Serbia	8	8	16
Total	**48**	**44**	**92**

Source: Authors' calculation.

Observed within each territorial group, the largest number of companies originates from Belgrade. In this group, 31 out of 55 companies (56.36%) prepare sustainability reports, while 24 companies (43.64%) do not. In Vojvodina, 9 out of 21 companies (42.86%) prepare sustainability reports, while 12 companies (57.14%) do not. In the group Rest of Serbia, the distribution is balanced, as 8 out of 16 companies (50.0%) prepare sustainability reports, while the remaining 8 companies (50.0%) do not.

From the perspective of companies that prepare sustainability reports, the majority originate from Belgrade, accounting for 64.6% of all reporting companies, followed by Vojvodina (18.8%) and Rest of Serbia (16.7%). A similar structure is observed among companies that do not prepare sustainability reports, where Belgrade also dominates (54.5%), followed by Vojvodina (27.3%) and Rest of Serbia (18.2%).

Although certain differences can be observed in the descriptive distribution, the results of the statistical analysis do not provide sufficient evidence to confirm a statistically significant relationship between regional affiliation and sustainability reporting. Therefore, Hypothesis H3 regarding the existence of a statistically significant relationship between the region of company establishment and sustainability reporting is not supported in this study.

5 Conclusion

In contemporary business conditions, sustainability reporting is increasingly moving beyond the scope of voluntary disclosure and becoming an important instrument of transparency, accountability, and communication with stakeholders. Under the influence of regulatory requirements, investor expectations, and the development of the sustainable business concept, the preparation of sustainability reports is increasingly becoming an integral part of modern corporate governance.

Building on this framework, the aim of this paper was to examine whether sustainability reporting practices are associated with certain company characteristics and governance structure features, namely the gender of the director, the presence of foreign nationals in governance bodies, and the region in which the company was established. The research was conducted on a sample of 92 of the most successful companies in the Republic of Serbia for the year 2023, using descriptive statistics and the Pearson chi-square test.

The results showed that slightly more than half of the observed companies prepare sustainability reports, indicating that the practice of non-financial reporting is present among a significant number of the most successful companies, but has still not been universally adopted. At the same time, the results confirmed the existence of a statistically significant association between the presence of foreign nationals in governance bodies and the preparation of sustainability reports, while director gender and regional affiliation were not confirmed as statistically significant factors.

The theoretical contribution of the study lies in providing empirical evidence from the context of the Republic of Serbia as an insufficiently explored transitional economy environment, as well as in deepening the understanding of the institutional and governance determinants of sustainability reporting. From a practical perspective, the findings suggest that sustainability reporting should not be viewed solely as a regulatory obligation, but also as an important element of modern corporate governance, transparency, and the strategic orientation of companies.

It is particularly important that the findings indicate that a significant number of companies do not prepare sustainability reports, despite the fact that non-financial reporting is normatively regulated and represents a legal obligation for certain categories of business entities in the Republic of Serbia. In this regard, the obtained results may serve as a signal for legislators and regulatory institutions regarding the need to strengthen mechanisms for monitoring the implementation of regulations, institutional support, and the education of business entities.

The limitations of the study relate to the fact that the sample includes only the most successful companies according to operating revenues, while the sample selection was additionally conditioned by continuity of operations and data availability. Therefore, the results cannot be fully generalized to all companies in the Republic of Serbia. Future research could include a larger number of companies, a longer time period, and additional variables such as ownership structure, profitability, board size, and the international connectedness of companies.

References

1. Bose, S., Khan, H.Z.: Sustainable development goals (SDGs) reporting and the role of country-level institutional factors: An international evidence. Journal of Cleaner Production 335, 130290 (2022)
2. Cicchiello, A.F., Kazemikhasragh, A., Monferrà, S., Girón, A.: Gender diversity on corporate boards: How Asian and African women contribute on sustainability reporting activity. Gender in Management: An International Journal 36(7/8), 801–820 (2021). https://doi.org/10.1108/GM-05-2020-0147
3. Dienes, D., Sassen, R., Fischer, J.: What are the drivers of sustainability reporting? A systematic review. Sustainability Accounting, Management and Policy Journal 7(2), 154–189 (2016). https://doi.org/10.1108/SAMPJ-08-2014-0050
4. Elkington, J.: Cannibals with Forks: The Triple Bottom Line of 21st Century Business. Capstone, Oxford (1997)
5. European Commission: The Commission adopts the European Sustainability Reporting Standards. Directorate-General for Financial Stability, Financial Services and Capital Markets Union (2023). https://finance.ec.europa.eu/news/commission-adopts-european-sustainability-reporting-standards-2023-07-31_en
6. Freeman, R.E.: Strategic Management: A Stakeholder Approach. Pitman, Boston (1984)
7. Gray, R., Kouhy, R., Lavers, S.: Corporate social and environmental reporting: A review of the literature and a longitudinal study of UK disclosure. Accounting, Auditing & Accountability Journal 8(2), 47–77 (1995). https://doi.org/10.1108/09513579510146996
8. Hahn, R., Kühnen, M.: Determinants of sustainability reporting: A review of results, trends, theory, and opportunities in an expanding field of research. Journal of Cleaner Production 59, 5–21 (2013). https://doi.org/10.1016/j.jclepro.2013.07.005
9. Hasan, A., Hussainey, K., Aly, D.: Determinants of sustainability reporting decision: Evidence from Pakistan. Journal of Sustainable Finance & Investment 12(1), 214–237 (2022). https://doi.org/10.1080/20430795.2021.1964813
10. Hussain, M.A., Alsayegh, M.F., Boshnak, H.A.: The impact of environmental, social, and governance disclosure on the performance of Saudi Arabian companies: Evidence from the top 100 non-financial companies listed on Tadawul. Sustainability 16(17), 7660 (2024). https://doi.org/10.3390/su16177660
11. Khan, M., Lockhart, J., Bathurst, R.: A multi-level institutional perspective of corporate social responsibility reporting: A mixed-method study. Journal of Cleaner Production 265, 121739 (2020). https://doi.org/10.1016/j.jclepro.2020.121739
12. Lehenchuk, S., Zhyhlei, I., Ivashko, O., Gliszczyński, G.: The impact of sustainability reporting on financial performance: Evidence from Turkish FBT and TCL sectors. Sustainability 15(20), 14707 (2023). https://doi.org/10.3390/su152014707

13. Ministarstvo finansija Republike Srbije: Smernice za nefinansijsko izveštavanje u Srbiji. Ministarstvo finansija Republike Srbije, Beograd (2025). https://www.mfin.gov.rs/upload/media/kQoEF9_691ed576a394b.pdf
14. Monteiro, S., Roque, V., Faria, M.: Does sustainability reporting impact financial performance? Evidence from the largest Portuguese companies. Sustainability 16(15), 6448 (2024). https://doi.org/10.3390/su16156448
15. Ortas, E., Álvarez, I., Jaussaud, J., Garayar, A.: The impact of institutional and social context on corporate environmental, social and governance performance of companies committed to voluntary corporate social responsibility initiatives. Journal of Cleaner Production 108, 673–684 (2015). https://doi.org/10.1016/j.jclepro.2015.06.089
16. Patten, D.M.: Exposure, legitimacy, and social disclosure. Journal of Accounting and Public Policy 10(4), 297–308 (1991)
17. Reverte, C.: Determinants of corporate social responsibility disclosure ratings by Spanish listed firms. Journal of Business Ethics 88(2), 351–366 (2009). https://doi.org/10.1007/s10551-008-9968-9
18. Shaban, O.S., Zarnoun, R.S.: Impact of sustainability reporting on financial performance and risks: Evidence from the emerging market. Risk Governance and Control: Financial Markets & Institutions 14(4), 96–109 (2024). https://doi.org/10.22495/rgcv14i4p10
19. Singhania, S., Singh, J., Aggrawal, D., Rana, S.: Board gender diversity and sustainability reporting quality: A generalized ordered logit approach. Kybernetes 53(8), 2679–2699 (2024). https://doi.org/10.1108/K-07-2022-0963
20. Todorović, T., Pavlović, A., Makivić, R., Savković, M.: The impact of sustainability practices on financial performance: A review of empirical studies. In: Forging the Future: Pioneering Approaches in Business, Management and Economics Engineering to Overcome Emerging Global Challenges, pp. 566–572 (2024). https://doi.org/10.24867/FUTUREBME-2024-063
21. Zakon o računovodstvu. Službeni glasnik RS, br. 73/2019 i 44/2021 – dr. zakon (2019)

AI-use declaration: During the preparation and revision of this manuscript, AI-based tools were used for language editing, text refinement, and assistance in structuring responses to reviewers. The authors reviewed, verified, and take full responsibility for the final content of the manuscript.

Original Research Article *https://doi.org/10.5281/zenodo.23475007*

To be Resilient, Committed or Branded: Exploring Predictors of Sustainable Career

Ivana Kovačević[1][0000-0002-7601-1000], Denisa Abrudan[2] [0000-0001-8310-7011], Mateja Manojlović[3][0000-0002-8856-2264], and Achilleas Anagnostopoulos[4] [0000-0003-3390-4498][4].

[1] University of Belgrade, Faculty of Organizational Sciences, Serbia; East-European Center for Research in Economics and Business, Faculty of Economics and Business Administration, West University of Timisoara, Romania
[2]West University of Timisoara, Faculty of Economics and Business Administration, Romania
[3] University of Belgrade, Faculty of Organizational Sciences, Serbia
[4] University of Thessaly, Business Administration, Greece

Abstract. In the context of social and economic changes of work environment, partly incited by technological development and scarcity of resources, contemporary career paths are increasingly shaped by employability and sustainability paradigms. Sustainable career emerged as the concept covering the phenomenon of managing one`s professional success in challenging circumstances while preserving well-being, satisfaction and positive social impact. This concept might be seen as part of the wider career ecosystem raising the interest for its correlates and predictors. The research idea was to examine the possibility to predict employee`s sustainable career based on his resilience, personal branding virtue and different aspects of his organizational commitment. In order to test this hypothetical model we administrated Employee Resilience Scale, Personal Branding Scale, Organizational Commitment Scale and Sustainable Career Scale to 177 participants that have at least one year of work experience in a corporate environment. Results indicate that it is possible to construct a predictive model of sustainable career based on linear combination of scores that explains approximately 40.2% of total variance and with all predictors included, except the Organizational Commitment that did not prove to be statistical significant. Explained variance becomes slightly higher (43.7%) when Organizational Commitment score is decomposed to Affective and Continuance Commitment, that were statistically significant predictors and on Normative Commitment that did not reach statistical significance. Results imply that hypotheses are partially confirmed, with low to moderate positive associations between Sustainable career and other examined concepts, except with Normative Commitment.

Keywords: Sustainable Career, Employee Resilience, Personal Branding, Organizational Commitment

1 Introduction

1.1 Theoretical framework

Defining career path in contemporary world of work becomes genuinely challenging endeavor. As in the previous period employment could cover the majority of work-related domain problems, today the concept of employability comes into the focus. It implies career contingencies based on turbulent, non-predictable work relations re-framing the context of work, redefining job roles, imposing new, sometimes unpredictable, competency requirements.

Sustainable career becomes the universal concept when explaining current vocational scope and scholars and practitioners are trying to identify predictors, enablers and consequences of this approach. This paper is in line with researches who explore sustainable career from the holistic perspective. It incorporates individual behavioral assets, like employee resilience, strategic behaviors of personal branding that aligns with new technological breach, as well as the relationship with organization, through commitment to the current work setting.

New workforce setting. Scholars have identified several disruptive factors that are reshaping the concept of contemporary careers. These include socioeconomic changes driven by rapid digitalization [1], particularly the introduction of artificial intelligence AI [2], [3] which has accelerated the automation of work processes [4], as well as the globalization of labor markets [5]. Consequentially, the employment arrangements are being replaced by alternative and sometimes precarious forms of work [4]. These developments contributed to greater career mobility and insecurity [6], more fluid career boundaries [7], and changing perceptions of career success [8].

At the same time, the long-standing prioritization of economic gain over environmental preservation has contributed to resource scarcity and increased concern about the impact of organizational practices on employees` physical and mental well-being. These challenges have stimulated the development of sustainable HRM strategies [7] and "Decent work agenda" initiatives [9], [10]. Further, increased longevity of population, leads towards the extension of work life and aging workforce, making careers longer and potentially more demanding for individuals [5].

Sustainable career. In this context, sustainable career, as new vocational construct, emerged, relying on some previous attempts to describe the new career phenomenon (protean, boudaryless, kaleidoscope, customized, post-corporate, see [12]). Following the "tradition" of Newman`s, three personal resources sustainable career - renewability, flexibility and integrity [11], Van der Heijden et al. [13], gave definition later expanded and operationalized by Russo et al. [9]. As there is a strong link between ESG paradigm, sustainability, sustainable HRM and sustainable career, one empirical approach assess the sustainable career readiness [14]. According to the ESG (environmental, social and governmental) framework the sustainable career awareness, background values and its operability in everyday HRM practice were measured.

Sustainable career meant employees` renewability based on life balance, flexibility and adaptability as the prerequisites of resiliency along the professional span, and integration through meaningful perspective [11]. Aligning with previous conceptions, it re-conceptualized dimensions of time and space incorporated with personal agency and meaning. This definition explains sustainable careers as *'sequences of career experiences reflected through a variety of patterns of continuity over time, thereby crossing several social spaces, characterized by individual agency, herewith providing meaning to the individual*' [13] (p 7). Russo et al. [15] started from the pillars of this definition: health, happiness and productivity [7] and expand it with the indicator of social empowerment [9]. Authors were lead by the idea to combine personal with contextual aspects when defining some career sustainable.

The later conception is wider and offer possibility to detach from the individual career perspective and incorporate all relevant stakeholders in career studies. It refers to the idea of new career ecosystem that assumes career-related development opportunities based on interaction between person, organizations and societal forces [3]. In "career ecosystem" we incorporate concepts of employee resilience, personal branding and organizational commitment, as potential predictors of sustainable career.

Employee resilience. Resilience is well known phenomenon in studying stress and burnout syndrome [16], [17], but authors recently started defining it as domain-specific and developmental individual resource. Hartmann et al. [18] systematized approaches of resilience on trait-like, state-like and process conceptions. Contrary to trait-like conceptions, state-like sees resilience as developable potential. Process perspective offer more operable definition, seeing it as the contingent responses leading toward positive adaptation. It gives resilience temporal dimension and posits it into the contextual frame, allowing it to be defined according to the specific domains [19].

While workplace becomes arena of pervasive and constant challenges, studying employee resilience gain on popularity. Also, the importance of constant need for upgrading capacities and competencies, the learning ability becomes synonym for adaptability. Authors insist on steady positive learning from adverse circumstances [20], and mistakes [21] in order to develop capacities, well-being, as well as the organizational resources [22]. Scholars often connect personal resilience at work with organizational resilience, seeing their relation reciprocal [21].

Kuntz at al. offered workplace specific, process definition derived from positive psychology: capability of using resources to continually adapt and flourish at work, even when faced with challenging circumstances [21]. Thus, employee resilience is adaptive behavior but dependent on organizational context, describing how much is employee skillful in using organizational resources to achieve personal aspirations while contributing to organizational goals. Adequate organizational context is often pictured in the terms of sustainable HRM practices, inspiring researches that combines them [23], [24]. Naswall et al. [25] developed and validate instrument to assess behaviors reflecting identification and utilization of available resources along with reflective learning and change oriented behavior [20].

Personal branding. As employee resilience covers leveraging social networks [21], and utilizing available resources [26], and sustainable careers assume strategic approach to career management [27], personal branding may be chosen strategy to en-

hance career sustainability. In order to stay employable in contemporary professional landscape employees have to embrace marketing strategies when defining themselves [28]. Although some of the basic ideas that lines in the conception of self-presenting is not novel, there are some new features that makes person more adaptable to present labor market circumstances [29].

As Gorbatov et al. [30] noticed, the concept is viable in the context of project based and temporary employment work arrangements. Further, it becomes popular as the tool to present oneself trough social media. Authors believe that AI utilization brings insecurity to employees if they do not learn how to use their benefits. They advocate that this situation might bring competitive advantage to those who learn to use opportunities of social media to promote their professional image [31]. Personal branding enables employee to stand out from competition thanks to well-developed self-promoting strategies [32].

Generally, personal branding is seen as the personal marketing activity that strategically creates, position and maintains desired positive impression. It presents selected personal and professional qualities to a target audience qualities and promises potential benefits through a strategically constructed narrative and image to signalize their value [30]. Authors identified three aspects of personal branding, strategic, differential and technologically supported [32]. Strategic is based on proactive approach to the perception others have, differential aspect leans on adjustment to targeted audience and all that with very meticulously and proficiently used technological tools. As it was found that personal branding are linked with positive career outcomes [33], financial benefits [34] and more career opportunities [35], it is relevant for contemporary career.

Organizational commitment. One`s career is never independent of social or organizational context. Also, sustainable career does not always imply changing organizations and professional transitions within multiple employments. It could be realized within one organization. Scholars believe that offering appropriate developmental program that foster career growth to employees result in enhanced organizational commitment and diminish turnover intentions [36]. Based on social exchange theory [37], organizational commitment reflects the employee-organization relationship and the psychological factors that determine whether employees choose to stay or leave the organization [38]. Commitment is determined by personal factors and expectations incorporated in psychological contract with organization [39]. Both external and internal determinants shape commitment, like the potential conflict between professional and organizational bonding, which may arise when individuals must choose between pursuing career development and advancing within a specific organization hierarchy [40].

Although workplace commitment in practice went through transformation from the organization as the target [41], to more localized targets (workgroups) [42], or professions in temporary short-term and flexible work engagements [43], the issue stays important. Ali and Swart [42] say it can encompass diverse work entities, from coworkers and supervisors to clients and teams or even be target-free [37]. It is in line with contemporary work setting and boundaryless organizational structures that favors more conscious and lucrative bonding, than emotional attachments to organization.

Nevertheless, theory recommends Three-Component Model (TCM) [44] of commitment, differentiating affective, continuance and normative one. These different commitment domains differ according to the nature of attachment to organization and also according to the consequences they produce. Affective commitment is based on need satisfaction of employees and their reciprocal behavior. Normative commitment creates organizational loyalty through moral obligation based usually on trust at organization and on feeling obligation toward organization, usually due to the specific employees` experience or mindset. Continuance commitment is based on calculating costs of leaving organization and depends on the incentives and benefits that organization provide compared to invested effort and alternatives.

1.2 Research problem

Sustainable careers are achieved in dynamic and an unpredictable circumstances that requires wider picture to determine what factors affects them. De Vos et al. [1] and Donald [3] followed the idea of broader career ecosystem encompassing internal personal domain, as well as the external circumstances. Although the concept of career is extensively studied and theoretically well established with different methodological approaches, there are few measures of sustainable career as such.

Operationalization of the sustainable career may be based on ESG framework that measures sustainable career perception through sustainability awareness, commitment to sustainable values, and the integration of sustainability principles into the HRM practices. According to this approach, it comprises nine dimensions reflecting flexibility, open-minded growth mindset of employee who has competencies for hybrid forms of work [14]. In this paper career sustainability is seen according to the Russo et al. [9], who extended the version of De Vos et al. [7]: satisfaction and meaning one career offer (happiness), well-being and balance based on match between capacities and stress (health) and perceiving good performance and employability (productivity) while contributing to wider society (social empowerment).

Career paths are inherently idiosyncratic and depend on individual strengths, such as employee resilience. Since career sustainability reflects an individual`s ability to adapt to changing work contexts [7], resilience may play a key role in its development. Research suggests that resilience is associated with high levels of well-being and more positive work-related outcomes because resilient employees are better equipped to cope with and overcome workplace challenges [3]. Resilience caries renewal potential and opportunities that are seen as the base of sustainable career [4]. Newman [11] claims that resilience is the consequence of embracing sustainable career, while other authors see resilience as the element of it [4]. In one study authors found that the relation with resilient behavior at work is complex and dependent on occupation [46], while in the other sustainable career perceptions are formed for high resilience employees [47].

Scholars claim that significant way to contribute to sustainable career is personal agency seen in constantly developing professional networks [48]. It is competency to present one`s strengths to wider audience. So, in order to keep career sustainable employees need to develop strategies and to approach to their vocational life proactively using new, technologically mediated ways of expressing. One tactic is seen in personal branding as the form of proactive work behavior. It means borrowing marketing

strategies to promote self professionally, and doing it strategically, targeted (differentiated) and technology based [32]. Proactive orientation and personal agency are two strongest links between sustainable career and this concept.

Finally, sustainable career is not strictly personal but has to align to context [49] that enables employees to satisfy their needs while working on organizational goals [13], with possible outcomes, like organizational commitment [50]. Organizational commitment is expected to be connected with career from the perspective of work setting that provides employees to develop their competencies and stay healthy. Researchers found positive relationship between organizational fostering of resilient capacities and career adaptability through organizational commitment developed by sustainable initiatives [51].

The main goal of our research was to explore potential predictors of sustainable career within the behavioral resources (employee resilience) and strategies (personal branding), as well as the perception of organizational context as worth to commit to (organizational commitment). By extending the predictors of sustainable career from individual trait-like and socially contextualized strategies, we wanted to contribute to the validation of the Russo`s measurement concept along with idea to make possible predictions of one`s career and potential employability [6], [9]. Further, predictors might enable scholars and practitioners to identify personal means of creating career success in contemporary circumstances of workforce aging [2], [5], [11], as well as on reliance at AI [2]. They could foster the development of these potentials among employees and workforce in general [27]. From the organizational point of view, linking with the organizational commitment, might point toward means to foster employee well-being [11], through job crafting [50] and redefining the commitment concept itself [43].

2 Methods

2.1 Hypotheses

H1. It is possible to construct a predictive model of sustainable career based on a linear combination of employee resilience, personal branding, and different facets of organizational commitment scores.

H2. Employee resilience is a positive predictor of sustainable career.

H3. Personal branding is a positive predictor of sustainable career.

H4. Different dimensions of organizational commitment will demonstrate differential relationships with sustainable career.

H4a. Affective organizational commitment will be a positive predictor of sustainable career.

H4b. Continuance organizational commitment will be a negative predictor of sustainable career.

H4c. Normative organizational commitment will be a negative predictor of sustainable career.

2.2 Sample

The sample in this study consisted of 177 participants. Eligibility criteria required that respondents have at least one year of work experience in a corporate environment. The sample comprised 68.4% female and 32.6% male participants. Participants' ages ranged from 20 to 60 years, with a mean age of M = 34.9 years and a standard deviation of SD = 10.86 years, indicating a relatively broad age distribution within the sample. Participants belongs to three different labour market contexts (Romanian 37.29 %, Greek – 38.42%, and Serbian 24.29 %).

2.3 Measures

Participants were administered a battery of tests comprising a total of 63 items derived from four distinct scales: the Employee Resilience Scale [25], the Personal Branding Scale [32], the Organizational Commitment Scale [44], and the Sustainable Career Scale [6]. For all measures, scores were computed by averaging respondents' answers across all items belonging to the respective target scale, thereby yielding composite scale scores for subsequent analyses.

Sustainable Career Scale. Sustainable Career Scale is self-report five-point Likert type scale. It comprises of 12 items along one dimension (3 reversed). Recent validity study [6] Cronbach's alpha reliability 0.89. On our sample registries reliability was of Cronbach's alpha 0.781.

Employee Resilience Scale. This instrument is a seven-point Likert-type self-report scale. It is unidimensional and consists of nine items. None of the items are reverse-coded. The scale assesses employee resilience. Validation studies conducted on three independent samples demonstrated high reliability, with Cronbach's alpha coefficients ranging between 0.88 and 0.93 [46]. In the present sample, reliability was slightly lower, with a Cronbach's alpha of 0.832.

Personal Branding Scale. This instrument is a five-point Likert-type self-report scale consisting of 18 items distributed across three dimensions: strategic, differentiated, and technologically savvy personal branding. Each subscale is represented by six items, ensuring balanced coverage of all three facets of the construct. The authors [32] recommend the use of both composite and subscale scores in empirical research, depending on the level of analytical detail required. In the original validation study, Cronbach's alpha for the composite score ranged between 0.87 and 0.90, while reliability estimates for the subscales were reported as 0.77–0.82 for strategic, 0.73–0.76 for differentiated, and 0.83–0.88 for technologically savvy personal branding. In the present study, only the composite score was utilized, which demonstrated acceptable internal consistency, with a Cronbach's alpha of 0.954.

Organizational Commitment Scale. This instrument is a seven-point Likert-type self-report scale consisting of 24 items organized into three dimensions: affective commitment, continuance commitment, and normative commitment. Each of the three subscales contains eight items. The sum of 8 items are reverse coded. The scale enables calculation of both a composite Organizational Commitment score and separate scores for each dimension, allowing for a more differentiated analysis of the construct. Although the original authors did not report reliability for the composite score, subsequent studies have suggested acceptable reliability levels ranging from 0.80 to 0.85 [47]. Allen and Meyer [44] reported Cronbach's alpha coefficients of 0.87 for affective commitment, 0.75 for continuance commitment, and 0.79 for normative commitment. In the present study, reliability coefficients indicated a Cronbach's alpha of 0.830 for the composite score, 0.737 for affective commitment, 0.837 for continuance commitment, and 0.691 for normative commitment, suggesting acceptable to good internal consistency across most dimensions.

2.4 Design and procedure

The design of this study is correlational–regression. The predictor variables included employee resilience, organizational commitment (both composite and its subscales), and personal branding, while the criterion variable was sustainable career. Two regression models were estimated: the first included employee resilience, composite organizational commitment, and personal branding as predictors; the second included employee resilience, affective, continuance, and normative commitment, along with personal branding. The survey was administered online, in English, Greek, or Serbian, depending on participants' preference. The questionnaires were completed in the following fixed order: demographic variables, sustainable career, personal branding, organizational commitment, and finally employee resilience.

3 Results

3.1 Descriptive and metric characteristics of the used scales

The data were analyzed using IBM SPSS Statistics, applying descriptive statistical analyses, the calculation of the Cronbach's alpha coefficient, as well as multiple regression analysis. The results of the descriptive and metric characteristics of the instruments will be presented first, followed by the results of the regression analysis.

Table 1 presents Cronbach's alpha coefficients, mean values, standard deviations, standardized skewness and kurtosis, as well as the range for each of the scales used. Inspection of the table indicates that all instruments demonstrate satisfactory reliability ($\alpha \geq .70$). The Sustainable Career Scale shows acceptable reliability, while the Employee Resilience Scale and Organizational Commitment Scale exhibit high reliability, and the Personal Branding Scale demonstrates excellent reliability. The symmetry of the distributions for all variables corresponds to that of a normal distribution, as indicated by the values of standardized skewness. With regard to kurtosis, the standardized kurtosis values suggest that the peakedness of the distributions approxi-

mates that of a normal distribution for all variables, with the exception of Organizational Commitment. In the case of this variable, a statistically significant leptokurtic distribution is observed, indicating that respondents' scores are more tightly clustered around the mean, resulting in reduced variability of the score distribution.

Table 1. Descriptive and metric characteristics of the applied scales

Variable	α	M	SD	zSK	zKU	Min	Max
Employee Resilience	0.832	5.794	0.771	-1.683	-1.074	3.670	7.000
Personal Branding	0.954	3.984	1.262	1.120	-1.468	1.280	7.000
Organizational Commitment	0.830	4.144	0.798	-0.885	3.080	1.420	6.670
Sustainable Career	0.781	4.402	0.726	0.650	0.127	2.580	6.330

3.2 Models of sustainable career score prediction

To examine the possibility of predicting the Sustainable Career score based on Employee Resilience, Organizational Commitment, and Personal Branding scores, a multiple linear regression analysis was conducted, with Employee Resilience, Organizational Commitment, and Personal Branding serving as predictor variables, and Sustainable Career score as the criterion variable. The tested model explains approximately 40.2% of the variance in Sustainable Career score ($R^2_{adj} = 0.402$), and is statistically significant ($F(3,173) = 40.381$, $p < .001$). Among the predictors included in the model, Employee Resilience and Personal Branding emerged as significant predictors, whereas Organizational Commitment did not reach the required level of statistical significance. Both significant predictors showed a positive relationship with the criterion variable, indicating that higher levels of Employee Resilience, as well as higher levels of Personal Branding, are associated with higher Sustainable Career scores.

Table 2. Structure of the prediction model

Variable	β	t	p	r_0	$r_{partial}$	r_{part}
Employee Resilience	0.277	4.539	<0.001	0.382	0.326	0.265
Organizational Commitment	0.024	0.388	0.699	0.195	0.029	0.023
Personal Branding	0.521	8.678	<0.001	0.578	0.551	0.506

Given that the authors of the Organizational Commitment Scale indicate that both a composite score and individual subscale scores can be meaningfully employed [45], an additional analysis was conducted to further explore this possibility. As a matter of

fact, scholars usually agree upon using these three scales separately implying the multifaceted but overlapping nature of the concept [39]. Specifically, we examined the extent to which the Sustainable Career score could be predicted based on a combination of Employee Resilience and Personal Branding scores, together with the individual dimensions of the Organizational Commitment Scale. This approach allowed for a more nuanced understanding of the contribution of specific facets of organizational commitment. Table 3 presents the descriptive and metric indicators of quality, as well as the results obtained on the subscales of the Organizational Commitment Scale.

Table 3. Descriptive and metric characteristics of the applied OC subscales

Variable	α	M	SD	zSK	zKU	Min	Max
Affective OC	0.737	4.533	1.048	-2.563	0.905	1.250	6.750
Continuance OC	0.837	4.138	1.263	0.142	-1.773	1.130	7.000
Normative OC	0.691	3.760	0.949	-0.348	0.792	1.000	6.250

The results indicate that this extended prediction model is also statistically significant, suggesting that the specified set of predictors can meaningfully account for variance in the Sustainable Career score ($F(5,171) = 28.375$, $p < .001$). In this case, the model explains 43.7% of the total variance ($R^2adj = 0.437$), representing a modest improvement compared to the previous model. Within this framework, Employee Resilience, Personal Branding, Affective Organizational Commitment, and Continuance Organizational Commitment emerged as statistically significant predictors, whereas Normative Organizational Commitment did not reach statistical significance. The pattern of results points to low to moderate positive associations between Sustainable Career and Employee Resilience, Personal Branding, and Affective Organizational Commitment. At the same time, a weak negative relationship was observed between the criterion variable and Continuance Organizational Commitment. These findings suggest differential roles of commitment dimensions in career sustainability. The detailed results are presented in Table 4.

Table 4. Structure of the prediction model

Variable	β	t	p	r_0	$r_{partial}$	r_{part}
Employee resilience	0.245	4.082	0.000	0.382	0.298	0.231
Personal branding	0.495	8.223	0.000	0.578	0.532	0.465
Affective OC	0.169	2.378	0.019	0.285	0.179	0.134
Continuance OC	-0.170	-2.854	0.005	-0.090	-0.213	-0.161
Normative OC	0.033	0.481	0.631	0.298	0.037	0.027

4 Discussion

Starting from the career ecosystem perspective [3] and Russo`s et al. [6] proposed model of measuring career sustainability that encompass individual and contextual aspects, we wanted to explore sustainable career further. Although there are many theoretical discussions and few methodological solutions for studying the concept [6], [14], we see it as still relatively neglected filed. Further, sustainability paradigm gain on popularity and influence globally and its impact is obvious in the work setting. Labor market perspectives are changing and challenging employees not only to upgrade but also to alter their professional profiles through reskilling programs to achieve more sustainable careers [52]. The aim of this study was to test the predictive model of sustainable career with concepts that we found that might be closely related to it. We identified three concepts of interest based on the content of the sustainable career: dynamic and adaptable, proactive venture that does not include only surmount adversities, but also to endure the whole life span, preserving well being and contributing to organizational goals and social welfare [1], [5], [7], [9].

Employee resilience is frequently connected with sustainability through organizational practices [24] and with sustainable career [46], [47] through adaptability [11]. Second, personal branding is seen as the strategic way to present self and to express agency. It means actively being involved in creating employability image through raising chances to be noticed and employed [28], [32]. Finally, we included organizational commitment consisted of three, sometimes overlapping, sometimes disparate dimensions: affective, continuance and normative [37], [44].

Predicting sustainable career. Results showed that most of our hypotheses were confirmed. The tested model is significant and explains about 40% of the variance of the sustainable career. This suggests that sustainable career can be predicted with moderate based on the included variables, particularly when employees express resilient behavior in the organization and are able to manage their professional image through social media. With organizational commitment situation is more complicated because of the discrepancy between different aspects of commitment. Only affective commitment contributes positive in predicting, while continuance is negatively related and normative has no effect. If we incorporate affective commitment and continuance, we see that model becomes slightly more accurate (only 43.7% of variance is predicted). So, employee who is emotionally connected and identified with organization probably has more sustainable career contrary to the one who stays in organization for lucrative reasons. It is in line with the idea that sustainable career revolves not only around productivity but also incorporate happiness and well-being [6], [9]. If we feel that our needs are satisfied in organizational setting we would probably see our career more sustainable. Normative commitment is shown to be irrelevant in prediction, which is not unexpected due to its connection with other personal dimensions out of the sustainability scope [37], [38].

The findings of this study indicate that sustainable career is more strongly related to individual agency than to general organizational attachment. In particular, employ-

ee resilience and personal branding, as adaptive and proactive characteristics, appear to have an important role in shaping sustainable career outcomes. At the same time, organizational commitment does not seem to operate as a unified predictor, as its different dimensions demonstrate distinct relationships with sustainable career. While affective commitment is positively related to sustainable career, continuance commitment shows a negative relationship, whereas normative commitment does not appear to be statistically significant. These results suggest that sustainable career is more closely associated with meaningful engagement and adaptability, rather than with obligation or cost-based attachment to the organization.

Several studies have addressed the relationship between sustainable career and employee resilience [46], [47], while other studies pointed to an indirect link between these concepts (see [5]). Also, some studies found links among predictors, like correlation between employee resilience and organizational commitment [53]. Personal branding is found to be indirectly linked with sustainable career through sustainable HRM framework [54]. We should point to the novelty gain in the study that connect these concepts together and put them in the role of career predictors.

Limitations of the study. This relatively ambitious trial to develop predicting model showed limited potential. Although metrics of the scales are satisfying, there are discrepancies in the level of theoretically and methodologically established concepts that are used as predictors. Also, the main concept of sustainable career is in its pioneering stage. Further, the socio-demographical profile of the sample is rather poor, so there is lack of context for deeper understanding the results. Sample is relatively small and heterogeneous in terms of ethnicity, cultural and economical background, as well as in terms of occupation and years of tenure and work experiences.

The main challenge of the paper is seen in its cultural heterogeneity. There is no common stand among scholars about generalizability what sustainable careers mean in different socioeconomic contexts, referring to local circumstances and differences among countries. Although sample comprises of respondents from different economies (Romanian, Greek, and Serbian), we did not systematically considered these potential differences for the purpose of this paper.

Although individual career studied longitudinally, by grasping the specific features of all career stages [12], might give better insight into the nature of the concept and predictors, we used cross-sectional design. This might be one of the limitations for deeper understanding of the mechanisms that effects career to be shaped as sustainable. Nevertheless, we believe that we can take into consideration obtained results and insights that might be valuable for inspiring further researches.

Implications for practice and further studies. At the level of individual career, it might be advisable to focus on insightful learning and grasping every opportunity to develop some new skills, especially if they involve technology. It should cover the strategically development skills of creating, managing and upgrading professional image through social media. And finally, find the workplace that is worth to commit and identify with, in order to have sustainable career.

From HRM and organizational perspective, there is an advice to develop sustainable HRM practices (like [24]) in order to create a resilient career system that can incorporate needs and bring welfare to all stakeholders involved. De Vos et al. [1] see it

as the organizational flexibility to respond to dynamics of market thanks to adaptable workforce, as the concept of shared responsibility. Based on this particular research, some guidelines should be given to scholars and practitioners.

Dedication toward HRM sustainability, might be a good path toward supporting sustainable career model. Yet, approaches focused on organizational context provoking commitment, might not be as fruitful as fostering employees` individual strategies of developing personal strengths. Personal branding fosters one`s career but it is also the way to fit the organizational goals [54]. Further, employee resilience as the personal psychological resource is relatively corroborated concept contributing to sustainable career [4], [7], [11], [46], [47]. So, it might be expected from HRM policies to focus on developing personal strengths and skills to shape employees career through trainings and to support their employability strategies.

Theoretical implications are revolving around the sustainable criteria of happiness, health, productivity and social empowerment. In previous studies these elements are considered in relation with individual employee resources, organizational, social, job related and more general contextual antecedents as well as in line with the individual and organizational outcomes [6]. [9], [15]. Novel in this predictive model are chosen variables to represent three complementary domains. While employee resilience is personal resource of protecting well-being and happiness while ensuring productivity [19 - 20], [23], [24] personal branding has role of proactive strategy for ensuring social visibility and personal control [32 - 35]. Organizational commitment signifies the quality of the organization-person relationship, and act as the potential to predict all four aspects, especially satisfaction and social fit [36], [39], [40], [51]. Further, all these variables are in line with sustainable career framework that recognizes the interactional nature of person, place and time redefinition in new work context [1].

Further research should incorporate the perception of employability [8] and proactive personality that is found to be the strongest predictor of networking [55] and career adaptability through actively seeking and using future opportunities and resources [27]. It might be also interesting to investigate the role of social belongingness, or social connectedness and trust in colleagues along with sustainable career [9].

Since researchers often pose the question of possibility to see career concept as universal regardless of the economical and geopolitical circumstances [3], in further researches local predictive models should be examined. Also, sustainable career research body could benefit from longitudinal approach in further studies.

Conclusion. Although the proposed model, which included employee resilience, personal branding, and organizational commitment as predictors of sustainable career, demonstrated limited predictive power, it provides valuable insights into the relationships among several important concepts relevant to contemporary employment. Specifically, the model integrates personal psychological resources, strategic career management behaviors, and employee-organization relations, all of which may contribute to the development of healthy, meaningful, productive, and socially responsible careers.

The findings suggest that individual resources and career management strategies have greater predictive potential for sustainable careers than organizational factors. Nevertheless, employees` emotional attachment to and identification with organization, emerged as significant predictor, whereas remaining in an organization primarily

for material reasons were the opposite. This findings highlights the importance of fostering employees` psychological resources and career management competencies through sustainable HRM practices, while also recognizing the role of positive employee-organization relationships, and looking for more predictive organizational and relational variables in supporting career sustainability.

The results further indicate that no single set of variables can fully explain career sustainability. Given the complexity and dynamic nature of contemporary careers, sustainable career is likely shaped by the interplay of multiple individual, organizational and contextual influences. Despite the limitations mentioned above, the present study contributes to the theoretical and practical understanding of employee resilience, personal branding, and organizational commitment within a single predictive framework. And of course, it contributes to defining the concept of sustainable career and its practical implications. Also, it should be noted that, according to our knowledge, these constructs have not previously been investigated together in relation to sustainable career.

Future research should incorporate additional personal, relational, and organizational variables in order to develop a more comprehensive understanding of sustainable career development and provide scholars and practitioners with more robust evidence-based recommendations.

AI Disclaimer: Artificial intelligence tools were used solely for language editing and proofreading purposes. No AI-generated content, analysis, or interpretations were included in the manuscript.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. De Vos, A.: Enhancing the sustainability of careers in disruptive times. Journal of the National Institute for Career Education and Counselling 53(1), 8–17 (2024). https://doi.org/10.20856/jnicec.5302
2. Bankins, S., Jooss, S., Restubog, S.L.D., Marrone, M., Ocampo, A.C., Shoss, M.: Navigating career stages in the age of artificial intelligence: A systematic interdisciplinary review and agenda for future research. Journal of Vocational Behavior 153, 104011 (2024). https://doi.org/10.1016/j.jvb.2024.104011
3. Donald, W.E., Van der Heijden, B.I., Baruch, Y.: Introducing a sustainable career ecosystem: theoretical perspectives, conceptualization, and future research agenda. Journal of Vocational Behavior 151, 103989 (2024). https://doi.org/10.1016/j.jvb.2024.103989
4. McDonald, K.S., Hite, L.M.: Conceptualizing and creating sustainable careers. Human Resource Development Review 17(4), 349–372 (2018). https://doi.org/10.1177/1534484318796318
5. Rapuano, V.: Toward sustainable careers: Literature review. Contemporary Research on Organization Management and Administration 8(1), 41–54 (2020). doi.org/10.33605/croma-012020-004
6. Russo, A., Zammitti, A., Zarbo, R., Magnano, P., Santisi, G.: The sustainable career scale (SCS): development, validity, reliability and invariance. Journal of Global Responsibility 17(1), 1–25 (2026). https://doi.org/10.1108/JGR-07-2024-0127

7. De Vos A, Van der Heijden BIJM, Akkermans J (2020) Sustainable careers: Towards a conceptual model. Journal of Vocational Behavior 117:103196. https://doi.org/10.1016/j.jvb.2018.06.011
8. Schweitzer, L., Lyons, S., Smith, C.J.: Career sustainability: Framing the past to adapt in the present for a sustainable future. Sustainability 15(15). Article 11800. https://doi.org/10.3390/su151511800
9. Russo, A., Zammitti, A., Ginevra, M.C., Santisi, G.: Sustainable career and decent work as mediators between the sense of community responsibility and well-being. BMC Psychology 13(1), 1192 (2025). https://doi.org/10.1186/s40359-025-03172-3
10. Hopner, V., Carr, S.C., Wloch, J.: Sustainable careers within greening economies. Australian Journal of Career Development 33(3), 286–293 (2024). https://doi.org/10.1177/10384162241285536
11. Newman, K.L.: Sustainable careers: Lifecycle engagement in work. Organizational Dynamics 40(2), 136–143 (2011). doi:10.1016/j.orgdyn.2011.01.008
12. Van der Heijden, B., De Vos, A., Akkermans, J., Spurk, D., Semeijn, J., Van der Velde, M., Fugate, M.: Sustainable careers across the lifespan: Moving the field forward. Journal of Vocational Behavior 117, 103344 (2020). https://doi.org/10.1016/j.jvb.2019.103344
13. Van der Heijden, B.I., De Vos, A.: Sustainable careers: Introductory chapter. In: Handbook of Research on Sustainable Careers, 1–19. Edward Elgar Publishing (2015). https://doi.org/10.4337/9781782547037.00006
14. Kovačević, I., Abrudan, D., Labrović, J.A.: The HR managers' perception of future career paths in the context of ESG perspective: Readiness for sustainable careers. In: Exploring ESG Challenges and Opportunities: Navigating Towards a Better Future, vol. 116, 145–173. Emerald Publishing Limited (2024). https://doi.org/10.1108/S1569-375920240000116009
15. Russo, A., Valls-Figuera, R.G., Zammitti, A., Magnano, P.: Redefining 'careers' and 'sustainable careers': A qualitative study with university students. Sustainability 15(24), 16723 (2023). https://doi.org/10.3390/su152416723
16. Norouzinia R, Ebadi A, Ferdosi M, Masoumi G, Tayebi Z, Yarmohammadian MH (2020) A systematic review of psychometric properties of workplace resilience measurement scales. TPM: Testing, Psychometrics, Methodology in Applied Psychology 27(2):251–269. https://doi.org/10.4473/TPM27.2.6
17. Aloba, O., Olabisi, O., Aloba, T.: The 10-item Connor–Davidson Resilience Scale: Factorial structure, reliability, validity, and correlates among student nurses in southwestern Nigeria. Journal of the American Psychiatric Nurses Association 22(1), 43–51 (2016). https://doi.org/10.1177/1078390316629971
18. Hartmann, S., Weiss, M., Newman, A., Hoegl, M.: Resilience in the workplace: A multilevel review and synthesis. Applied Psychology 69(3), 913–959 (2020). https://doi.org/10.1111/apps.12191
19. Fisher, D.M., Ragsdale, J.M., Fisher, E.C.: The importance of definitional and temporal issues in the study of resilience. Applied Psychology 68(4), 583–620 (2019). https://doi.org/10.1111/apps.12162
20. Caniëls, M.C., Baaten, S.M.: How a learning-oriented organizational climate is linked to different proactive behaviors: The role of employee resilience. Social Indicators Research 143(2), 561–577 (2019). https://doi.org/10.1007/s11205-018-1996-y
21. Kuntz, J.R., Malinen, S., Näswall, K.: Employee resilience: Directions for resilience development. Consulting Psychology Journal: Practice and Research 69(3), 223 (2017). https://psycnet.apa.org/doi/10.1037/cpb0000097
22. Tonkin, K., Malinen, S., Näswall, K., Kuntz, J.C.: Building employee resilience through wellbeing in organizations. Human Resource Development Quarterly 29(2), 107–124 (2018). https://doi.org/10.1002/hrdq.21306

23. Shao, D., Ji, Y., Zhou, R., Liu, T.: Unpacking the mixed influences of sustainable human resource management on employee resilience: effects of challenge–hindrance appraisal and trait mindfulness. Asia Pacific Journal of Human Resources 63(1), e12433 (2025). https://doi.org/10.1111/1744-7941.12433
24. Trunk Širca, N., Riaz, Z., Hamid, S., Žurauskė, G., Stankevičiutė, Ž.: The interplay among employee resilience, sustainable HRM and work engagement. Human Systems Management 43(4), 573–587 (2024). https://doi.org/10.3233/HSM-230204
25. Näswall, K., Malinen, S., Kuntz, J., Hodliffe, M.: Employee resilience: Development and validation of a measure. Journal of Managerial Psychology 34(5), 353–367 (2019). https://doi.org/10.1108/JMP-02-2018-0102
26. Kuntz, J., Connell, P., Näswall, K.: Workplace resources and employee resilience: The role of regulatory profiles. Career Development International 22(4), 419–435 (2017). https://doi.org/10.1108/CDI-11-2016-0208
27. Zhao, L., Li, W., Zhang, H.: Career adaptability as a strategy to improve sustainable employment: a proactive personality perspective. Sustainability 14(19), 12889 (2022). https://doi.org/10.3390/su141912889
28. Gorbatov, S., Khapova, S.N., Oostrom, J.K., Lysova, E.I.: Personal brand equity: Scale development and validation. Personnel Psychology 74(3), 505–542 (2021). https://doi.org/10.1111/peps.12412
29. Scheidt, S., Gelhard, C., Henseler, J.: Old practice, but young research field: A systematic bibliographic review of personal branding. Frontiers in Psychology 11, 1809 (2020). https://doi.org/10.3389/fpsyg.2020.01809
30. Gorbatov, S., Khapova, S.N., Lysova, E.I.: Personal branding: Interdisciplinary systematic review and research agenda. Frontiers in Psychology 9, 2238 (2018). https://doi.org/10.3389/fpsyg.2018.02238
31. Huhta-Koivisto T (2025) Crafting the digital self: Expert personal branding on social media in the age of AI. Master's thesis, Aalto University School of Business, Helsinki, Finland.
32. Gorbatov, S., Khapova, S.N., Lysova, E.I.: Get noticed to get ahead: The impact of personal branding on career success. Frontiers in Psychology 10, 495947 (2019). https://doi.org/10.3389/fpsyg.2019.02662
33. Gandini, A.: Digital work: Self-branding and social capital in the freelance knowledge economy. Marketing Theory 16(1), 123–141 (2016). https://doi.org/10.1177/1470593115607942
34. Rangarajan, D., Gelb, B.D., Vandaveer, A.: Strategic personal branding—And how it pays off. Business Horizons 60(5), 657–666 (2017). https://doi.org/10.1016/j.bushor.2017.05.009
35. Minor-Cooley, D., Parks-Yancy, R.: The power of the brand: personal branding and its effect on job seeking attributes. Journal of Internet Commerce 19(3), 241–261 (2020). https://doi.org/10.1080/15332861.2020.1777028
36. Al Balushi, A.K., Thumiki, V.R.R., Nawaz, N., Jurcic, A., Gajenderan, V.: Role of organizational commitment in career growth and turnover intention in public sector of Oman. PLOS ONE 17(5), e0265535 (2022). https://doi.org/10.1371/journal.pone.0265535
37. Klein, H.J., Brinsfield, C.T., Cooper, J.T.: The experience of commitment in the contemporary workplace: An exploratory reexamination of commitment model antecedents. Human Resource Management 60(6), 885–902 (2021). https://doi.org/10.1002/hrm.22040
38. Jaros, S., Culpepper, R.A.: Some unresolved issues in the person-centered approach to organizational commitment within the Meyer/Allen paradigm. Management Research Review 48(6), 849–862 (2025). https://doi.org/10.1108/MRR-11-2022-0747
39. Kayar, S., Yeşilada, T.: Quartet of sustainable job security, job performance, organizational commitment, and motivation in an emerging economy: Focusing on Northern Cyprus. Sustainability 16(16), 6764 (2024). https://doi.org/10.3390/su16166764

40. Meyer, J.P., Morin, A.J., McInerney, D.M.: Multi-target commitment profiles and their links to motivational orientation, turnover intentions, and personal adjustment: An investigation with Hong Kong teachers. Current Psychology 44(4), 2594–2611 (2025). https://doi.org/10.1007/s12144-024-07235-3
41. van Rossenberg, Y.G.T., Klein, H.J., Asplund, K., Bentein, K., Breitsohl, H., Cohen, A., Yalabik, Z.Y.: The future of workplace commitment: Key questions and directions. European Journal of Work and Organizational Psychology 27(2), 153–167 (2018). https://doi.org/10.1080/1359432X.2018.1443914
42. Ali, N., Swart, J.: "It is all about me"–an investigation of interns' workplace commitment. Personnel Review 54(7), 1794–1809 (2025). https://doi.org/10.1108/PR-03-2024-0209
43. Van Rossenberg, Y.G., Swart, J.A., Yalabik, Z.Y., Cross, D., Kinnie, N.: Commitment systems in cross-boundary work. British Journal of Management 34(2), 746–766 (2023). https://doi.org/10.1111/1467-8551.12610
44. Allen, N.J., Meyer, J.P.: The measurement and antecedents of affective, continuance and normative commitment to the organization. Journal of Occupational Psychology 63(1), 1–18 (1990). https://doi.org/10.1111/j.2044-8325.1990.tb00506.x
45. Näswall, K., Kuntz, J., Hodliffe, M., Malinen, S.: Employee resilience scale (EmpRes) measurement properties. Resilient Organizations Research Programme: Christchurch, New Zealand (2015). URL: https://www.resorgs.org.nz/wpcontent/uploads/2019/08/ Resilient_Organisations_2015-04_Employee_Resilience_Scale.pdf
46. Kovačević, I., Bétourné, N., Stankevičiūtė, Ž., Màthé, I., Anagnostopoulos, A., Abrudan, D., Anđelković Labrović, J.: Employee resilience and career perception across countries and occupations. In: ITAIS 2025 Conference, Castellanza, Italy, 17–18 October 2025 (2025).
47. Kovačević, I., Abrudan, D., Anagnostopoulos, A., Anđelković Labrović, J.: Career sustainability perception of the resilient employees: Contributions to regenerative work setting. In: 12th Responsible Management Education Research Conference 2025, Belgrade, Serbia, 21st–23rd October 2025, Book of Extended Abstracts, 260–265 (2025). ISBN 978-86-7680-501-3
48. Akkermans, J., Hirschi, A.: Career proactivity: Conceptual and theoretical reflections. Applied Psychology 72(1), 199–204 (2023). https://doi.org/10.1111/apps.12444
49. Chudzikowski, K., Gustafsson, S., Tams, S.: Constructing alignment for sustainable careers: Insights from the career narratives of management consultants. Journal of Vocational Behavior 117, 103312 (2020). https://doi.org/10.1016/j.jvb.2019.05.009
50. Valcour, M.: Facilitating the crafting of sustainable careers in organizations. In: Handbook of Research on Sustainable Careers, 20–34 (2015). https://doi.org/10.4337/9781782547037.00007
51. Mujajati, E., Ferreira, N., Du Plessis, M.: Fostering organisational commitment: A resilience framework for private-sector organisations in South Africa. Frontiers in Psychology 15, 1303866 (2024). https://doi.org/10.3389/fpsyg.2024.1303866
52. Obradović, V., Kovačević, I., Kužet, I., Manojlović, M.: The sustainability of reskilling projects based on employees' readiness for a career shift: Pursuing sustainable careers by transitioning into IT professions. Sustainability 16(2), 709 (2024). https://doi.org/10.3390/su16020709
53. Youssef, C.M., Luthans, F.: Positive organizational behavior in the workplace: The impact of hope, optimism, and resilience. Journal of Management 33(5), 774–800 (2007). https://doi.org/10.1177/0149206307305562
54. Szántó, P., Papp-Váry, Á.: Sustainable personal branding: Aligning individual identity with organizational goals in human resource management. J. Cult. Anal. Soc. Change 10(4), 5417–5428 (2025). https://doi.org/10.64753/jcasc.v10i4.4577

55. Pearson, J.L., Schneer, J.A., Van der Heijden, B.I., De Vos, A., Steele, C., Kadiyil, P.R.: The role of agentic disposition and orientation on the career competencies of graduating students: A four-country study. Career Development International 30(5), 532–545 (2025). https://doi.org/10.1108/CDI-11-2024-0469

THEMATIC SECTION

Data and Business Analytics

Resilient and Intelligent Organizations
in Times of Permanent Disruption

SYMORG 2026 • CONFERENCE PROCEEDINGS

Original Research Article *https://doi.org/10.5281/zenodo.23475038*

A two-level framework for joint optimization of feature engineering and hyperparameter tuning in predictive maintenance classification

Martin Guyon[1][0009-0002-2369-4909]

[1] ISAE-SUPAERO, Toulouse, Occitanie, France

Abstract. This study presents a structured framework to optimize time-series classification in predictive maintenance using the C-MAPSS dataset. A two-level approach jointly tunes feature engineering and model hyperparameters via GridSearchCV, combined with group-based validation to prevent data leakage. To mitigate class imbalance, a trajectory-based stochastic sampling strategy prioritizes near-failure observations. The framework also analyzes the train–test distribution mismatch caused by truncated test trajectories. Results show that sliding windows are the main driver of performance, while feature extraction provides modest but more stable gains. The approach achieves up to 0.89 test F1-score, with SVM demonstrating the best performance for the data at hand.

Keywords: Predictive Maintenance, Classification, C-MAPSS dataset.

1 Introduction

Traditional maintenance strategies, such as corrective maintenance — reacting after failure, and preventive maintenance — servicing at fixed intervals, often lead to either unexpected failures or unnecessary interventions. In contrast, predictive maintenance (PdM) leverages data-driven techniques to anticipate failures before they occur, enabling more efficient, condition-based decision-making that reduces costs and improves system reliability. A widely studied benchmark in this domain is the NASA C-MAPSS dataset, traditionally used for estimating Remaining Useful Life (RUL) as a regression problem. However, in practical applications, maintenance decisions are often binary in nature, requiring clear indications of whether a system is approaching failure within a given time horizon. This has motivated a growing interest in reformulating RUL prediction as a classification task.

Despite promising results, time-series classification for predictive maintenance presents several challenges. These include the effective representation of temporal dynamics, the selection of optimal feature engineering strategies, the handling of severe class imbalance due to long healthy operating periods, and the mismatch between training and test data distributions caused by full and truncated trajectories, respectively. Addressing these issues is essential for developing robust and reliable PdM models.

In this study, we propose a structured framework for optimizing time-series classification performance on the C-MAPSS FD001 dataset. Our approach combines sliding window-based feature engineering with a two-level hyperparameter optimization

strategy using GridSearchCV and group-based validation to prevent data leakage. In the machine learning community, the proposed procedure corresponds to a two-stage optimization or nested optimization [17]. Additionally, we introduce a trajectory-aware stochastic sampling method to handle class imbalance and we investigate the train–test distribution mismatch. Through a series of controlled experiments, we aim to identify the key factors that drive performance and provide practical insights for building effective predictive maintenance classification systems.

2 Related Work

The NASA C-MAPSS dataset has served as a standard benchmark for remaining useful life (RUL) prediction, predominantly formulated as a continuous regression problem. However, the operational requirements of predictive maintenance (PdM) often favor threshold-based decision-making. Recent literature has increasingly explored reframing RUL prediction into a binary classification task. For instance, studies have demonstrated that predicting failure within a fixed cycle horizon yields highly interpretable maintenance alerts, with classical models like gradient-boosted trees and deep learning models showing strong classification performance [1], [2], [16]. Similarly, [4] highlighted that segmenting the continuous degradation process into discrete health stages allows classification algorithms to more effectively identify near-failure instances and handle operational data constraints. [15] further investigated the trade-offs of these predictive architectures, confirming that while classification is highly effective for operational alerts, model performance remains extremely sensitive to hyperparameter tuning and the specific selection of temporal sliding window sizes.

To capture the temporal dynamics of engine degradation, sliding window techniques have become standard practice. [5] and [8] utilized sliding windows coupled with attention mechanisms to process sequential data, significantly improving prediction stability over single-cycle observations. Furthermore, [7] evaluated the impact of varying sliding window sizes by converting time-series into 2D sequence images, mirroring our objective of mapping optimal window sizes. While recent trends lean towards deep feature extraction using autoencoders [6] or complex hierarchical transformers [14], explicitly engineered features remain highly efficient and computationally viable. [3] proposed linear frameworks to partition the measurement space based on hyperplane boundaries, and [11] emphasized that rigorous preprocessing, such as removing constant sensor features to reduce redundancy, is critical for creating robust temporal health indicators.

The combination of windowing and classification inherently introduces severe class imbalance, as redundant healthy early-life data vastly outnumbers critical near-failure events. Traditional random sampling often fails to capture the specific context of mechanical degradation. To address this, [9] utilized piecewise linear labeling combined with Weibull distributions to manage early-life data, while [10] employed k-means clustering to safely drop redundant "before fault" data without losing predictive accuracy. Our trajectory-based stochastic sampling strategy builds upon these concepts by

probabilistically prioritizing the decision boundary, reducing early-life redundancy without entirely discarding it.

Finally, a persistent but often overlooked challenge in C-MAPSS is the distribution mismatch between complete training trajectories and truncated test trajectories. [12] explicitly tackled this distributional shift using Deep Domain Adaptation techniques to align source and target domains. Other works have explored alternative, highly complex architectures, such as Hybrid Quantum Neural Networks [13], to improve model generalization. In contrast, our approach aims to isolate and assess this train-test mismatch by implementing a strict test-based validation strategy within a GridSearchCV framework.

This work introduces novel contributions. Most previous studies on the C-MAPSS dataset focus either on deep learning models for RUL prediction or on specific aspects of predictive maintenance classification taken separately. In contrast, this work proposes a unified framework that studies several methodological components, including feature engineering, temporal windowing, sampling strategies, and validation procedures.

Moreover, recent works such as [5], [8], and [14] mainly rely on complex architectures such as attention-based LSTMs or transformers to improve prediction performance. In this study, the objective is different. Rather than introducing a new deep learning architecture, the focus is placed on systematically optimizing classical machine learning models through a two-level GridSearchCV-based framework combining feature extraction and hyperparameter tuning.

In particular, the proposed framework:

- introduces a trajectory-aware stochastic sampling strategy designed for predictive maintenance trajectories;
- assesses the train–test distribution mismatch caused by truncated test trajectories in C-MAPSS;
- employs strict group-based validation using engine identifiers to prevent temporal leakage between folds;
- provides a mapping between prediction horizons, window sizes, and optimal machine learning pipelines.

Overall, the goal of this work is not only to achieve competitive classification performance, but also to provide a structured and interpretable experimental framework to analyze how temporal context, feature engineering, sampling, and validation strategies might influence predictive maintenance classification results.

3 Methodology

3.1 Overview of the Proposed Approach

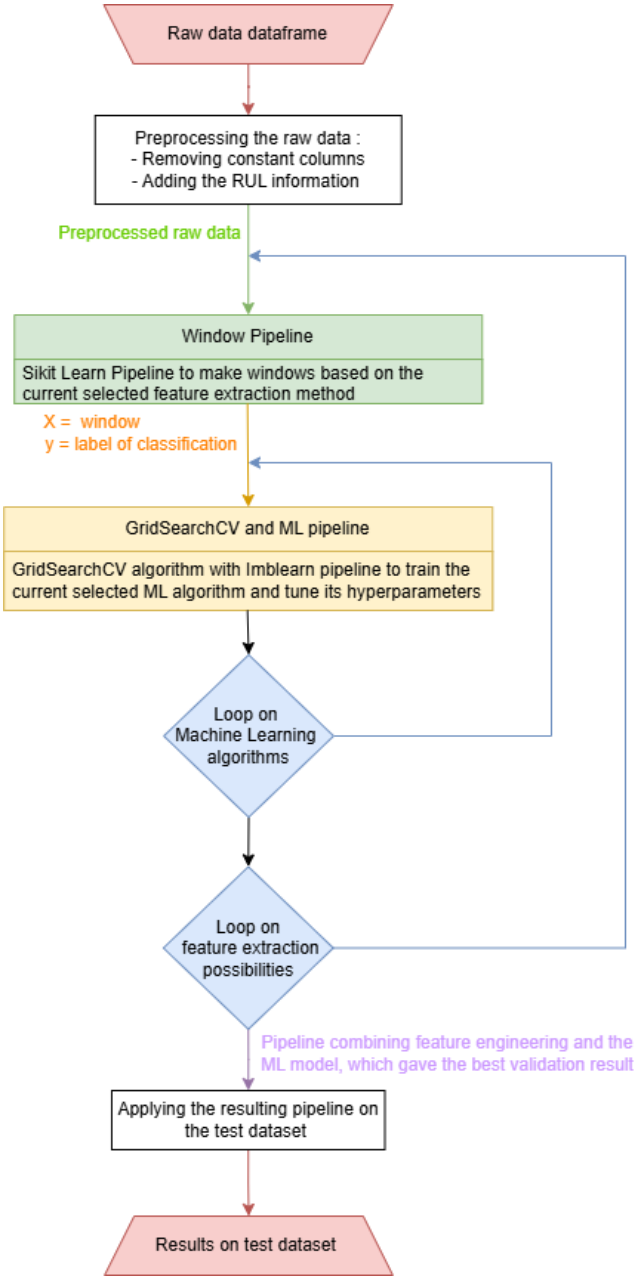


Fig. 1. Two-level optimization architecture for feature engineering and ML model selection

This study proposes a structured methodology for optimizing predictive performance in industrial time-series classification by combining feature engineering and machine learning model selection within a controlled experimental framework. The approach is based on a two-level optimization strategy, as shown in Fig. 1. A first loop is dedicated to feature engineering, where raw time-series data are transformed into window-based representations using tunable parameters, and a second loop relies on GridSearchCV to optimize machine learning architectures and their corresponding hyperparameters.

3.2 Dataset Description and Preprocessing

The experiments are conducted on the C-MAPSS dataset, which consists of four sub-datasets: FD001, FD002, FD003, and FD004. This study focuses on the FD001 subset.

The training dataset contains 100 engines, each represented by its full degradation trajectory from initial operation to failure. For each engine, the Remaining Useful Life (RUL) at cycle t is computed as:

$$RUL(t) = T_{\text{failure}} - t \tag{1}$$

where T_{failure} is the maximum cycle of the engine, here corresponding to the last cycle observed in the trajectory. Constant sensor features are removed to reduce redundancy and improve model efficiency.

The test dataset contains 100 different engines with truncated trajectories at $t_{lastcycle}$, meaning that the failure point is not directly observed. The true $RUL_{lastcycle}$ at the maximum cycle is provided separately. The RUL is computed in the same way as for the train dataset (1), where T_{failure} equals the sum of $RUL_{lastcycle}$ and $t_{lastcycle}$.

The same preprocessing steps applied to the training data, removal of constant features and addition of the RUL column, are consistently applied to the test data.

The corresponding RUL to each cycle of an engine is shown in Fig. 2 in both the case of complete trajectories in the train dataset, and truncated trajectories in the test dataset.

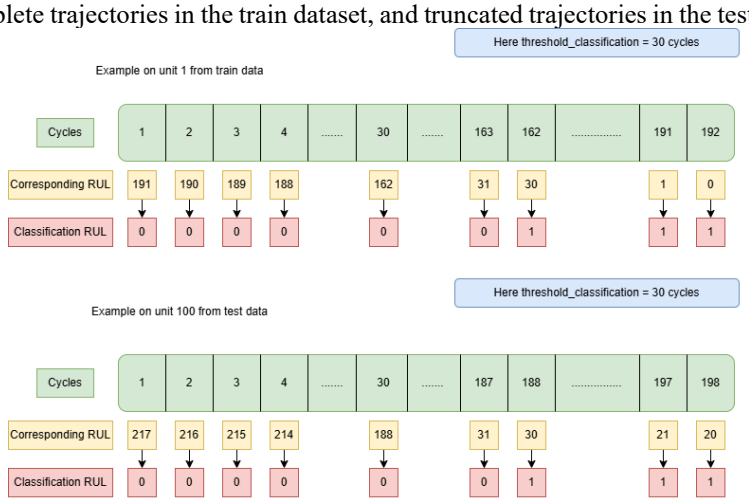


Fig. 2. Transformation of RUL regression into a binary classification task

3.3 Window-Based Feature Engineering

3.3.1 Binary Target Transformation. The RUL regression problem is transformed into a binary classification task, Fig. 2. For each cycle:

$$Classification\ RUL = \begin{cases} 1 & if\ RUL \leq threshold_classification \\ 0 & otherwise \end{cases} \quad (2)$$

The hyperparameter threshold_classification defines the prediction horizon.

3.3.2 Sliding Window Construction. A sliding window approach is applied independently to each engine trajectory, as explained in Fig. 3. Each window contains a fixed number of cycles defined by window_size and moves forward one cycle at a time. The number of windows generated for a given engine i is therefore:

$$N_i = T_i - W + 1 \quad (3)$$

where T_i is the trajectory length of engine i and W is the window size.
The total number of windows of a fixed size on all the different trajectories is :

$$N = \sum_i (T_i - W + 1) \quad (4)$$

3.3.3 Feature Extraction. For each window, a set of features is extracted from sensor data, as shown in Fig.3. Those features are of different types, such as :

- Statistical Features : Mean, Standard deviation, Minimum, Maximum, Median, Quartiles, Interquartile range.
- Dynamic Features : a polynomial regression of a certain degree p is fitted for each sensor s on the relative time $t \in [1, window_size]$:

$$s(t) \approx \sum_{k=0}^{p} a_k\, t^k \quad (5)$$

  Only the coefficients a_k corresponding to degree higher than 0, so with k > 0, are kept.

A key aspect of this methodology is to combine static statistical descriptors with dynamic temporal representations.

3.3.4 Feature Vector Construction. All extracted features are concatenated into a feature vector X. The target label y is defined as the binary RUL value at the last cycle of the window. Fig. 3 shows in detail how the process of selecting the windows, extracting the features and making the input data for classifier models is done. Additionally, the engine_id information is preserved to enforce group-based validation in the gridsearchCV.

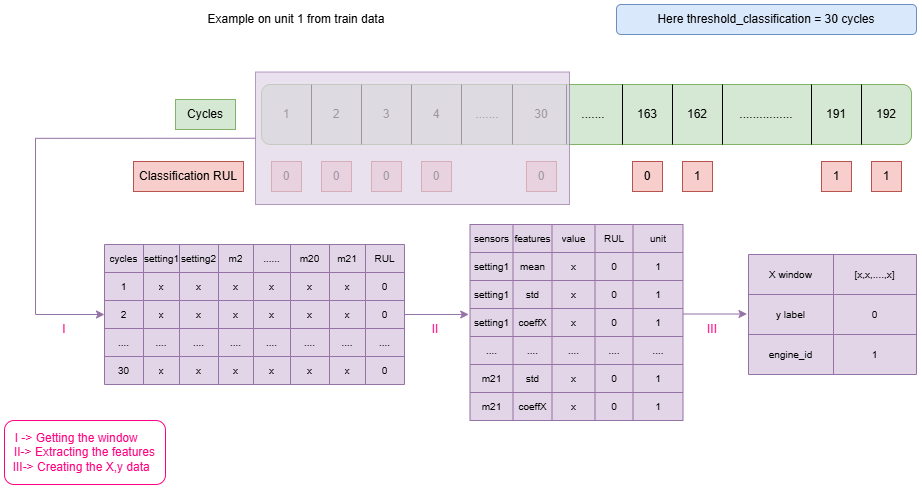


Fig. 3. Sliding window generation and feature extraction process

3.4 Machine Learning Pipeline and Hyperparameter Optimization

A machine learning pipeline that applies scaling, sampling, and a classification model is used within the GridSearchCV algorithm. GridSearchCV takes as input the feature vectors X, the labels y, and the corresponding engine IDs to which each vector belongs. A group-based cross-validation strategy based on engine IDs is employed within the GridSearchCV framework to prevent data leakage caused by windows originating from the same engine. In this way, windows from the same engine are always assigned to the same fold, ensuring that no information from a given trajectory appears in both the training and validation sets.

3.4.1 Feature Scaling. Several standard scikit-learn scaling techniques are considered, such as Min-Max scaling, Standard scaling, and Robust scaling. The scaling transformation is fitted on the training folds and subsequently applied to the validation folds within the cross-validation process. The resulting transformed data, represented as numerical arrays, can then be directly used for model training.

3.4.2 Mitigating Class Imbalance. The windowing strategy inherently leads to a highly imbalanced dataset, as the prediction horizon from 20 up to 40 cycles, is much smaller than the average trajectory length of 206 cycles. As a result, most windows are labeled as negative 0, with only a small proportion of positive samples 1.

Standard techniques such as Random Undersampling and Oversampling are commonly used but remain context-agnostic, treating all samples equally. In RUL prediction, however, early-life negative samples are far less informative than those near the failure boundary. Such methods risk removing critical near-boundary samples while preserving redundant early data.

To address this, we propose a trajectory-based stochastic sampling strategy that prioritizes windows based on their proximity to the decision boundary, focusing learning on the transition from healthy to degraded states while reducing overall imbalance.

For cycles prior to t_{start} (6), windows are sampled according to the probability defined in (7):

$$t_{start} = max(1, \lfloor t_{max} - K * threshold_classification \rfloor) \quad (6)$$

$$P_{keeping_window}(t) = \frac{1}{t_{start} - t + 1} \quad (7)$$

where t_{max} is the maximum cycle reached at failure, and K is positive a constant.

This approach achieves two objectives. First, it aims to reduce redundancy in early-life data by sparsely sampling early regions. Second, it creates a data density gradient in which the sampling rate increases as it approaches the failure threshold and reaches 100% for $t \geq t_{start}$, thereby improving discrimination at the decision boundary.

Moreover, a method with constant sampling probability p for $t < t_{start}$ is proposed.

However, these two approaches cannot be integrated with standard GridSearchCV, as stochastic sampling must be applied only to training data, preventing the use of automatic K-fold cross validation. In these setups, the traditional scikit-learn samplers are disabled to evaluate the intrinsic effectiveness of the proposed methods.

3.5 Experimental Design

For the experiments, unless explicitly stated otherwise, as in Study 2 in Section 4.1.2 and in Study 3 in Section 4.1.3, a group-based cross-validation strategy is employed within the GridSearchCV framework using three folds. Grouping is performed using the engine_id associated with each generated window, Fig.3. The scoring parameter in GridSearchCV is set to "f1". For the experiments, a fixed random seed of 42 is used to ensure reproducibility throughout the framework and the hyperparameters explored for each model are presented in Table 1, while all other associated hyperparameters are kept at their default values.

Table 1. Model hyperparameter search space

MLP	*max_iter : [50,100,200]* *hidden_layer_sizes : [(100,), (128,128)]* *learning_rate_init : [0.001]* *alpha : [0.1,0.01]*
RandomForest	*n_estimators : [100,200,500]* *max_depth : [None, 10]*
XGBoost	*n_estimators : [100,200]* *max_depth : [3,6]* *learning_rate : [0.01,0.1]* *colsample_bytree : [0.7, 1]* *subsample : [0.7]*
SVM	*C = [0.1, 1, 10]* *kernel = ['poly', 'rbf', 'linear']*

The sampling strategies used are RandomUnderSampler with a sampling strategy of 0.5 and RandomOverSampler with a sampling strategy of 1. The experiments are implemented in Python using the XGBoost, scikit-learn, and imbalanced-learn libraries.

3.5.1 Core Methodological Experiments: These experiments aim to address the issues outlined in the introduction. Three main methodological aspects are investigated.
Study 1: Impact of Temporal Windowing and Feature Extraction. This study compares the performance of models trained on raw time-series data and on engineered features derived from sliding windows, for a prediction in a 30 cycles horizon. The raw setting corresponds to window sizes of 1 and 30, whereas feature extraction is applied to windows of size 30. The goal is to assess the impact of using multi-cycle windows and the additional contribution of feature extraction for predictions.
Study 2: Evaluation of Sampling Strategies. This study, as described in Section 3.4.2 compares the proposed trajectory-based sampling method with standard imbalanced learning techniques. Four configurations are evaluated: (i) no sampling, (ii) scikit-learn implementations of RandomOverSampler and RandomUnderSampler, (iii) stochastic decay sampling (7), and (iv) constant early-life sampling. The goal is to assess the impact of addressing class imbalance and whether focusing on the transition zone near the decision boundary improves classification performance. As mentioned at the end of Section 3.4.2, the automatic cross-validation in GridSearchCV is not used, and a predefined split between the training and validation datasets is employed.
Study 3: Effect of Validation Strategy Under Distribution Shift. This study compares standard GridSearchCV validation strategy, where validation data is drawn from the training set, with a test-based validation strategy. In this alternative setup, part of the test engines is used for validation during hyperparameter tuning, while the rest is kept as a final hold-out test dataset. This approach requires setting to "False" the refit parameter within GridSearchCV and also prevents the use of K fold cross-validation. The goal is to evaluate whether aligning validation with the test distribution improves generalization on unseen truncated trajectories while training is performed on full trajectories. This experiment is intended as an exploratory analysis of the train–test distribution mismatch rather than as a standard evaluation protocol.
3.5.2 Pipeline Optimization Across Window size and Prediction Horizon. An experiment is designed to establish a comprehensive mapping between the problem formulation and the optimal pipeline. Specifically, for each combination of *window_size*, which defines the amount of past information considered, and *threshold_classification*, which determines the prediction horizon, the search pipeline described in Fig. 1 identifies the best-performing machine learning model, its optimal set of hyperparameters, and the most effective feature extraction configuration. This resulting pipeline is then applied to the test dataset to generate predictions.

4 Results

4.1 Core Methodological Experiments

4.1.1 Study 1: Impact of Temporal Windowing and Feature Extraction. The full CMAPSS FD001 dataset, identical ML hyperparameter space, StandardScaler, RandomOverSampler, GridSearchCV and a prediction horizon of 30 cycles that corresponds to the threshold_classification, are used in the experiments.

We first evaluated a naïve baseline using raw single-cycle data, so window_size is set to 1. Because single-cycle observations are highly sensitive to noise, simply expanding the input to raw sliding windows, with window_size of 30 cycles, improved the F1-score by +16.19% on average across all models. The results are displayed in Table 2.

Table 2. Impact of temporal windowing on F1-Scores on test dataset using raw data

Models	*F1-Scores for window_size = 1*	*F1-scores for window_size = 30*	*Improvement*
MLP	*0.6574*	*0.8018*	*+21.97%*
RandomForest	*0.7049*	*0.7455*	*+5.76%*
XGBoost	*0.8014*	*0.8081*	*+18.59%*
SVM	*0.6868*	*0.8183*	*+18.42%*

Next, we applied feature extraction to these 30-cycles windows, by computing classical statistics such as the mean, standard deviation, median, first and third quartiles, and a polynomial fit of degree 1. As shown in Table 3, this yielded an additional but modest average F1-score improvement of +4.11% over the raw 30-cycles windows. Interestingly, in a supplementary experiment using only a 20-engine subset instead of the full training dataset, raw data windows resulted in 15% lower F1-scores whereas the feature-based approach maintained consistent results.

Ultimately, feature-based windows are the preferred approach. They ensure more stable performance across varying dataset sizes and drastically reduce input dimensionality.

Table 3. F1-Scores on test dataset: raw data vs. feature-engineered windows

Models	*F1-Scores on raw data windows*	*F1-scores on featured windows*	*Improvement*
MLP	*0.8018*	*0.8094*	*+0.95%*
RandomForest	*0.7455*	*0.8251*	*+10.08%*
XGBoost	*0.8081*	*0.8344*	*+3.25%*
SVM	*0.8183*	*0.8309*	*+2.16%*

4.1.2 Study 2: Evaluation of Sampling Strategies. The experiments are conducted using a classification threshold of 30 and window size of 30. As described in Section 3.4.2 of the methodology section, four experimental settings are evaluated. Table 4 presents experiments (i), (ii), (iii), and (iv).

For experiments (iii) and (iv), the initial training data are split at the engine level into training and validation sets, using a random seed of 42 and a validation proportion of 0.3. The reason is that the designed sampling strategies should only be applied to the training fold and never to the validation fold. For experiments (i) and (ii), the same training split is used to ensure a fair comparison with experiments (iii) and (iv).

In experiment (ii), two sampling strategies are considered within the GridSearchCV procedure: a RandomUnderSampler with a sampling strategy of 0.5, and a RandomOverSampler with a sampling strategy of 1. Both samplers use a fixed random seed of 42. Ultimately, the RandomOverSampler is selected by GridSearchCV, as it yields the best performance.

In experiment (iii), as detailed in Section 3.4.2, the stochastic decay follows Equation (7), with $K \in \{0.5,1,1.5,2,2.5,3,3.5,4\}$, Equation (6). Finally, for experiment (iv), the explored parameters are the constant sampling probability $p \in \{0.1,0.2,0.3,0.4,0.5,0.6,0.7\}$ and $K \in \{0.5,1,1.5,2,2.5,3,3.5,4\}$.

For methods (iii) and (iv), it is worth noting that the additional use of standard scikit-learn sampling methods with RandomUnderSampler and RandomOverSampler, leads to very similar results. This can be explained by the fact that the proposed approaches already address the class imbalance issue by reducing the number of negative windows through a probabilistic sampling mechanism.

The results presented in Table 4 show a slight improvement when using a sampling method, with the method (iv) providing the best results in average.

Table 4. F1-Scores on test dataset across evaluated sampling strategies

Models	*No sampler(i)*	*Scikit learn samplers (ii)*	*Stochastic decay sampling (iii)*	*Stochastic constant sampling (iv)*
MLP	*0.8052*	*0.8163*	*0.8096*	*0.8282*
Random Forest	*0.8199*	*0.8367*	*0.8233*	*0.8370*
XGBoost	*0.8233*	*0.8206*	*0.8469*	*0.8508*
SVM	*0.8311*	*0.8602*	*0.8390*	*0.8305*

4.1.3 Study 3: Effect of Validation Strategy Under Distribution Shift. It is observed that a model achieving the best performance on a validation set derived from the training data does not necessarily yield the best results on the final test set. Moreover, in the first two studies, a consistent discrepancy is observed between the F1-scores obtained during cross-validation on validation dataset, around 0.92, and those achieved on the final test set, around 0.82. This gap does not appear to stem from traditional overfitting, as the F1-score on a held-out "test-from-train" subset remains close to 0.92. Instead, it is likely due to a distributional mismatch in the dataset, where training trajectories reach the failure point while test trajectories are truncated before the failure point.

To address this observation, a test-based validation strategy is considered. Specifically, 70% of the test set is used as the validation dataset within the GridSearchCV algorithm, while the remaining 30% is reserved as the final test set. To ensure a fair comparison with the classical approach, where the validation set is drawn from the training data by GridSearchCV, the final test set in the classical approach is restricted to the same 30% subset of engines used in the test-based validation strategy. The experiments are conducted using a classification threshold of 30 and window size of 30.

Interestingly, this test-based validation strategy does not seem to improve performance on the final test set; in fact, it results in an average decrease of 5.10%, as shown in Table 5. This may be explained by the fact that, under the classical approach, the model is trained on full trajectories and therefore has more training examples from which to learn and generalize across all stages of the engines' life cycles. Additionally, unlike standard GridSearchCV procedure, this test-based validation strategy does not automatically allow for K-fold cross-validation, as explained in Section 4.3, which may further limit its effectiveness.

Table 5. Impact of validation strategy on F1-Scores on test dataset

Models	*Train-based validation data*	*Test-based validation data*	*Improvement*
MLP	*0.8219*	*0.8211*	*-0.10%*
RandomForest	*0.8215*	*0.7684*	*-6.46%*
XGBoost	*0.8361*	*0.8083*	*-3.32%*
SVM	*0.8605*	*0.7700*	*-10.52%*

4.2 Pipeline Optimization Across Window Size and Prediction Horizon

The two-level optimization framework proposed in Fig. 1 was used to perform an extensive grid search jointly exploring machine learning architectures (MLP, Random Forest, XGBoost, and SVM), their associated hyperparameters shown in Table 1, the scaling strategies described in Section 3.4.1, and the feature extraction hyperparameter space described in Section 3.3.3 restricted to the mean, standard deviation, median statistical features, and polynomial regression of degree 1. Based on the results of Section 4.1.2, a RandomOverSampler with a sampling strategy of 1 and a random seed of 42 was used. Based on the results of Section 4.1.3, validation was performed using internal 3-fold cross-validation automatically done by GridSearchCV and grouped by engine ID. This process identified the optimal pipeline for each combination of *window_size*, representing the amount of past information considered, and *threshold_classification*, representing the prediction horizon, which was then evaluated on the test dataset. As observed in Study 1, when applying feature extraction to sliding windows, using only 20 engines for training can yield performance comparable to that obtained with the full dataset; therefore, due to computational constraints, 20 out of the initial 100 training engines were randomly selected using a fixed seed of 42, while the full test dataset was retained for evaluation. The results are illustrated in Fig. 4.

Performance increases with both window size and threshold, reaching its peak with a test F1-score of 0.89 at higher values ranging from 35 to 40, with SVM architectures dominating most configurations.

One may expect prediction performance to improve as the window size increases relative to the prediction horizon, since larger windows contain more temporal information. However, since the extracted features in this experiment are limited to basic statistics and first-degree polynomial trends, large window sizes may smooth out informative short-term variations. In contrast, smaller windows are better suited to capturing rapid and meaningful temporal dynamics. This trade-off may explain the approximate symmetry observed around the diagonal in the window size–threshold plane in Fig. 4, where both large windows capturing longer-term historical information and small windows emphasizing short-term dynamics yield comparable performance.

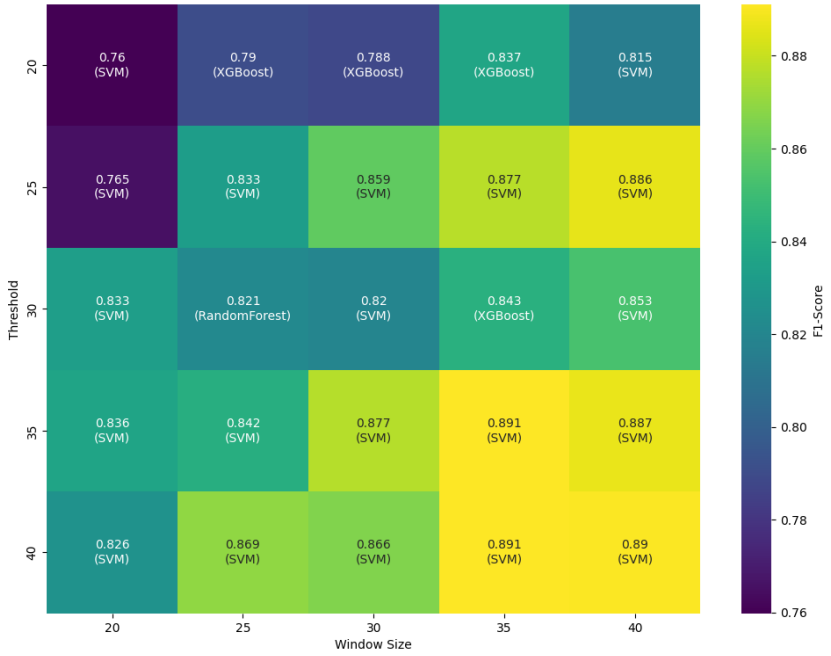


Fig. 4. Heatmap of maximum test F1-Scores across window size and prediction horizon

4.3 Limitations

Sampling should be applied only to the training folds and never to the validation folds in order to avoid data leakage and biased evaluation. In this study, to reduce computational cost, the proposed sampling strategies (iii) and (iv) described in Section 4.1.2 are performed during the window generation stage, prior to the feature extraction process explained in Section 3.3.3. Consequently, although data are grouped by engine ID, the use of automatic K-fold cross-validation in GridSearchCV would result in validation folds containing already sampled trajectories. Therefore, in Study 2, the automatic K-fold cross-validation in GridSearchCV is disabled, and the training and validation splits are predefined in advance.

This limitation may partly explain the results reported in Section 4.1.2, where the standard scikit-learn sampling method (ii) achieves performance comparable to the more advanced stochastic sampling strategies (iii) and (iv). Although methods (iii) and (iv) prioritize boundary-adjacent windows and better exploit trajectory context than the context-agnostic RandomOverSampler used in method (ii), the latter benefits from integrated K-fold cross-validation within the GridSearchCV pipeline.

In Study 3, a similar limitation arises from the absence of K-fold cross-validation, as a predefined subset of the test set is used as the validation set, while the remaining samples are used for final testing. The original training set is used as the training fold.

For both Study 2 and Study 3, where no K-fold procedure is implemented, a more advanced and computationally expensive approach could address this limitation by

iterating over multiple train–validation splits in Study 2, and multiple validation–test splits in Study 3, prior to entering the framework pipeline.

Another limitation concerns the restricted search space for feature engineering in Sections 4.1.1 and 4.2, which was constrained by high computational cost and resource requirements, particularly in Experiment 4.2.

Finally, due to computational constraints limiting the number of simulations, no variability measures across folds such as variance estimates and confidence intervals could be computed. In addition, only F1-scores are reported.

5 Conclusion

This study demonstrates that incorporating temporal context is critical when reframing RUL prediction into a binary classification task, as using sliding windows improves the F1-score by 16% compared to single-cycle observations, with feature extraction providing an additional 4% improvement. This feature-based approach is highly preferred, as it drastically reduces input dimensionality while showing limited sensitivity to the number of training trajectories.

To address class imbalance, the obtained results show that context-aware stochastic sampling does not outperform classical methods such as Random Oversampling.

Furthermore, attempting to mitigate the train-test distribution mismatch by validating on a subset of truncated test data counterintuitively decreased final performance; models seem to generalize better when trained on complete full-life trajectories.

Finally, the proposed two-level optimization architecture, Fig. 1, revealed that increasing both the window size and prediction horizon, up to 40 cycles, maximizes predictive performance, reaching a peak F1-score of 0.89 on the C-MAPSS FD001 test dataset, with SVM demonstrating the best performance across tested configurations, Fig. 4.

6 Discussion

In [1], classical machine learning algorithms such as LightGBM achieve a macro F1-score of 0.86, while LSTM-based models reach a macro F1-score of 0.92. These results are comparable to those obtained in our experiments, where an SVM optimized alongside a tailored feature extraction pipeline and a model tuning pipeline achieves an F1-score of 0.89 and a corresponding macro F1-score of 0.93. Due to the window construction described in Section 3.3.4 and the truncation of trajectories, the test set is strongly imbalanced, with a majority of negative samples. Moreover, the standard F1-score, which balances precision and recall, focuses on the positive class, whereas the macro F1-score equally weights both classes, including the dominant negative class. Consequently, the macro F1-score is typically higher in this setting, as performance on the negative class is generally stronger.

Further comparisons can be drawn with [2], which reports an F1-score of 0.92 using an LSTM, and with [16], where the combination of a Temporal Convolutional Network (TCN) and class imbalance handling yields an F1-score of 0.915. Notably, the

referenced studies [1], [2], and [16] all evaluate classification performance on the FD001 dataset using a 30-cycle prediction horizon.

Moreover, [15] demonstrates that, for classification tasks on the FD001 dataset, larger window sizes ranging from 40 to 70 cycles lead to improved performance on the test set. This finding is consistent with our observations, as our experimental results in Section 4.2 confirm that increasing the window size up to 40 cycles directly enhances classification performance.

Overall, while the results obtained in this study do not strictly outperform the highest state-of-the-art results reported in the literature, this slight gap can be explained by the use of classical machine learning architectures (SVM, XGBoost, Random Forest) and a standard deep learning Multi Layer Perceptron (MLP) architecture, as well as by the restricted hyperparameter search space used for feature extraction and model tuning—a constraint deliberately chosen to reduce computational costs. Nevertheless, the proposed two-level optimization framework provides a structured approach for systematically exploring the impact of feature engineering choices and model configurations on predictive performance.

Future work will focus on expanding the search space for feature extraction and model hyperparameters and including repeated runs to assess variability across folds. The proposed two-level optimization framework will also be evaluated on additional classification datasets to test generalization. Finally, the methodology will be extended to RUL regression and compared with alternative method such as survival analysis–based methods.

Acknowledgments. The author would like to thank Professor Boris Delibašić, Professor Miloš Jovanović, Dr. Sandro Radovanović and Dr. Andrija Petrović for their supervision throughout the preparation of this paper. The author also thanks the reviewers for their constructive feedback, which helped improve the quality of the manuscript. The author used artificial intelligence tools to improve language clarity and grammatical accuracy of this work.

Disclosure of Interests. The author has no competing interests to declare.

References

1. H. Özcan, "Predictive maintenance in aircraft engine maintenance using the C-MAPSS dataset: performance comparison and evaluation of machine learning classification algorithms," *Artificial Intelligence for Engineering Design, Analysis and Manufacturing (AI EDAM)*, vol. 40, Feb. 2026, doi: 10.1017/S0890060426100249.
2. M. Kimollo and X. Liu, "Aircraft Engine Remaining Useful Life (RUL) Prediction Using Machine Learning," *2024 International Conference on Machine Learning and Applications (ICMLA)*, Miami, FL, USA, 2024, pp. 1759-1763, doi: 10.1109/ICMLA61862.2024.00271.
3. U. Yıldırım and H. Afşer, "Linear Methods for Predictive Maintenance: The Case of NASA C-MAPSS Datasets," *Applied Sciences*, vol. 15, no. 18, p. 9945, Sep. 2025, doi: 10.3390/app15189945.
4. M. K. Kılınç, F. K. Ece, and A. Erciyes, "Hierarchical RUL Prediction for Turbofan Engines Based on Health Stage Classification and Change Point-Guided Data Augmentation," *IEEE Access*, vol. 12, pp. 11200-11215, 2024, doi: 10.1109/ACCESS.2025.3593206.

5. M. Dida, A. Cheriet, and M. Belhadj, "Remaining Useful Life Prediction Using Attention-LSTM Neural Network of Aircraft Engines," *International Journal of Prognostics and Health Management*, vol. 16, no. 2, Jul. 2025, doi: 10.36001/ijphm.2025.v16i2.4274.
6. Z. Yu, N. Lei, Y. Mo, X. Xu, X. Li, and B. Huang, "Feature Extraction Based on Self-Supervised Learning for Remaining Useful Life Prediction," *Journal of Computing and Information Science in Engineering*, vol. 24, no. 2, p. 021008, Feb. 2024, doi: 10.1115/1.4062599.
7. S. Kalyani, K. V. Rao, and A. M. Sowjanya, "A TimeImageNet Sequence Learning for Remaining Useful Life Estimation of Turbofan Engine in Aircraft Systems," *Structural Durability & Health Monitoring*, vol. 15, no. 4, pp. 317-334, Nov. 2021, doi: 10.32604/sdhm.2021.016975.
8. J. Lan, Z. Li, and S. Xue, "Engine remaining useful life prediction method based on deep residual network and attention mechanism," *Discover Computing 29"*, 109 (2026), doi: 10.1007/s10791-026-10000-8
9. S. Zhou, Y. Yao, A. Liu, F. Wang, L. Chen, and R. Xiong, "Multiform Informed Machine Learning Based on Piecewise and Weibull for Engine Remaining Useful Life Prediction," *Sensors*, vol. 23, no. 12, p. 5669, Jun. 2023, doi: 10.3390/s23125669.
10. P. Mao, Y. Lin, S. Xue, and B. Zhang, "Remaining Useful Life Estimation of Aircraft Engines Using Differentiable Architecture Search," *Mathematics*, vol. 10, no. 3, p. 352, Jan. 2022, doi: 10.3390/math10030352.
11. U. Thakkar, "Remaining Useful Life Prediction of a Turbofan Engine Using Deep Layer Recurrent Neural Networks," Master's Thesis, Carleton University, Ottawa, Canada, 2021.
12. Y. Wang, M. Ragab, Y. Hou, Z. Chen, M. Wu, and X. Li, "Deep Domain Adaptation for Turbofan Engine Remaining Useful Life Prediction: Methodologies, Evaluation and Future Trends," *arXiv preprint arXiv:2510.03604*, Oct. 2025, doi: 10.48550/arXiv.2510.03604.
13. O. Tsurkan, A. Konstantinova, A. Sedykh, A. Senokosov, D. Tarpanov, M. Anoshin, A. Sagingalieva, and A. Melnikov, "Hybrid quantum recurrent neural network for remaining useful life prediction," *arXiv preprint arXiv:2504.20823*, Apr. 2025, doi: 10.48550/arXiv.2504.20823.
14. Z. Fan, W. Li, and K.-C. Chang, "A Two-Stage Attention-Based Hierarchical Transformer for Turbofan Engine Remaining Useful Life Prediction," *Sensors*, vol. 24, no. 3, p. 824, Jan. 2024, doi: 10.3390/s24030824.
15. E. Jeon, D. Jin, and Y. Kim, "Effects of Window and Batch Size on Autoencoder-LSTM Models for Remaining Useful Life Prediction," *Machines*, vol. 14, no. 2, p. 135, Jan. 2026, doi: 10.3390/machines14020135.
16. R. D. Sainz Calderon, "A Predictive Maintenance Framework for Remaining Useful Life Classification on IoT Enabled Devices using Neural Networks," Master's Thesis, School of Computing, National College of Ireland, Dublin, Ireland, 2022.
17. S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge University Press, 2004.

Original Research Article *https://doi.org/10.5281/zenodo.23475078*

Differences in Statistics Anxiety Across Academic Years: The Role of Perceived Usefulness of Statistics

Luna Radević[1[0000-0003-1046-8910]], Milica Maričić[1[0000-0003-0441-9899]] and Tatjana Von Rosen[2[0000-0002-8610-0365]]

[1] University of Belgrade, Faculty of Organizational Sciences
Jove Ilića 154, Belgrade, Serbia
luna.radevic@ff.uns.ac.rs; milica.maricic@fon.bg.ac.rs;

[1] Stockholm University
Albanovägen 12, Hus 4, Stockholm, Sweden
tatjana.vonrosen@stat.su.se;

Abstract. Statistics anxiety represents a specific form of academic anxiety that can function as an affective barrier to learning and applying statistical methods. Previous studies have highlighted the importance of motivational beliefs such as perception of relevance and value of statistics as relevant antecedents of statistics anxiety. The main aim of this study was to examine whether the association between perceived usefulness of statistics and statistics anxiety varied across three student cohorts sampled in 2023, 2024 and 2026. To this end, a moderation analysis using multi-categorical moderator was conducted. Results indicated that perception of statistics as a useful and relevant discipline is consistently negatively associated with lower levels of statistics anxiety across all cohorts. This finding suggests that motivational belief such as utility-related perception may function as stable protective mechanism against anxiety, independently of broader contextual factors. Although the overall interaction effect did not reach statistical significance, a closer insight into bootstrap confidence intervals revealed a significant moderation effect in 2023-2024 contrast, with the protective effect of perceived usefulness being strongest in 2023, and less pronounced in later years. Taken together, these findings suggest that while the relationship between perceived usefulness and statistics anxiety is generally robust, it may be modestly shaped by contextual factors. Findings are discussed in terms of their implications for statistics education and motivational interventions targeting utility value perceptions.

Keywords: Statistics Anxiety, Perceived Usefulness of Statistics, Moderation Analysis.

1 Introduction

In contemporary higher education, statistical literacy has emerged as a fundamental skill required across a wide range of academic fields. The ability to collect, analyze, and meaningfully interpret quantitative data is no longer limited to traditionally quantitative disciplines, but is expected of students of humanities, behavioral, health and social sciences as well [1]. Statistics education therefore equips students with the tools and knowledge necessary for conducting research, managing data and making evidence-based decisions in a society that is increasingly data-driven. However, despite its importance and clear role in professional development, statistical courses are often considered as overly difficult and are frequently associated with discomfort and anxiety, which may lead to poorer performance and avoidance. [2]. This gap between objective relevance and students' subjective experience represents an important challenge in statistics education and served as a foundation for the present study.

The paper is structured in the following sections. The introductory part briefly describes the importance of statistical literacy in higher education and professional development. The second section provides a literature review of research focused on statistics anxiety and its correlates, specifically perceived usefulness as a motivational disposition. In the next section, we introduce the present study and cover the research design and methodology; main aim of the study, our proposed conceptual moderation model and its' hypotheses, statistical model for its' validation, sample characteristics and information about used instruments. The fourth section provides the main results of the study while the last section summarizes key findings, discusses the practical implications, outlines limitations of the present study and future research directions.

2 Literature Review

Statistics anxiety is commonly used to describe a complex affective response that includes worry, stress, and avoidance arising from engagement with statistical content in academic, professional, or everyday contexts [3]. It represents a specific form of academic anxiety that is related to but distinct from other anxiety forms, such as test, spatial and math anxiety [4]. Numerous studies have shown that higher levels of statistics anxiety are negatively associated with statistics performance, attitudes toward statistics, perceived self-efficacy and positively related to tendency toward procrastination [5, 6]. Beyond these consequences, elevated statistics anxiety often leads to avoidance of courses focused on research methods and statistics as well as poorer professional engagement with quantitative data [7]. Overall, consequences of statistics anxiety can be detrimental for both students' academic and professional development, underscoring the importance of research aimed at better understanding the factors that contribute to it.

The first line of research on statistics anxiety aimed to explore its antecedents. For example, Trassi et al. [8] in their systematic review found that learning strategies, procrastination and self-efficacy appeared to be reliable predictors of statistics anxiety in numerous studies. The other line was focused on construction and validation of

instruments and measures, which consequently led to a debate on construct dimensionality. To date, The Statistics Anxiety Rating Scale (STARS), developed by Cruise et al. [3] has remained the most commonly used instrument, measuring multiple dimensions such as interpretation anxiety, test and class anxiety, computational self-concept, fear of asking for help and fear of statistics teacher. Vigil-Colet et al. [9] developed another widely used scale named Statistical Anxiety Scale (SAS) which focused on three main anxiety domains - examination anxiety, interpretation anxiety and asking for help anxiety. At the same time, these researchers began integrating statistics anxiety's correlates (such as attitudes towards statistics) in their theory and showed strong associations with statistics anxiety and its outcomes. However, despite the continued development of this research field, many questions remain a matter of debate.

2.1 Attitudes towards statistics and perceived usefulness

Research focused on dimensionality as well as on the antecedents of statistics anxiety consistently identified the importance of both dispositional and situational factors. However, it seems that recent studies have been slightly more focused on exploring dispositional ones. For instance, Cui et al. [10] offered a framework for organization of dispositional factors of statistics anxiety into three distinct categories. The first group of factors gathered attitudes towards statistics, motivational aspects and previous experiences with the subject. The second encompasses behavioral patterns (e.g. learning strategies), while the third domain referred to individual characteristics and cognitive variables.

Among these dispositional characteristics, motivational factors have received substantial empirical attention. Researchers have examined a range of attitudes toward statistics, such as self-confidence in statistical abilities, expectations regarding task difficulty and performance as well as perceived relevance of the subject [11]. Among these, perceived relevance is of particular importance, as it reflects the extent to which students see statistics as valuable for their academic path and future career. Moreover, this construct corresponds to the utility value component of academic motivation within the framework of well-established Situated Expectancy-Value Theory (SEVT) [12, 13]. According to this theory, utility value is expected to reduce the emotional distress of engaging with demanding academic content as statistics is usually considered, by providing a motivational rationale for that cost. In other words, students who perceive statistics as generally useful are expected to experience less anxiety in relation to it.

The role of perceived value in buffering domain-specific anxiety is not unique to statistics. Research in adjacent quantitative and technical disciplines consistently demonstrates that motivational value beliefs operate as protective factors against academic anxiety. In the context of mathematics anxiety, a construct closely related to statistics anxiety and sharing several common antecedents [5], studies have shown that students' subjective valuation of mathematics is significantly associated with lower anxiety levels [14]. Similarly, in foreign language learning research, higher utility value and expectancy-value motivation have been linked to reduced classroom anxiety, with students who perceive strong instrumental relevance of a subject, showing systematically lower anxiety responses [15]. Taken together, these findings suggest that the

association between perceived usefulness and lower anxiety reflects a broader motivational mechanism consistent with SEVT. Because statistics is often perceived as difficult and anxiety-provoking, yet its perceived usefulness can be shaped through instruction, research on this topic has important theoretical and practical implications. Accordingly, our research represents a case study of the relationship between perceived usefulness and anxiety in the domain of statistics..

3 Present Study

Following the rationale of SEV Theory, the general aim of our study was to examine whether perceived usefulness of statistics can predict level of statistics anxiety. In addition, an important but underexplored question is whether this relation is stable across time and independent of broader contextual conditions. This question became particularly important during the COVID-19 pandemic which caused disruption to higher education and had enduring impacts on students' psychological functioning as well as academic motivation [16]. With that in mind, our second aim was to explore potential differences in relation between perceived usefulness and statistics anxiety among student cohorts sampled in three academic years: 2023, 2024 and 2026.

3.1 Proposed conceptual model and hypotheses

Following the SEVT framework and empirical studies reviewed so far, the present study proposes a moderation model in which perceived usefulness predicts statistics anxiety, while academic year of data collection (or student cohort) moderates the strength of this relationship. The model is presented in Fig. 1.

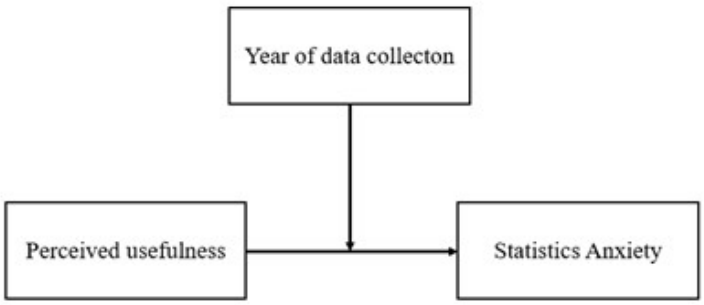


Fig. 1. Proposed moderation model.

Two research hypotheses can be formed based on the proposed conceptual model:

Hypothesis 1: Perceived usefulness of statistics has an impact on statistics anxiety.

Hypothesis 2: Year of data collection has a moderating effect on the relation between perceived usefulness of statistics and statistics anxiety.

3.2 Moderation Analysis Procedure

A moderation analysis was conducted using the PROCESS macro for SPSS v.4.2 [17], specifying Model 1. Perceived usefulness of statistics was entered as the predictor (X) and statistics anxiety as the outcome variable (Y). Both variables were computed as mean scores of their respective scale items. The moderator (W) was the academic year of data collection, treated as categorical variable with three levels (2023, 2024, 2026). Given the multi-categorical nature of the moderator, it was dummy-coded with 2023 as the reference category. As a result, two indicator variables were created (W1: 2024 vs. 2023 and W2: 2026 vs. 2023).

The model included direct effects of the predictor and the dummy moderator variables, as well as their interaction terms (X * W1 and X * W2). Statistical inference was based on 5,000 bootstrap resamples with 95% bias-corrected percentile confidence intervals.

3.3 Research instrument

The questionnaire consisted of several sections. First, some general information about students were collected such as gender, year of study and study program they attend. Second, the 5-item adaptation from [9] was used to assess students' statistics anxiety, specifically in relation to the interpretation of statistical results. Third, a set of seven items was adapted from [18] in order to assess students' perceived usefulness of statistics in their future careers as well as in everyday life. The items assessed beliefs about the value, relevance, and applicability of statistics across professional and personal contexts and included the following statements: "Statistics is worthless", "Statistics should be a required part of my professional training", "Statistical skills will make me more employable", "Statistics is not useful to the typical professional", "Statistical thinking is not applicable in my life outside my job", "I use statistics in my everyday life"," I will have no application for statistics in my profession".

3.4 Sample and procedure

Our sample comprised 513 students of Faculty of Organizational Sciences, University of Belgrade. The survey was administrated during online lectures in summer semester of each academic year (2023, 2024, 2026). Students were asked to fill out the questionnaire and to forward it to other students from the class. The sample was relatively balanced across years, with 189, 214, and 110 participants in each year, respectively. The majority of respondents were female (61.4%), but this disproportion was somewhat expected since more female students are enrolled the Faculty of Organizational sciences [19]. Most of the participants were second-year students (86%) enrolled in the statistics course at the time of data collection, while the rest of them were third- or fourth-year students who had not yet passed the course exam.

There was also a slight disproportion regarding the study program students attend, with 69.8% were students of the ICT, while others attended the Business and Management program.

4 Results

Table 1 presents descriptive statistics for the two main variables. In average, students reported relatively low to moderate levels of statistics anxiety, while perceived usefulness had a higher mean, suggesting that students generally viewed statistics as a useful discipline. Both instruments demonstrated satisfactory internal consistency with Cronbach's alpha > 0.60 [20]. Both variables' distributional characteristics indicated approximated normality with both skewness and kurtosis within acceptable ranges [21].

Table 1. Descriptive statistics of main variables.

Variable	Mean	Std	Sk	Ku	Cronbach's α
Statistics anxiety	2.31	0.91	0.69	-0.25	0.84
Perceived usefulness	3.82	0.69	-0.35	-0.44	0.81

Note. *Std* – Standard deviation, *Sk* – Skewness, *Ku* – Kurtosis.

The overall moderation model was statistically significant F (5, 501) = 18.56, p < 0.001, and accounted for approximately 16% of the variance in statistics anxiety (R^2 = .16). The summary of all model coefficients is presented in Table 2. The main and negative effect of perceived usefulness of statistics was significant, indicating that higher perceived usefulness of statistics was associated with lower statistics anxiety. Also, this was the strongest predictor in the model.

Neither of the moderator main effects appeared to be significant which means that levels of statistical anxiety didn't differ between years once perceived usefulness and the interaction terms were accounted for.

The first tested interaction effect (Perceived usefulness * W1) approached conventional level of significance, with p level being exactly 0.05. Moreover, while the standard 95% CI marginally included zero, the bootstrap CI narrowly excluded it which taken together suggests this moderating effect should be interpreted with caution. The interaction between the predictor and other dummy moderator (Perceived usefulness * W2) did not reach statistical significance by either inferential approach.

Table 2. Overall model coefficients.

Predictor	*b*	*SE*	*t*	*p*	95% CI	Boot 95% CI
Constant	4.66	0.34	13.84	< 0.001	[3.99, 5.32]	[3.97, 5.31]
Perceived usefulness	-0.63	0.09	-7.39	**< 0.001**	[−0.80, −0.46]	[−0.79, −0.46]
W1 (2024 vs. 2023)	-0.90	0.50	-1.81	0.07	[−1.88, 0.07]	[−1.85, 0.08]
W2 (2026 vs. 2023)	-0.63	0.53	-1.20	0.23	[−1.68, 0.41]	[−1.82, 0.58]
Perceived usefulness * W1	0.25	0.13	1.96	**0.05**	**[−0.00, 0.50]**	**[0.01, 0.49]**

Perceived usefulness * W2	0.22	0.14	1.64	0.10	[−0.04, 0.49]	[−0.08, 0.53]

Note. Reference category for W = 2023. b = unstandardized regression coefficient.

Test of the highest-order unconditional interaction (Perceived usefulness * W, Table 3) evaluated whether year of data collection as a whole significantly moderated the perceived usefulness-statistics anxiety relation. This interaction appeared to be insignificant. In other words, when two interaction terms were combined, together they added less than 1% of explained variance beyond other effects and consequently, failed to reach statistical significance.

Table 3. Test of highest-order interaction.

	ΔR²	*F*	*df1*	*df2*	*p*
Perceived usefulness * W	0.01	2.33	2	501	0.10

Table 4 presents the simple slopes, precisely, the effects of perceived usefulness of statistics on statistics anxiety separately within each year. This analysis revealed that the negative effect of the predictor was statistically significant in all three years, with the strongest relation observed in 2023. The effects observed in 2024 and 2026 were similar and somewhat smaller, indicating that the effect didn't change progressively over time, but rather reflected a shift between the 2023 and the two later cohorts.

Overall, all three slopes were statistically significant and negative, meaning that in every analyzed year, students who found statistics as more useful, were less anxious about it.

Table 4. Conditional effects of the Perceived usefulness on Statistics Anxiety by the year of data collection

Year	*b*	*SE*	*t*	*p*	95% CI
2023	-0.63	0.09	-7.39	< 0.001	[-0.81, -0.46]
2024	-0.38	0.09	-4.07	**< 0.001**	[−0.57, −0.20]
2026	0.22	0.11	-3.83	<0.001	[−0.62, -0.20]

5 Conclusion

Out of two research hypotheses derived from proposed conceptual model, one has been confirmed. Our results revealed that perceived usefulness of statistics is a robust and significant negative predictor of statistics anxiety related to data interpretation in every examined cohort. This finding could be interpreted within the framework of expectancy-value theory of academic motivation [14, 15] which points out the role of perceived task value in shaping emotional responses. Moreover, this result is in line with prior research showing that attitudes toward statistics are important dispositional

predictors influencing anxiety in statistics education contexts [22]. Overall, these findings highlight an important practical implication. Potential interventions aimed at reducing statistics anxiety should be focused on reshaping students' motivational beliefs by strengthening perception of the value and relevance of statistics.

Our second hypothesis was not fully supported. However, the weaker usefulness-anxiety association observed in 2024 in contrast to 2023 supported by the bootstrap confidence interval is a theoretically interesting and noteworthy finding, despite the non-significant overall interaction effect. Students in 2023 cohort entered their studies during a period of greater uncertainty and more dispersed general anxiety following the pandemic which may explain why the role of perceived usefulness had greater protective role in their emotional responses. Prior research examining student cohorts before, during and after the pandemic found that perceived level stress has largely returned to pre-pandemic level after 2020, but aspects of academic motivation and study commitment have remained relatively lower for longer period of time [23]. At the same time, findings across multiple cohorts point to a gradual decline in moderate anxiety among university students as the post-pandemic period progressed [24], which is consistent with proposed interpretation. Although tentative due to the lack of direct measures of pandemic-related factors, this interpretation offers a coherent account of the observed pattern and a useful starting point for future research. The varying strength of this relation across time periods, indicates that the effectiveness of such intervention may depend on broader contextual and cohort-specific differences.

Taken together, our findings have several practical implications. Enhancing students' perception of relevance and usefulness of statistics in professional and everyday life can reduce statistics anxiety and as such should be integrated into course design.

Several limitations should be considered. First, cross-sectional design of our study prevents causal conclusions about the relationship between perceived usefulness and statistics anxiety, leaving open the possibility of reverse causation or the influence of third factors, such as academic motivation, success in previous quantitative experience etc. In addition, longitudinal research design should be applied to examine individual changes over time, rather than differences across cohorts. Second, the study did not include measures that may explain differences among cohorts in more details, such as contextual factors, variation in academic environments and post-pandemic-related stress.

Future research should aim to replicate these findings in more diverse samples, ideally using longitudinal designs. Examining potential mechanisms such as baseline anxiety, prior experience in statistics and perceived level of self-efficacy would provide more precise understanding of the effects of contextual factors on usefulness-anxiety relation. Additionally, testing targeted interventions aimed at strengthening perceived usefulness may offer practical strategies for reducing statistics anxiety in higher education.

Acknowledgments. The authors have no acknowledgments to declare.

Differences in Statistics Anxiety Across Academic Years: The Role of Perceived Usefulness of Statistics

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Burrill, G., & Pfannkuch, M.: Emerging trends in statistics education. ZDM Mathematics Education **56**, 19–29 (2024). https://doi.org/10.1007/s11858-023-01501-7
2. Vuco, N., & Buljan, I.: Patterns of Use and Attitudes Toward Statistics in Psychological Profession: A Cross-Sectional Study. Psychology Learning & Teaching **24**(3), 257-275 (2025)
3. Cruise, R. J., Cash, R. W., & Bolton, D. L.: Development and validation of an instrument to measure statistical anxiety. In Proceedings of the American Statistical Association, Section on Statistical Education, 92–97, American Statistical Association (1985).
4. Onwuegbuzie, A. J.: Academic procrastination and statistics anxiety. Assessment & Evaluation in Higher Education **29**(1), 3-19 (2004). https://doi.org/10.1080/0260293042000160384
5. Paechter, M., Macher, D., Martskvishvili, K., Wimmer, S., & Papousek, I.: Mathematics Anxiety and Statistics Anxiety. Shared but Also Unshared Components and Antagonistic Contributions to Performance in Statistics. Frontiers in Psychology, **8,** 1196 (2017). https://doi.org/10.3389/fpsyg.2017.01196
6. Chew, P. K. H., & Dillon, D. B.: Statistics Anxiety Update: Refining the Construct and Recommendations for a New Research Agenda. Perspectives on Psychological Science, **9**(2), 196–208 (2014). https://doi.org/10.1177/1745691613518077
7. Hunt, B. W., Mari, T., Knibb, G., Christiansen, P., & Jones, A.: Statistics anxiety and predictions of exam performance in UK psychology students. PLoS One, **18**(8), e0290467 (2023). https://doi.org/10.1371/journal.pone.0290467
8. Trassi, A. P., Leonard, S., Rodrigues, L. D., Rodas, J. A., & Santos, F. H.: Mediating factors of Statistics Anxiety in university students: a systematic review and meta- analysis. Annals of the New York Academy of Sciences, **1512**, 76–97 (2022). https://doi.org/10.1111/nyas.14746
9. Vigil-Colet, A., Lorenzo-Seva, U., & Condon, L.: Development and validation of the statistical anxiety scale. Psicothema, **20**(1), 174-180 (2008).
10. Cui, S., Zhang, J., Guan, D., Zhao, X., & Si, J.: Antecedents of statistics anxiety: An integrated account. Personality and Individual Differences **144**, 79-87 (2019). https://doi.org/10.1016/j.paid.2019.02.036
11. Peiró-Signes, Á., Trull, O., Segarra-Oña, M., & García-Díaz, J. C.: Anxiety towards statistics and its relationship with students' attitudes and learning approach. Behavioral sciences, **11**(3), 32 (2021). https://doi.org/10.3390/bs11030032
12. Eccles, J. S., & Wigfield, A.: From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. Contemporary educational psychology, **61**, 101859 (2020). https://doi.org/10.1016/j.cedpsych.2020.101859
13. Eccles, J. S., & Wigfield, A.: Expectancy-value theory to situated expectancy-value theory: Reflections on the legacy of 40+ years of working together. Motivation Science, **9**(1), 1-12 (2023). https://doi.org/10.1037/mot0000275
14. Faber, G., & Drexler, H.: Predicting education science students' statistics anxiety: The role of prior experiences within a framework of domain-specific motivation constructs. Higher Learning Research Communications, **9**(1) (2019). https://doi.org/10.18870/hlrc.v9i1.435

15. Dong, L., Liu, M., & Yang, F.: The Relationship Between Foreign Language Classroom Anxiety, Enjoyment, and Expectancy-Value Motivation and Their Predictive Effects on Chinese High School Students' Self-Rated Foreign Language Proficiency. Frontiers in Psychology, **13**, 860603 (2022). https://doi.org/10.3389/fpsyg.2022.860603
16. Troncone, A., Caldarelli, G., Cosenza, M., Affuso, G., Sacco, M., Ciccarelli, M., & Pizzini, B.: Psychological problems and academic motivation in university students: A cross-sectional study one year after the COVID-19 lockdown in Italy. Psychiatric Quarterly, **96**(3), 683-704 (2025). https://doi.org/10.1007/s11126-025-10138-6
17. Hayes, A. F.: Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression Based Approach. 2nd ed. Guilford publications, New York (2017)
18. Cruise, Robert J., and Elizabeth M. Wilkins. "STARS: Statistical anxiety rating scale." Unpublished manuscript, Andrews University, Berrien Springs, MI 20 (1980).
19. University of Belgrade: Izveštaj univerzitetske komisije za upis studenata u prvu godinu studijskih programa. Beograd (2020)
20. Loewenthal, K. M.: An introduction to Psychological Tests and Scale, 2nd ed. Psychology Press, Hove (2004)
21. Tabachnick, B. G., & Fidell, L.-S.: Using multivariate statistics, 6th ed. Pearson, Boston (2013)
22. Rosenzweig, E. Q., Wigfield, A., & Eccles, J. S.: Beyond utility value interventions: The why, when, and how for next steps in expectancy-value intervention research. Educational Psychologist, **57**(1), 11-30 (2022). https://doi.org/10.1080/00461520.2021.1984242
23. Dresen, V., Staggl, S., Fischer-Jbali, L., Canazei, M., & Weiss, E.: Stress, Burnout and Study-Related Behavior in University Students: A Cross-Sectional Cohort Analysis Before, During, and After the COVID-19 Pandemic. Brain Sciences, **15**(7), 718 (2025). https://doi.org/10.3390/brainsci15070718
24. Chen, Y., Ke, X., Liu, J., Du, J., Zhang, J., Jiang, X., ... & Xiao, X.: Trends and factors influencing the mental health of college students in the post-pandemic: four consecutive cross-sectional surveys. Frontiers in psychology, **15**, 1387983 (2024). https://doi.org/10.3389/fpsyg.2024.1387983.

Original Research Article *https://doi.org/10.5281/zenodo.23475116*

Reproducibility in IS research as a Weak Signal: A bibliometric Foresight Analysis

Sara Koprivica[1][0000-0001-7077-8780], Aleksandra Hornjak[2][0009-0005-1211-7091], and Darko Stefanovic[3][0000-0001-9200-5092]

[1, 2, 3] University of Novi Sad, Faculty of Technical Sciences
[1]havzisara@uns.ac.rs, [2]aleksandra.hornjak@uns.ac.rs,
[3]darko.stefanovic@uns.ac.rs

Abstract. This paper aims to examine whether research reproducibility constitutes a weak signal in Software Engineering, Computer Science, and Information Systems literature. Using bibliometric analysis as a foresight instrument, four indicators derived from Hiltunen's weak signal framework are applied to a corpus of 57 publications retrieved from Scopus. Three indicators are fully satisfied, while one is partially satisfied, pointing toward reproducibility as an emerging research concern with demonstrable early-stage growth, fragmented disciplinary presence, and limited, but accelerating visibility. The findings suggest that reproducibility in SE, CS, and IS has not yet reached mainstream status but is showing early signs of potential transition from a weak signal toward a recognized research priority.

Keywords: reproducibility, foresight, weak signals, bibliometric analysis, software engineering, computer science, information systems

1 Introduction

Reproducibility is regarded as the foundation of scientific credibility, as the ability of independent researchers to obtain consistent results under comparable conditions. Without it, the validity of published findings remains uncertain, and the nature of scientific progress is undermined [1]. Findings are considered reproducible when the used data are openly accessible, and the applied methods are clearly described [2]. Additionally, it supports and promotes the dissemination of research [3, 4].

A survey conducted in 2016 [5] has shown that the reproducibility of scientific findings has emerged as a growing concern across multiple disciplines. A great number of research studies across disciplines may be unreliable, driven by pressure to publish or other questionable research practices, rather than a commitment to producing verifiable, high-quality findings [6]. The reproducibility crisis has affected numerous fields, ranging from the natural sciences, such as biology and chemistry, to the social sciences, including economics and education [3]. Computing research faces particular reproducibility challenges due to its multidisciplinary nature, complex socio-technical artifacts, and poor documentation practices. Despite increasing awareness, empirical

assessments of reproducibility in this domain remain rare, and no widely accepted evaluation framework exists [3].

The reproducibility crisis was not an issue that arose overnight. It developed gradually, manifesting across isolated publications or fragmented discussions [7]. The forces that may shape the future are often already present, manifesting as weak signals that remain largely unrecognized [7, 8]. This raises the question of how such concerns can be identified and tracked before they reach mainstream visibility. Foresight is a systematic process for anticipating the future to inform present decision-making. Weak signals represent little pieces of information or minor events that carry potential future significance, yet remain too small to exert immediate impact [7, 9].

This paper employs bibliometric analysis [10] as a foresight tool for tracking the growth and structural development of research topics over time, enabling the detection of weak signals in academic literature. Therefore, this paper aims to examine whether research reproducibility constitutes a weak signal in Software Engineering (SE), Computer Science (CS), and Information Systems (IS) literature. To this end, four bibliometric indicators are analyzed within the defined corpus.

The remainder of this paper is structured as follows. Section 2 presents the theoretical background, covering the concepts of reproducibility, foresight, and weak signals. Section 3 describes the research methodology, including the bibliometric approach and data collection process. Section 4 presents the results of the bibliometric analysis. Section 5 discusses the findings in the context of weak signal theory, Section 6 concludes the paper with implications for IS researchers and directions for future research, and Section 7 describes limitations associated with the study.

2 Related Work

The terminology around reproducibility remains inconsistent across disciplines. The terms reproducibility and replicability are often used interchangeably in the literature, without a clear distinction [1].

Plesser [11] deals exclusively with the terminological confusion surrounding the terms repeatability, replicability, and reproducibility as three distinct concepts. He traces two conventions: Claerbout and ACM conventions. The Claerbout convention defines reproducibility as obtaining the same results by running the same code on the same data, while replicability refers to independently reimplementing a study and obtaining similar results. The ACM convention, on the other hand, reverses this. As Plesser notes, these two conventions are in direct conflict, which is a source of confusion in the literature.

This inconsistency further complicates efforts to systematically address the reproducibility problem [1]. This paper uses the term reproducibility to refer to obtaining the same results using the same data and methods, and replicability to refer to obtaining consistent findings using new data [11]. Given this terminological inconsistency, the bibliometric analysis presented in this paper includes both terms in the search query, as detailed in Section 3.

An early sign of the replication crisis was shown in a work from 2005 by Ioannidis [12], who mathematically demonstrated that a majority of published research findings may be false due to statistical biases, small sample sizes, and publication bias. It was a mathematically supported argument highlighting a weakness in research methodology. This argument was later confirmed on a large scale.

A key moment in growing scientific awareness of the reproducibility problem was the large-scale study conducted by the Open Science Collaboration [13] in which 100 published psychological studies were attempted to be replicated. Only around 36% of replications have shown statistically significant results, and the effect sizes were about half of what the original studies reported. The authors attributed these findings to a combination of factors, including publication bias, small sample sizes, pressure to publish positive results, and selective reporting. Notably, the Open Science Collaboration uses the term reproducibility to refer to what this paper defines as replicability, reflecting the broader terminological inconsistency. However, this study is mentioned as empirical evidence of the scale of the problem, regardless of how the terms are defined.

When the crisis was globally confirmed, the reproducibility problem was later empirically confirmed in computer science as well [14]. Collberg and Proebsting [15] empirically examined repeatability in computer systems research by analyzing 601 papers from ACM conferences and journals, attempting to locate and build their associated source code. They found that only 32.3% of papers were repeatable under time-constrained conditions. They also use the term repeatability to describe what this paper would define as reproducibility. These findings suggest that reproducibility challenges in computing are not isolated incidents but an issue requiring attention [14]. This problem was further confirmed within the IS community.

The authors of the study [3] empirically assessed the state of reproducibility in computing research by analyzing 100 papers from the ICIS 2019 conference proceedings. They evaluated reproducibility across three factors: method, data, and experiment, using a 25-variable checklist. Their findings revealed that while 95% of papers were method reproducible, only 28% documented experimental details sufficiently. The authors attributed this to poor documentation practices and proposed a pre-submission reproducibility checklist as a remedy.

These findings suggest that reproducibility concerns in CS and IS have been documented, present across the literature, but without a coherent disciplinary response. Whether this pattern of growing but fragmented attention constitutes a weak signal of an emerging research concern is the central question this paper seeks to answer.

The concept of weak signals was first introduced by Ansoff [16], and further developed by Hiltunen [7], who distinguishes weak signals from emerging issues. An emerging issue refers to an event or cluster of events, but a weak signal is the early evidence pointing toward such an event. Weak signals are also distinguished from trends, which already exhibit a recognizable pattern, and from megatrends, which are already dominant. A weak signal precedes all.

To detect such signals, this paper employs bibliometric analysis, which has been recognized as an explorative foresight method for mapping the development of research fields [17]. Applied here, it serves to examine whether research reproducibility constitutes a weak signal in SE, CS, and IS literature.

3 Methodology

This section describes the methodological design used to address the central research question of this study: *Does research reproducibility constitute a weak signal in Software Engineering, Computer Science, and Information Systems literature?* To answer this question, a bibliometric analysis was conducted, but as a foresight instrument, primarily for its ability of detecting weak signals through quantifiable patterns in publication activity, source distribution, and citation behavior.

The section consists of three segments: data source selection and search strategy, a two-phase selection process, the treatment of replication studies, and the operationalization of weak signal indicators.

3.1 Data Source Selection and Search Strategy

The bibliographic database Scopus is used for the paper. The database, launched by Elsevier, was selected for its broad cross-publisher coverage, including IEEE, Springer, ACM, and Elsevier content, and for its standardized metadata schema, which enables consistent bibliometric extraction.

Limitations to using a single database will be further addressed in Section 7.

A single search query was formulated to target titles, abstracts, and keywords of publications within the defined subject areas. The query is presented below:

("reproducibility" OR "replicability" OR "repeatability" OR "reproducible research" OR "reproducibility crisis")

AND

("information systems" OR "information technology" OR "software engineering" OR "empirical software engineering" OR "IS research" OR "IT research")

The subject area filter was applied to ensure coverage of the three target disciplines as defined by Kitchenham [18]: Computer Science (COMP), Software Engineering covered through Engineering (ENGI), and Information Systems covered through Business, Management and Accounting (BUSI) and Decision Sciences (DECI). Only journal articles and conference papers published in English were included.

The query returned 1,617 results. The results were obtained and exported on April 14th, 2026. No publication year filter was applied to the search query. The temporal range of the selected studies (2011–2026) emerged from the data. The year 2011 represents the earliest publication identified in the resulting corpus, rather than a predefined lower boundary. This unconstrained temporal scope is methodologically appropriate for weak signal detection. Following the database search, a selection process was conducted to identify publications relevant to the research question.

3.2 A Two-Phase Selection Process and the Treatment of Replication Studies

The selection process was conducted in two phases: screening by title and abstract, and full-text review. The first phase is applied to all 1,617 results retrieved from Scopus.

Publications were screened against the inclusion and exclusion criteria using only the title, abstract, and keywords of the studies. The defined inclusion and exclusion

criteria applied in both phases are shown in Table 1. The third inclusion criterion was partially ensured through the Scopus filter, which restricts results to peer-reviewed journal articles and conference papers, and further verified during full-text review. The third exclusion criterion was applied only in the second phase.

Table 1. Inclusion and Exclusion Criteria

Type	No.	Criterion
Inclusion	I1	Addresses reproducibility, replicability, or replication crisis in SE, CS, and IS concepts
	I2	Investigates or measures reproducibility, proposes tools, methods, guidelines, critically discusses it, or calls for action
	I3	Published in a peer-reviewed journal or conference
Exclusion	E1	Belongs exclusively to domains outside of SE, CS, and IS
	E2	Mentions reproducibility without substantive analysis
	E3	No access to the full text publication
	E4	Duplicated record

A study was excluded if it failed to meet any inclusion criterion or matched any exclusion criterion. This phase resulted in 182 publications advancing to full-text review, while 1,435 were excluded. 836 papers were excluded at the title level, and 599 at the abstract level.

The remaining 182 studies were reviewed against the inclusion and exclusion criteria at the full-text level, with the second exclusion criterion requiring a more detailed examination, as incidental mentions are often only detectable through full-text reading. Moreover, full-text access to the studies had to be permitted, hence the third exclusion criterion. After this phase, the corpus of 57 publications was identified for a bibliometric analysis. Screening was conducted by a single researcher, which constitutes a methodological limitation discussed in Section 7.

To describe the screening decision, three examples are presented. The publication [19] was excluded under E1, as it addressed computational reproducibility within the social sciences, a domain outside the defined scope of SE, CS, and IS. Publication [20] was excluded under E2, as it did not investigate, measure, or discuss reproducibility as a research concern. The term was used solely to characterize the experimental setup. In contrast, publication [21] was included under the I2 criterion, as it empirically investigated the current state of research artifact publication and reproducibility practices in empirical software engineering, proposing a structured framework for improving reproducibility of human-oriented experiments.

Additionally, 28 publications were identified as studies that explicitly replicated or reproduced the results of a previously published work. These publications were excluded from the primary corpus as their contribution lies in validating prior findings rather than advancing knowledge about reproducibility itself. Their presence is noted here as evidence of an active replication community within SE, CS, and IS, which can be seen as a sign of growing engagement with the topic. The existence of those studies, along with their temporal distribution, will be discussed further in the paper.

3.3 The Operationalization of Weak Signals

Hiltunen's conceptual framework [7] characterizes weak signals through low visibility, fragmented presence, and early-stage growth dynamics. Building on this framework, four bibliometric indicators are developed to empirically operationalize weak signals.

The indicators with thresholds and operationalizations are presented in Table 2. Publication counts and their development over time reflect how visible and recognized a topic is, while the distribution across source venues indicates whether a topic has yet consolidated around established venues.

The first indicator, Low Initial Publication Volume, reflects Hiltunen's observation that weak signals go unnoticed and circulate in a small number of channels. The second, Sustained Growth Trend, captures the gradual emergence of a research area, tolerating short-term fluctuations consistent with early-stage instability. The third, Growth Acceleration, identifies the point at which a topic transitions from marginal to recognized, marked by a measurable increase in publication rate following an inflection point. The fourth, Source Fragmentation, reflects the absence of an established disciplinary core, assessed through the dispersion of publications across source venues.

Table 2. Indicators for Weak Signals

No.	Indicator	Operationalization	Threshold
1	Low Initial Publication Volume	Mean annual publication count in the pre-inflection subperiod	< 3 publications per year
2	Sustained Growth Trend	Overall direction of annual publication counts; final third vs. initial third of observation period	Final third substantially exceeds initial third; isolated single-year declines tolerated
3	Growth Acceleration	Ratio of post-inflection to pre-inflection mean annual publication rate	Post-inflection mean ≥ 2× pre-inflection mean
4	Source Fragmentation	a) Mean publications per source venue b) proportion of single-publication venues	a) < 2 publications/venue b) ≥ 80% single-publication venues

The thresholds were operationalized as quantitative translations from Hiltunen's qualitative criteria, as her framework does not specify quantitative thresholds.

The threshold of fewer than three publications per year for Low Initial Publication Volume was set to reflect Hiltunen's description of weak signals as appearing in very few channels and remaining largely unrecognized.

The threshold for Sustained Growth Trend was operationalized descriptively rather than numerically, reflecting Hiltunen's characterization of weak signals as showing signs of development with early-stage fluctuations before stabilization.

The 2× ratio for Growth Acceleration was selected as the minimum acceleration that would be meaningfully distinguishable from normal year-to-year fluctuation, doubling the baseline rate. Under a relaxed ratio of 1.5× for Growth Acceleration, the indicator would still be satisfied, as the observed ratio of approximately 4× substantially exceeds any reasonable minimum threshold.

The 80% threshold for Source Fragmentation was set as a boundary reflecting near-total source dispersion. This value was selected to ensure that the majority of venues contribute only once, which would indicate the absence of any emerging disciplinary core. A lower threshold, such as 70%, would also be consistent with fragmentation. Under that relaxed criterion, the fourth indicator would be fully satisfied, and the overall conclusion would remain unchanged.

The thresholds are treated as provisional operationalizations open to refinement in future work.

Bibliometric and descriptive methods are applied across temporal, source, geographic, keyword, and citation dimensions of the corpus of 57 selected studies. Metadata was exported from Scopus in CSV format and processed in Microsoft Excel.

The data, including years, author keywords, and affiliations, were manually extracted and analyzed by the researcher. Author keywords were grouped into thematic clusters inductively from the keyword set. The manual approach was chosen due to the small size of the corpus.

Duplicate detection was handled within the selection process as criterion E4. Citation counts reflect the values reported by Scopus on the export date (April 14th, 2026) and represent a snapshot rather than a longitudinal measure.

To enable a comparative subfield analysis, the 57 publications were manually classified into three subfields (SE, CS, and IS), based on a review of each paper's title, abstract, and author keywords. A paper was classified as SE if its primary focus was on reproducibility within software development processes, empirical studies, testing, artifact evaluation, or software repository mining. A paper was classified as CS if its primary focus was on reproducibility in computational systems, scientific workflows, high-performance computing, or machine learning platforms. A paper was classified as IS if its primary focus was on reproducibility within IS research methods or IS-specific empirical practices. Papers with overlapping signals were assigned to the dominant subfield at the researcher's discretion.

In the subsequent sections, the findings are presented using visual representations and evaluated against the four predefined weak-signal indicators.

4 Results

Figure 1 presents the annual publication trends for the period 2011–2026. The first period (2011-2019) yielded between one and three publications per year, with a total of 15 publications, and a mean of 1.67.

The growth was recorded from 2020 onward. The annual count increased, ranging from four to nine publications per year and totaling 42 publications across seven years, of which 40 were published in the complete years 2020–2025. The mean annual publication rate for this period is 6.67 publications per year, excluding 2026 as a partial year. Two publications were recorded in 2026 at the time of data export (April 14th, 2026), and are included in the total corpus count but excluded from rate calculations to avoid underestimation.

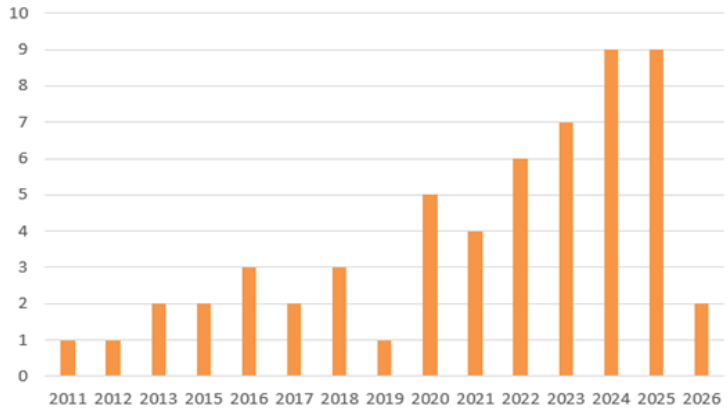


Fig. 1. Distribution of studies by year

Figure 2 illustrates the distribution of document types. Of the 57 included publications, 36 are conference papers (63%), and 21 are journal articles (37%).

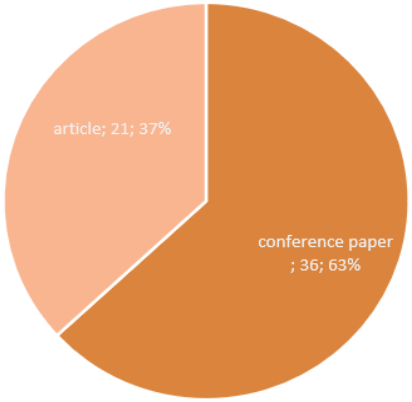


Fig. 2. Distribution by publication type

All 57 publications are distributed among 36 source venues. Figure 3 presents the 12 most productive sources by publication count. One source contributed five publications, two sources contributed four, two sources contributed three, and seven sources contributed two publications each. The remaining 24 sources each contributed a single publication, representing 66% of all contributing venues.

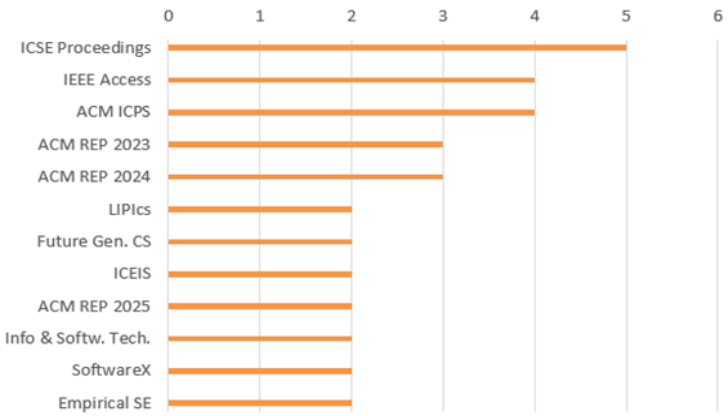


Fig. 3. Distribution by source venue

The geographic distribution of the studies spans 26 countries across six continents. Based on author affiliations, which may include multiple countries per publication, European institutions contributed the largest share (46), followed by North American (20) and South American (11) institutions.

Four publications included Asian affiliations, three included Australian or Oceanian affiliations, and no African affiliations were present in the corpus. The full country-level distribution is presented in Figure 4.

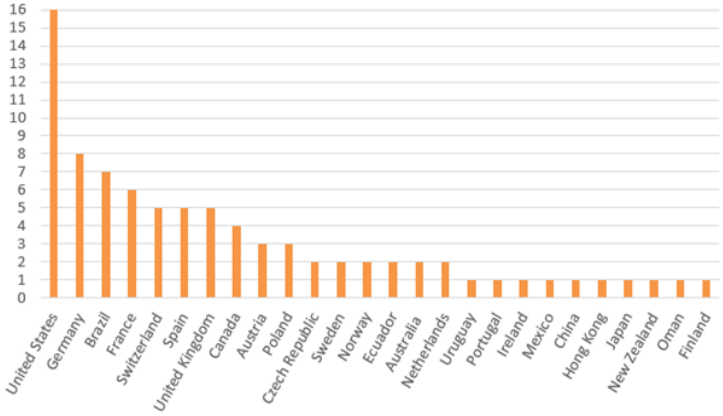


Fig. 4. Distribution across countries

The author keywords were grouped into four thematic clusters, presented in Table 3: a core concept cluster, an open science cluster, a CS/SE/IS concept cluster, and an AI/ML concept cluster. The most frequent keyword across the corpus was reproducibility, followed by reproducible research and open science.

The terms replicability and replication appeared considerably less frequently, despite both being foundational to the topic. Notably, the term replication crisis, used in

psychology and biomedical fields, does not appear in the corpus, suggesting that the SE, CS, and IS community approaches the topic through a distinct conceptual vocabulary.

Table 3. Keyword Clusters

Cluster	Keywords
Core concept	reproducibility (25), reproducible research (7), replicability (2), replication (3), scientific reproducibility (2)
Open Science	open science (6), open data (3), artifact evaluation (3), FAIR data (2), open research (2), research artifacts (2), provenance (2)
CS/SE/IS concept	empirical software engineering (5), software engineering (4), systematic review (4), software engineering research (3), research software engineering (3)
AI/ML concept	machine learning (2), data science (2), Jupyter notebooks (2), scientific workflows (2), high performance computing (2), simulation (2)

Table 4 presents the total number of citations, which ranges from 0 to 110, with a mean of 14.1 and a median of 4. The substantial difference between mean and median indicates a highly skewed distribution: 41 of the 57 publications (72%) received fewer citations than the mean, while the 16 publications above the mean collectively account for 83% of all citations in the corpus.

Table 4. Distribution within citation range

Range	Number of studies
0 citations	9
1-10 citations	29
11-50 citations	15
50 or more citations	4

Figure 5 shows the subfield distribution. The manual classification of the 57 publications resulted in 38 SE papers, 17 CS papers, and 2 IS papers. Both SE and CS show consistent growth after 2020, with post-inflection means of 4.33 and 2.0 publications per year, respectively. Publications from 2026 were excluded from the mean calculation because it is a partial year. IS contributed only 2 publications across the entire observation period, with no publications recorded before 2021.

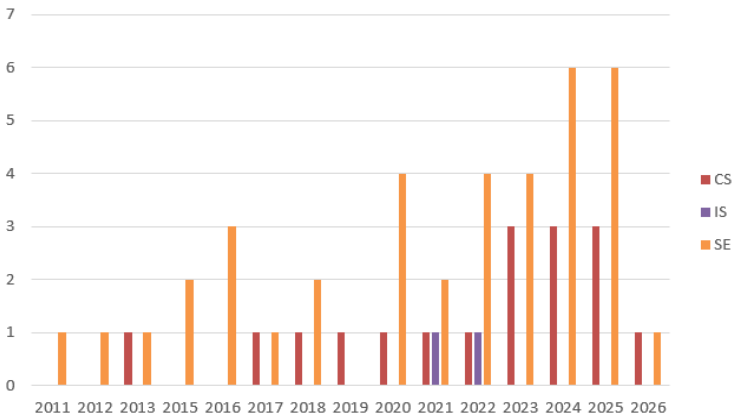


Fig. 5. Subfield distribution

Beyond the primary corpus of 57 publications, a full-text review identified 28 additional studies as direct replication or reproduction attempts of previously published work. While excluded from the bibliometric analysis, their temporal distribution is noted here as an observation. Of the 28 studies, 5 were published before 2020, while 23 were published in 2020 or later, with a notable concentration in 2025 (14 studies) and 2026 (7 studies). The temporal analysis is shown in Figure 6.

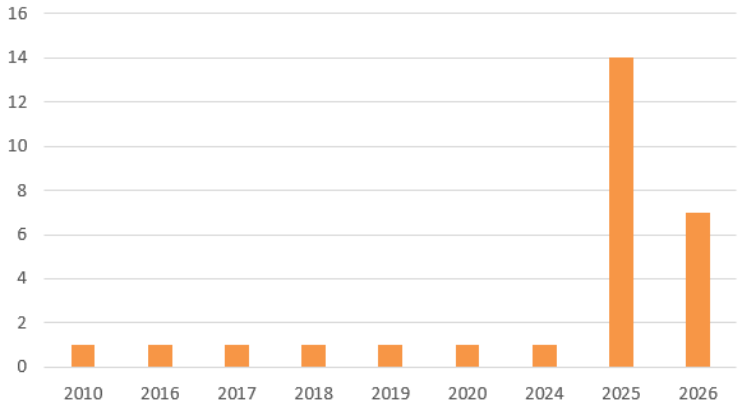


Fig. 6. Temporal analysis of additional studies

5 Discussion

The results will be discussed in terms of four bibliometric indicators for weak signal operationalization evaluated in Section 3.

Low Initial Publication Volume is confirmed by the temporal distribution. Between 2011 and 2019, the mean annual publication count was approximately 1.67 publications per year, falling below the defined threshold of three. The absence of any publication

in 2014 further reflects the marginal status of the topic during this period. This pattern is consistent with Hiltunen's [7] characterization of weak signals as circulating in a small number of channels and remaining largely unrecognized by the broader research community.

Sustained Growth Trend is confirmed across the observation period. Annual publication counts increased from one publication in both 2011 and 2012 to nine in 2024 and 2025. A single-year decline in 2019, one publication following three in each of the two preceding years, does not invalidate the trend, as short-term fluctuations are expected in early-stage research areas [7]. The final third of the observation period substantially exceeds the initial third, satisfying the indicator threshold.

The inflection point in 2020 was identified through analysis of publication counts by year, which show a sustained increase from that year onward (one publication in 2019 to five in 2020), followed by consistently higher counts in subsequent years, after a period of low publication numbers between 2011 and 2019. The shift in publication rate following the inflection point identified in 2020 confirms growth acceleration. The pre-inflection mean (2011–2019), which was approximately 1.67 publications per year, rose to 6.67 publications per year (excluding 2026 as a partial year), exceeding the threshold of two times the pre-inflection mean, with the post-inflection rate approximately four times higher than the pre-inflection rate. This acceleration coincides with broader institutional recognition of reproducibility as a research concern, including the establishment of the ACM Conference on Reproducibility and Replicability in 2023, a dedicated venue whose emergence suggests that the topic is showing early signs of transition from a marginal concern toward a recognized research agenda.

Source Fragmentation is partially satisfied by the distribution across 36 distinct source venues. The mean publication count per venue is 1.58, below the threshold of two. Single-publication venues account for 24 of 36 contributing sources (66%), not reaching the 80% threshold. While this falls short of the defined threshold for the second fragmentation measure, the absence of a dominant core venue, with contributions scattered across journals and conferences from multiple disciplines, is consistent with a topic that has not yet consolidated around an established publication venue.

Beyond the four indicators, the broader pattern of results points in the same direction. The dominance of conference papers is common in early-stage SE, CS, and IS research, where new ideas surface at conferences before a field develops a stable journal presence. The citation data reflects a similar immaturity: most publications have received very few citations, while a small cluster of older works accounts for 83% of all citations in the corpus. The absence of the term replication crisis from the keyword analysis suggests that SE, CS, and IS researchers are working through this problem on their own terms rather than importing the framing from psychology and biomedicine. Geographically, contributions are heavily concentrated in Europe and North America, consistent with a topic that has not yet reached global attention. The 28 replication studies identified but excluded from the corpus further suggest that some researchers are already acting on reproducibility concerns at the level of practice, which may be an early sign of emerging institutionalization.

A comparative subfield analysis further suggests that the weak signal pattern is not limited to a single discipline. Both SE and CS show growth ratios that go beyond the

2× threshold of the third indicator, suggesting that awareness of reproducibility as a research concern has developed across both communities.

The near-absence of IS publications, despite IS being explicitly included in the search scope, suggests that reproducibility has not yet emerged as a recognized research concern within the IS community, further reinforcing the weak signal interpretation.

The results of the temporal distribution of the 28 replication studies follow the same growth pattern as the primary corpus, suggesting that actively reproducing published results has accelerated similarly with the corpus of conceptual literature on the topic. This is consistent with a topic showing early signs of transition from a weak signal toward a recognized research priority.

Three of the four indicators are fully satisfied, and the fourth is partially satisfied, pointing toward research reproducibility in SE, CS, and IS, as exhibiting characteristics consistent with a weak signal.

6 Conclusion

The main aim of this paper was to answer the question of whether research reproducibility constitutes a weak signal in SE, CS, and IS literature, using bibliometric analysis as a foresight instrument.

Four indicators derived from Hiltunen's conceptual framework were applied to a corpus of 57 publications retrieved from Scopus. Three indicators were fully satisfied, while source fragmentation was only partially satisfied. The evidence represents an emerging research concern that has not yet reached mainstream visibility, but is showing clear signs of growth and increasing disciplinary attention.

These findings align with related work documenting reproducibility challenges in SE, CS, and IS research. The paper also extends that work by situating it within a foresight perspective. This paper asks whether reproducibility is becoming a recognized problem. The bibliometric evidence suggests that reproducibility is becoming a recognized concern, though the process remains gradual and geographically concentrated. Researchers and institutions in SE, CS, and IS should treat reproducibility as an emerging priority rather than a marginal concern, as the trajectory of the field suggests that its visibility is likely to continue to grow.

7 Limitations

Several limitations of this study should be acknowledged.

First, the analysis relies exclusively on Scopus, which may have excluded relevant publications indexed only in Web of Science, IEEE Xplore, or ACM Digital Library.

Second, full-text access was required as an inclusion criterion, which introduces selection bias toward openly accessible work and may partly explain the limited representation of publications from some venues.

Third, screening was conducted by a single researcher, which may have introduced subjective variability in inclusion and exclusion decisions and could have affected corpus composition.

Fourth, Scopus subject area classifications are assigned at the journal level rather than the article level, meaning some relevant publications may have been excluded or misclassified by the subject area filter.

The subfield classification was performed manually, which introduces interpretive subjectivity similar to the single researcher limitation noted above. Although this approach avoids the issues of venue-level subject area assignment, borderline cases between SE, CS, and IS were decided based on the researcher's judgment.

Fifth, the four indicator thresholds are provisional operationalizations derived from Hiltunen's qualitative framework and have not been validated against other known weak signals. Alternative threshold choices were considered in Section 3.3, and the overall conclusion remains stable under relaxed criteria.

Finally, citation counts reflect a single snapshot taken on April 14th, 2026, which structurally disadvantages more recent publications regardless of their quality or relevance.

Acknowledgments. This research has been supported by the Ministry of Science, Technological Development and Innovation (Contract No. 451-03-34/2026-03/200156) and the Faculty of Technical Sciences, University of Novi Sad, through project "Scientific and Artistic Research Work of Researchers in Teaching and Associate Positions at the Faculty of Technical Sciences, University of Novi Sad 2026" (No. 01-3609/1).

The authors used Claude (Anthropic) for language editing and proofreading, and take full responsibility for the contents of this manuscript.

References

1. [1] M. Miłkowski, W. M. Hensel, and M. Hohol, 'Replicability or reproducibility? On the replication crisis in computational neuroscience and sharing only relevant detail', *J. Comput. Neurosci.*, vol. 45, no. 3, pp. 163–172, Dec. 2018, doi: 10.1007/s10827-018-0702-z.
2. [2] K. M. Merz *et al.*, 'Editorial: Method and Data Sharing and Reproducibility of Scientific Results', *J. Chem. Inf. Model.*, vol. 60, no. 12, pp. 5868–5869, Dec. 2020, doi: 10.1021/acs.jcim.0c01389.
3. [3] W. Raghupathi, V. Raghupathi, and J. Ren, 'Reproducibility in Computing Research: An Empirical Study', *IEEE Access*, vol. 10, pp. 29207–29223, 2022, doi: 10.1109/ACCESS.2022.3158675.
4. [4] D. Ghoshal *et al.*, 'Experiences with Reproducibility: Case Studies from Scientific Workflows', in *Proceedings of the 4th International Workshop on Practical Reproducible Evaluation of Computer Systems*, in P-RECS '21. New York, NY, USA: Association for Computing Machinery, Jun. 2021, pp. 3–8. doi: 10.1145/3456287.3465478.
5. [5] M. Baker, '1,500 scientists lift the lid on reproducibility', Nature. Accessed: Mar. 31, 2026. [Online]. Available: https://www.nature.com/articles/533452a
6. [6] D. Fanelli, 'Is science really facing a reproducibility crisis, and do we need it to?', *PNAS*, vol. 115, no. 11, pp. 2628–2631, Mar. 2018, doi: https://doi.org/10.1073/pnas.1708272114.

7. [7] E. Hiltunen, 'Weak Signals in Organizational Futures Learning', Doctoral Thesis, Aalto University School of Economics, Helsinki, 2010.
8. [8] E. Hiltunen, 'The future sign and its three dimensions', *Futures*, vol. 40, no. 3, pp. 247–260, Apr. 2008, doi: 10.1016/j.futures.2007.08.021.
9. [9] S. Mendonça, G. Cardoso, and J. Caraça, 'The strategic strength of weak signal analysis', *Futures*, vol. 44, no. 3, pp. 218–228, Apr. 2012, doi: 10.1016/j.futures.2011.10.004.
10. [10] N. Donthu, S. Kumar, D. Mukherjee, N. Pandey, and W. M. Lim, 'How to conduct a bibliometric analysis: An overview and guidelines', *J. Bus. Res.*, vol. 133, pp. 285–296, Sep. 2021, doi: 10.1016/j.jbusres.2021.04.070.
11. [11] H. E. Plesser, 'Reproducibility vs. Replicability: A Brief History of a Confused Terminology', *Front. Neuroinformatics*, vol. 11, Jan. 2018, doi: 10.3389/fninf.2017.00076.
12. [12] J. P. A. Ioannidis, 'Why Most Published Research Findings Are False', *PLOS Med.*, vol. 2, no. 8, p. e124, Aug. 2005, doi: 10.1371/journal.pmed.0020124.
13. [13] Open Science Collaboration, 'Estimating the reproducibility of psychological science', *Science*, vol. 349, no. 6251, p. aac4716, Aug. 2015, doi: 10.1126/science.aac4716.
14. [14] T. Hummel and J. Manner, 'A Literature Review on Reproducibility Studies in Computer Science', in *Proceedings of the 16th ZEUS Workshop on Services and their Composition (ZEUS 2024)*, in CEUR Workshop Proceedings, vol. 3673. Ulm, Germany: CEUR-WS.org, pp. 54–62. [Online]. Available: https://ceur-ws.org/Vol-3673/paper9.pdf.
15. [15] C. Collberg and T. A. Proebsting, 'Repeatability in computer systems research', *Commun ACM*, vol. 59, no. 3, pp. 62–69, Feb. 2016, doi: 10.1145/2812803.
16. [16] H. I. Ansoff, R. P. Declerck, and R. L. Hayes, 'From Strategic Planning to Strategic Management', in *Strategische Unternehmungsplanung / Strategische Unternehmungsführung*, D. Hahn and B. Taylor, Eds., Heidelberg: Physica-Verlag HD, 1997, pp. 105–143. doi: 10.1007/978-3-662-41482-8_5.
17. [17] E. Gibson, T. U. Daim, E. Garces, and M. Dabic, 'Technology Foresight: A Bibliometric Analysis to Identify Leading and Emerging Methods', *Foresight STI Gov.*, vol. 12, no. 1, pp. 6–24, 2018, doi: 10.17323/2500-2597.2018.1.6.24.
18. [18] B. Kitchenham and S. Charters, 'Guidelines for performing Systematic Literature Reviews in Software Engineering', Keele University and Durham University, EBSE-2007-01, Jan. 2007.
19. [19] F. Momeni, M. T. Khan, J. Kiesel, and T. Ross-Hellauer, 'Checklists for Computational Reproducibility in Social Sciences: Insights from Literature and Survey Evaluation', in *Proceedings of the 3rd ACM Conference on Reproducibility and Replicability*, in ACM REP '25. New York, NY, USA: Association for Computing Machinery, Oct. 2025, pp. 179–191. doi: 10.1145/3736731.3746161.
20. [20] P. Zou, 'DBSel-CV: Residual-Learned Cardinality Estimation with Control-Variate Correction over SQLite', in *Proceedings of the 2025 3rd International Conference on Mathematics and Machine Learning*, in ICMML '25. New York, NY, USA: Association for Computing Machinery, Jan. 2026, pp. 85–90. doi: 10.1145/3783779.3783794.
21. [21] C. Guevara-Vega, B. Bernárdez, M. Cruz, A. Durán, A. Ruiz-Cortés, and M. Solari, 'Research artifacts for human-oriented experiments in software engineering: An ACM badges-driven structure proposal', *J. Syst. Softw.*, vol. 218, p. 112187, Dec. 2024, doi: 10.1016/j.jss.2024.112187.

Original Research Article *https://doi.org/10.5281/zenodo.23475148*

Analysis of the Level of Acceptance and Application of Artificial Intelligence in Medical Diagnostics

Tijana Ponjarac[1][0009-0007-2576-4823]

[1] University of Belgrade – Faculty of Organizational Sciences, Jove Ilića 154, Belgrade, Serbia
tp20253152@student.fon.bg.ac.rs

Abstract. This paper explores the perception and opinions of respondents about the use of artificial intelligence (AI) in medical diagnostics, with a focus on the application of tools such as ChatGPT. The aim of the paper is to examine the level of trust of respondents in the use of AI tools as a supporting tool in making medical diagnoses. For the purposes of this paper, a survey was conducted on 160 respondents in Serbia. Within the study, it was examined how sociodemographic factors affect the level of patient trust when using AI in medicine, as well as in self-diagnosis. The results indicate that age and level of education have a significant impact on trust and the use of AI tools, while gender has a limited impact. Factor analysis determined that there is a single construct of trust, while cluster analysis identified different groups of respondents who have specific characteristics and behavior patterns. The application of machine learning models showed high predictive ability, and trust in the use of artificial intelligence stood out as the most significant factor affecting the willingness to use it for medical diagnostic purposes. A special scientific contribution of the paper is the application of mediation analysis, which confirmed that trust represents a key mechanism through which previous experience influences the behavior of respondents. The obtained results contribute to a better understanding of the factors that influence the acceptance and use of artificial intelligence tools in medical diagnostics and can be used as a basis for further research and improvement of the implementation of AI systems in healthcare.

Keywords: Artificial Intelligence in Medicine, AI Adoption, Self-diagnosis, Medical diagnostics

1 Introduction

Artificial intelligence (AI) is a tool that has become a part of everyday life for a large number of individuals, whose ways of use are improving, expanding and changing day by day. In medicine, AI is increasingly being used and is becoming an important tool of modern healthcare, opening new directions in diagnostics, risk prediction, treatment, drug discovery, robotic surgery and prevention through daily developments in these fields [4, 32, 20, 48, 30]. AI systems analyze patient records, conduct imaging studies, map genomics, are integrated in wearable devices with the aim of detecting patterns,

and have a significant impact in the fields of radiology, cardiology, genomics, and public health [32, 48]. In cancer research alone, artificial intelligence supports early detection through image analysis and adapts treatments and therapies individually [48]. AI systems have shown improved diagnostic accuracy, which is sometimes equal to or even better than that of human experts, especially in radiology, pathology, cardiology and dermatology [39, 32, 20]. They reduce errors that occur during diagnosis and shorten the time required to establish a diagnosis, which enables earlier treatment. One of the significant contributions of AI in medicine is the simplification and reduction of administrative processes, which reduces waiting times and allows a larger number of patients to be served. Hospitals are introducing bidirectional systems, through platforms such as CFlow, for automating data collection, routine tasks, treatment processes, patient prioritization and record keeping, which significantly reduces the workload of doctors and leads to saving valuable time, placing patient care in the center of focus instead of paperwork [12]. Despite all these advancements, experts still believe that AI will not replace medical professionals [9].

AI has its limitations when it comes to its application in healthcare. There may be an increase in health inequality due to bias in training datasets, and an insufficient level of transparency and accountability in the AI decision-making processes that can threaten patient safety [19]. Although AI has significantly advanced in medicine, there are still many barriers that prevent its use in real-world treatment: data privacy and security [28, 33], lack of standardized datasets and interoperability [19], regulatory, interpretability ("black box" models) [22], clinician resistance and training gaps [11], and cost barriers in resource-limited settings [26].

With the development of large language models such as ChatGPT, new possibilities have opened in the field of diagnostics in medicine. These models are capable of analyzing symptoms described by patients, providing assumptions about possible diagnoses and suggesting next steps in further treatment. According to recent studies, ChatGPT has shown promising results in symptom classification and the speed of making preliminary diagnostic conclusions, which represents a significant step toward automation and more accurate decision-making [18, 21].

The aim of this paper is to analyze the perception of the use of artificial intelligence as a supporting tool in medicine. Through this paper, it will be examined how sociodemographic factors affect patient trust when using artificial intelligence in medicine, as well as in self-diagnosis, putting the focus on trust and willingness to use ChatGPT for self-diagnosis. To achieve this goal, a survey was conducted, and its results were interpreted using SPSS and R. Literature review and theoretical background

2 Literature review and theoretical background

Artificial intelligence is present in everyday activities, both in the lives of individuals and in large systems such as hospitals. The goal of its integration in medicine is to reduce the amount of unnecessary and routine work, as well as to eliminate errors that occur due to subjective decision-making, which will be explained in more detail in this chapter.

2.1 Application of artificial intelligence in MRI analyses

AI, especially deep learning, has significantly changed the landscape of magnetic resonance imaging (MRI) by enhancing image acquisition, reconstruction, segmentation, diagnosis, and workflow efficiency across a range of clinical applications. AI-driven approaches have enabled faster scan times, improved image quality, automated segmentation, and even radiologist-level diagnostic accuracy in certain domains [43, 46, 15]. AI-driven tools streamline radiology workflows by automating repetitive tasks, scan planning, triage prioritization, segmentation for example, and by reducing both scanning and reading times for radiologists [37]. Automated reporting systems based on large language models are emerging but require further validation.

2.2 Application of artificial intelligence in X-ray analyses

Chest X-ray imaging is one of the most commonly used diagnostic methods in medicine, and its timely and accurate interpretation is crucial for high-quality patient care. Studies show that physicians misinterpret approximately 30% of images, which can lead to delayed treatment, increased risk for patients, and unnecessary costs. A large number of physicians analyzing X-ray images are not specialized in radiology, which further increases the likelihood of errors [6]. Deep learning models, especially convolutional neural networks, have achieved excellent results in lung disease classification, fracture detection, and dental caries identification [22, 31, 44, 5, 34, 45]. AI also enables low-dose imaging protocols, thereby reducing radiation exposure [13].

2.3 Application of artificial intelligence in CT analyses

AI is rapidly transforming computed tomography (CT) imaging, offering significant advancements in image quality, diagnostic accuracy, workflow efficiency, and patient safety. Recent studies have shown AI's role in enhancing image analysis, automating detection and classification tasks, optimizing radiation dose, and supporting clinical decision-making across a range of applications including cardiovascular, oncologic, neurological, and musculoskeletal imaging [30, 1]. Methods of using AI for diagnostic purposes have led to high-quality images with reduced radiation dose, which has a significant contribution to patient safety. AI is applied in these purposes across different fields: cardiovascular imaging, oncology, neurology, musculoskeletal and trauma imaging, and lung cancer screening.

2.4 Trust in AI-based medical diagnostics

Trust is a central and most important determinant when it comes to adoption and effective usage of AI in medical diagnostics. Even though AI systems have demonstrated high accuracy and efficiency in disease detection, clinical integration is followed by concerns over transparency, explainability, accountability, bias, and the preservation of human oversight [38, 33, 49, 16]. Patients and healthcare workers both express greater

trust in AI when it is used with the clinicians, rather than when it replaces human expertise. [42, 33, 41, 3]. However, trust depends on user factors, healthcare context, technology design, regulatory framework, and ethical considerations. [49, 40 8, 29]. Meta-analyses that were conducted using UTAUT/TAM frameworks confirm that trust is of the strongest predictors of intention to use AI decision-making systems, surpassing traditional factors which include performance expectations or effort expectations [14].

2.5 Sociodemographic factors and AI acceptance

Sociodemographic factors, including age, gender, and education, play a significant role in shaping acceptance and trust of AI across different domains and daily life. Recent studies have shown that younger people have a more positive attitude towards the usage of AI [36, 7, 47], while older adults often express skepticism or lower trust in AI systems due to unfamiliarity or perceived risk. [2, 27, 36] Also, there is a difference in the AI acceptance when it comes to gender – females have fewer positive attitudes toward AI usage in general, but also in medicine [10, 36, 23, 49]. People with higher education or income also exhibit more positive attitudes toward AI usage in general. [35 , 23, 36].

2.6 Research questions and hypotheses

Based on the previous literature on technology acceptance and trust in healthcare, when it comes to using modern technology, the study is guided by the following research questions:
RQ1: What is respondents' level of trust in the use of artificial intelligence and its tool in medical diagnostics?
RQ2: How do sociodemographic factors such as gender, age, and level of education influence and affect trust in AI-based medical diagnostics and willingness to use ChatGPT for self-diagnosis?
RQ3: Does trust in AI mediate the relationship between respondents' previous experience when it comes to using AI tools and willingness to use it as a self-diagnosis tool?
RQ4: Can respondents be grouped into distinct profiles based on their trust, technological literacy, and patterns in using AI?
In line with these questions, the following hypothesis are to be tested:
H1: Respondents' gender is associated with willingness to use ChatGPT for self-diagnosis. This hypothesis is based on the assumption that gender may be influencing willingness to make health decisions using digital tools.
H2: Age is significantly associated with the frequency of using ChatGPT and AI tools for self-diagnosis. It is expected that younger respondents use ChatGPT more frequently, even for self-diagnosis.
H3: Age is significantly associated with trust in AI for medical diagnostics. This hypothesis is being tested to examine whether trust in AI is different across age groups when AI is used to interpret MRI, X-Ray, and CT scans.
H4: Education level is significantly associated with trust in AI-based medical diagnostics. It is expected that the respondents with higher education are expected to show

greater familiarity with AI tools, and different levels of trust in their use for medical diagnostic.
H5: Education level is significantly associated with the frequency of ChatGPT use. Respondents with higher education are expected to use ChatGPT more frequently in everyday life.
H6: Trust in AI-based diagnostics represents a single latent construct across different diagnostic methods, including MRI, X-ray, and CT interpretation. The hypothesis is tested through reliability analysis and factor analysis.
H7: Trust in AI mediates the relationship between previous experience with AI tools and willingness to use AI for self-diagnosis. This hypothesis aims to test the idea that the previous experience does not influence future willingness only directly, but also indirectly by shaping the respondents' level of trust in AI tools.
In addition to these hypotheses, cluster analysis and decision tree analysis are used as exploratory methods in this paper. To identify the profiles of respondents, cluster analysis was applied, and to explore which variables are most relevant for predicting the willingness to use ChatGPT for self-diagnosis, the chosen and used model is decision tree.

3 Description and results of the conducted survey research

The research was conducted with the aim of analyzing the perception of the use of artificial intelligence in medicine. In order to collect data, a survey was conducted with the aim of examining the impact of sociodemographic factors on patient trust in the use of artificial intelligence in medicine and in self-diagnosis, as well as to identify key factors that influence trust and the willingness of respondents to use artificial intelligence in medicine. A total of 160 responses were collected through the survey, with the oldest respondent born in 1959 and the youngest in 2008. The survey was conducted among respondents living in the Republic of Serbia, in Belgrade. The survey was conducted between July and November 2025, primarily among students of the Faculty of Organizational Sciences. However, it was also distributed beyond this group in order to acquire responses from participants with different sociodemographic characteristics. It is important to emphasize that the distribution by year of birth is not uniform, meaning that there is not an equal number of respondents across different years, which may to some extent affect the results. The male-to-female ratio is evenly distributed, with 50% women and 50% men. Out of all respondents, 47.5% are currently enrolled in undergraduate studies, while 44.4% are not currently formally educating. Out of 160 respondents, 53 of them always use ChatGPT, while only 19 respondents always verify the information that it provides in its answers. Less than half of the respondents, 40.6% to be exact, have used ChatGPT for the purposes of self-diagnosis, of whom only 3 do it frequently. Figure 1 shows frequencies for age, gender, cuurent education, and previous experience in self-diagnosis using ChatGPT.

Variable	Category	Frequency	Percentage
Age group	18-28	87	54%
	29-39	23	14%
	40-50	22	13.7%
	51-57	28	17.5%
Gender	Male	80	50%
	Female	80	50%
Current education	Primary school	1	0.6%
	High school	2	1.25%
	Higher vocational studies	2	1.25%
	Undergraduate studies	76	47.5%
	Master studies	6	3.75%
	Not currently enrolled	73	45%
Previous experience in self-diagnosis using ChatGPT	Yes	65	40.6%
	No	95	59.4%

Fig. 1. Frequency table for variables age, gender, current education, and previous experience in self-diagnosis using ChatGPT

The assumption of the paper is that in the observed area there is still not a high level of trust in artificial intelligence, especially when it comes to medicine. It is also assumed that there is greater trust in the use of artificial intelligence among younger generations compared to older ones, as well as a higher level of use of ChatGPT for self-diagnosis. The level of education may have a relevant impact on trust in the results of artificial intelligence, while gender does not have a significant impact.

4 Hypothesis testing results

4.1 Analysis of the impact of gender on the use of artificial intelligence for diagnostic purposes

In this part of the paper, H1 was tested, and the impact of respondents' gender on trust in the use of artificial intelligence in medical diagnostics was analyzed, as well as the frequency of using ChatGPT for these purposes and in everyday life. The aim is to examine whether there are differences between men and women in terms of: The level of trust in the use of artificial intelligence systems for interpreting different types of medical images and making medical diagnoses; The willingness to use artificial intelligence tools such as ChatGPT for self-diagnosis. Given that ordinal and nominal variables were used for these analyses, the following tests were applied: Mann-Whitney test: to examine differences between men and women in terms of continuous or ordinal variables, such as trust in AI diagnostic tools, as well as the frequency of using ChatGPT and the frequency of its use for self-diagnosis; Chi-Square test of independence: to examine the relationship between gender and categorical variables, such as

experience in using ChatGPT for self-diagnosis and willingness to use it for that purpose in the future.

The results of these tests will allow an assessment of whether the gender of respondents has a statistically significant impact on the use and perception of artificial intelligence tools for medical diagnostic purposes.

Gender * Willingness to use ChatGPT for self-diagnosis

			Willingness to use ChatGPT for self-diagnosis		
			yes	no	Total
Gender	Man	Count	43	37	80
		% within Gender	33.5	46.5	80.0
		%within Willingness to use ChatGPT for self-diagnosis	53.8%	46.3%	100.%
			64.2%	39.8%	50.0%
	Woman	Count	24	56	80
		% within Gender	33.5	46.5	80.0
		%within Willingness to use ChatGPT for self-diagnosis	30.0%	70.0%	100.0%
			35.8%	60.2%	50.0%
Total		Count	67	93	160
		Expected Count	67.0	93.0	160.0
		% within Gender	41.9%	58.1%	100.0%
		%within Willingness to use ChatGPT for self-diagnosis	100.0%	100.0%	100.0%

Fig. 2. Frequency table for variables gender and willingness to use ChatGPT for self-diagnosis

Based on the Chi-Square test conducted on the variables gender and respondents' willingness to use ChatGPT for self-diagnosis, it can be observed that the p-value is equal to 0.002, which indicates that there is a statistically significant difference between genders when it comes to respondents' willingness to use ChatGPT for self-diagnosis.

Based on the frequency table shown in Figure 2, it can be clearly observed that 53.8% of male respondents are willing to use ChatGPT for self-diagnosis, while almost half as many, 30% of female respondents, are willing to use it for these purposes. Given that there is a statistically significant difference, it can be concluded that men are more willing to use artificial intelligence tools to make self-diagnoses.

All other tests that were conducted were not statistically significant, which leads to the conclusion that there is no statistically significant difference between male and female respondents when it comes to their trust in the use of artificial intelligence in the interpretation of MRI (Mann-Whitney U=2807, p=0.126), X-ray (Mann-Whitney U=2925, p=0.283) and CT scans (Mann-Whitney U=2967, p=0.359), as well as in the frequency of self-diagnosis using ChatGPT (Mann-Whitney U=2964, p=0.392) and the frequency of using ChatGPT in general (Mann-Whitney U=2792, p=0.151).

4.2 Analysis of the impact of respondents' age on the use of artificial intelligence for diagnostic purposes

In this part of the paper, the impact of respondents' age on their trust in the use of artificial intelligence in medical diagnostics was analyzed, as well as the frequency of using ChatGPT for these purposes and in everyday life. Through this part, H2 and H3 were tested, with the aim of examining whether there are differences between age in terms of: The level of trust in the use of artificial intelligence systems for interpreting different types of medical images and making medical diagnoses and the willingness to use artificial intelligence tools such as ChatGPT for self-diagnosis.

Age is a continuous variable, so Spearman and Mann-Whitney tests were used to examine its relationship with ordinal and continuous variables. The tests were conducted in the following ways: Mann-Whitney test: to examine the relationship between age and categorical variables, such as experience in using ChatGPT for self-diagnosis and willingness to use it for that purpose in the future; and Spearman test: to examine differences between age and continuous or ordinal variables, such as trust in AI diagnostic tools, as well as the frequency of using ChatGPT and the frequency of its use for self-diagnosis.

The results of these tests will show whether there is a significant difference between respondents' age when it comes to the use and perception of artificial intelligence tools for medical diagnostic purposes.

For examining the relationship between age and trust in the use of artificial intelligence for interpreting MRI, CT and X-ray scans, a Spearman test was conducted and it indicates that there is a statistically significant relationship. The correlation coefficients ($r=0.235$ for trust in MRI; $r=0.298$ for trust in X-ray; $r=0.301$) for trust in using the AI in MRI, X-ray CT scans are slightly positive, which leads to the conclusion that as respondents' age increases, trust in the use of artificial intelligence for these purposes slightly increases.

A Spearman test was also conducted on the variables of respondents' age in years and the frequency of using ChatGPT, as well as the frequency of self-diagnosis using it ($p=0$ for frequency of using ChatGPT; $p=0$ for frequency of using ChatGPT for self-diagnosis), based on which it can be concluded that there is a statistically significant relationship. The correlation coefficients are slightly negative ($r=-0.571$ for frequency of using ChatGPT; $r=0.341$ for frequency of using ChatGPT for self-diagnosis), which indicates that as age increases, the frequency of using ChatGPT and the frequency of self-diagnosis using artificial intelligence decrease.

A Mann-Whitney test was conducted on the variables of respondents' age and the existence of previous experience in self-diagnosis using ChatGPT, and based on the p-value, which is equivalent to 0.005, it can be concluded that there is a statistically significant difference in respondents' behavior on these issues depending on their age. The same test was also conducted to examine the relationship between age and respondents' willingness to use ChatGPT for self-diagnosis, but it proved to be statistically insignificant.

Based on all conducted tests, which were all statistically significant except for the last one examining the relationship between age and willingness to use artificial intelligence for self-diagnosis, it can be concluded that age has an impact on trust in the use of artificial intelligence for self-diagnosis.

4.3 Analysis of the impact of education on the use of AI for diagnostic purposes

In this part of the paper, the impact of respondents' level of education on trust in the use of artificial intelligence tools in medical diagnostics and the frequency of using ChatGPT for these purposes and in everyday life will be analyzed, with the goal of testing the H4 and H5. Respondents were divided into three categories based on education: lower, secondary, and higher. The lower education category includes respondents who have completed or are currently attending primary school, the secondary education category includes respondents who have completed or are currently attending secondary school or college, while the higher education category includes respondents who have completed or are currently attending undergraduate, master's, or doctoral academic studies. The aim of these analyses is to examine whether there are differences between education categories in terms of:

- The level of trust in the use of artificial intelligence systems for interpreting medical images and making diagnoses with the help of artificial intelligence;
- The willingness to perform self-diagnosis using artificial intelligence tools such as ChatGPT.

The education category is an ordinal variable, while the dependent variables are nominal and ordinal, so non-parametric Kruskal-Wallis and Chi-Square tests were used for the analysis in the following ways: Kruskal-Wallis test: used to examine the relationship between the level of education and trust in the use of artificial intelligence for interpreting medical images and making diagnoses, as well as the frequency of using ChatGPT in everyday life and for self-diagnosis; Chi-square test: used to examine the relationship between the level of education and experience in using ChatGPT for self-diagnosis and the willingness to use it for that purpose.

Test Statistics a,b

	Trust of respondents in the use of AI in the interpretation of MRI scans	Trust of respondents in the use of AI in the interpretation of X-Ray scans	Trust of respondents in the use of AI in the interpretation of CT scans
Chi-Square	10.899	5.149	7.517
df	1	1	1
Asymp. Sig.	.001	.023	.006

a. Kruskal Wallis test
b. Grouping variable: Education category

Fig. 3. Kruskal-Wallis test for variables education and trust in use of AI for MRI, X-ray and CT scans

Test Statistics a,b

	Frequency of using ChatGPT
Chi-Square	36.941
df	1
Asymp. Sig.	.000

a. Kruskal Wallis test
b. Grouping variable: Education category

Fig. 4. Kruskal-Wallis test for variables education and frequency of using ChatGPT

Based on the Kruskal-Wallis test, whose results can be seen in Figure 3, which was conducted on the variables of education category and respondents' trust in the use of artificial intelligence for interpreting MRI, X-ray and CT scans and making medical diagnoses, it can be concluded that there is a statistically significant difference among different education groups.

By conducting the Kruskal-Wallis test on the variables of education category and the frequency of using ChatGPT, it can be concluded that in this case as well there is a statistically significant difference between different groups, which is shown on Figure 4.

Education category * Frequency of using ChatGPT Crosstabulation

			Frequency of using ChatGPT					Total
			never	rarely	some-times	often	always	
Education category	Secondary level of education	Count	12	10	10	6	2	40
		Expected Count	4.5	5.0	7.5	9.5	13.5%	40.0
		% within Education category	30.0%	25.0%	25.0%	15.0%	5.0%	100.0%
		% withing Frequency of using ChatGPT	66.7%	50.0%	33.3%	15.8%	3.7%	25.0%
	Higher level of education	Count	6	10	20	32	52	120
		Expected Count	13.5	15.0	22.5	28.5	40.5	120.0
		% within Education category	5.0%	8.3%	16.7%	26.7%	43.3%	100.0%
		% withing Frequency of using ChatGPT	33.3%	50.0%	66.7%	84.2%	96.3%	75.0%

		Frequency of using ChatGPT					Total
		never	rarely	some-times	often	always	
Total	Count	18	20	30	38	54	160
	Expected Count	18.0	20.0	30.0	38.0	54.0	160.0
	% within Education category	11.3%	12.5%	18.8%	23.8%	33.8%	100.0%
	% withing Frequency of using ChatGPT	100.0%	100.0 %	100.0 %	100.0 %	100.0%	100.0%

Fig. 5. Contingency table for variables education and frequency of using ChatGPT

In Figure 5 there is a contingency table, based on which it can be concluded that respondents with a secondary level of education mostly rarely or never use ChatGPT, while respondents with a higher level of education use this tool much more frequently, with responses indicating frequent use being dominant. These results indicate a clear relationship between the level of education and the frequency of using ChatGPT, meaning that as the level of education increases, the willingness of respondents to use artificial intelligence tools also increases.

5 Analysis of the conducted multivariate analysis

5.1 Cluster analysis

Cluster analysis was conducted in R to better understand the differences between groups in terms of the use of and trust in artificial intelligence tools. The aim of this analysis is to identify groups of respondents who have similar behavior patterns and attitudes regarding their level of technological literacy, the use of artificial intelligence, the perception of the accuracy of results provided by ChatGPT, and their tendency and willingness to perform self-diagnosis using artificial intelligence tools.

For the purposes of this cluster analysis, the K-means algorithm and the Elbow method were used to determine the optimal number of clusters, with the goal of minimizing differences within clusters and maximizing differences between clusters.

The selected variables for the cluster analysis include respondents' technological literacy, frequency of using ChatGPT, frequency of verifying the information it provides, frequency of self-diagnosis using the internet, as well as using advice from family and friends, and frequency of self-diagnosis using artificial intelligence tools.

Based on the Elbow method, it can be concluded that the optimal number of clusters is three, and this number will be used for the K-means cluster analysis.

Variable	Cluster 1	Cluster 2	Cluster 3
Technological literacy	-1.1158	0.3064	0.3985
AI usage frequency	-1.2405	0.3357	-0.4468

Variable	Cluster 1	Cluster 2	Cluster 3
Trust in AI	-0.7125	0.7076	-0.1186
Self-diagnosis (internet)	-0.4551	-0.6688	0.7409
Self-diagnosis (family/friends)	-0.3121	-0.5764	0.5938
Self-diagnosis (AI)	-0.7771	-0.1270	0.5255

Fig. 6. Cluster centers based on selected variables

Cluster	Number of respondents
1	39
2	51
3	70

Fig. 7. Distribution of respondents across clusters

In Figures 6 and 7, respondents are divided into three groups, with 39 respondents in the first group, 51 in the second, and 70 in the third. Based on the ratio between_ss and total_ss, it can be concluded that 36.6% of the variability between respondents in these clusters is explained, which indicates that the data are not highly differentiated, but it is still possible to identify certain patterns and group respondents based on their characteristics.

Based on the values of the cluster centers, it is possible to draw conclusions about the characteristics of these groups and categorize them as follows:
First group of respondents (39 respondents):

- Lower level of technological literacy compared to the average
- Rarely use artificial intelligence tools
- Lower trust in the accuracy of results provided by artificial intelligence
- Less frequently use the internet for self-diagnosis
- Rarely rely on advice from family and friends
- Do not frequently use artificial intelligence for making diagnoses

It can be concluded that respondents in this group have a low level of trust in technology and artificial intelligence, as well as a low level of digital engagement.
Second group of respondents (51 respondents):

- Slightly above-average level of technological literacy
- Moderately use artificial intelligence tools
- Moderate trust in the accuracy of results provided by artificial intelligence
- Rarely use the internet for self-diagnosis
- Rarely rely on advice from family and friends
- Moderately use artificial intelligence for making diagnoses

It can be concluded that respondents in this group have a moderate level of trust in and use of artificial intelligence but are not fully reliant on these technologies.

Third group of respondents (70 respondents):

- Higher level of technological literacy
- Frequently use artificial intelligence tools
- Moderate trust in the accuracy of results provided by artificial intelligence
- Frequently use the internet for self-diagnosis
- Frequently rely on advice from family and friends
- Moderately to frequently use artificial intelligence for making diagnoses

It can be concluded that respondents in this group have a higher level of digital literacy and more frequently use technology but still maintain a certain level of caution when it comes to fully trusting the results of artificial intelligence.

Fig. 8. Visualization of clusters obtained using the K-means algorithm

In Figure 8, a graphical representation of the distribution of respondents across clusters can be seen. The figure also shows the cluster centers, and thus the distance of each point (respondent) from its center, which represents the typical profile of a cluster member, since the centers are average values across variables.

5.2 Decision tree

A decision tree is a method used to predict a dependent variable based on a set of independent variables. The dependent variable selected is ai_self_diagosis, which represents the respondents' willingness to use ChatGPT for self-diagnosis, while the independent variables include all other variables. The whole dataset was divided into training and test sets so that the model could learn from one part of the data, which is the test set, and be evaluated on another. The data was split 80-20, in which 80% was assigned to the training set, while the remaining 20% was assigned to the test set. Regardless, the decision tree results should be interpreted with caution, due to the relatively small sample size.

	actual	
predicted	no	yes
no	15	0
yes	3	13

Fig. 9. Confusion matrix of the decision tree model

Accuracy	Precision	Recall	F1
0.9032258	0.8125000	1.0000000	0.8965517

Fig. 10. Classification results of the decision tree model

Based on the confusion matrix, shown on Figure 9, it can be observed that the model correctly predicted 15 cases in which respondents do not use ChatGPT for self-diagnosis, as well as 13 cases in which they do use it. The model made three errors, predicting that respondents use ChatGPT ("yes") in situations where they actually do not ("no"), while it did not make any errors in predicting that respondents do not use ChatGPT in situations where they actually do use it. (see Figure 9)

Based on the evaluation metrics shown in Figure 10, it can be observed that the accuracy of the model is 0.903, which indicates that the model correctly classified over 90% of the cases. Precision is 0.813, which means that about 81% of the cases predicted by the model as belonging to the "yes" category was correctly classified. Recall, also called sensitivity, is 1, which indicates that the model successfully identified all actual cases in which respondents use ChatGPT for self-diagnosis. However, when interpreting precision and recall values together, it suggests that the model is more effective in detecting positive cases than in avoiding false positives, which can indicate the imbalance between the “yes” and “no” category. The F1 score, which represents the balance between precision and recall, is 0.897, indicating a very good overall performance of the model.

Variable	Importance
trust	36.931543
ai_self_diagnosis_experience	14.192774
ai_self_diagnosis_frequency	12.960636
age	11.429034
year_of_birth	11.429034
age_group	10.988238
completed_education	3.498137
trust_ct_ai	3.069590
final_education	2.862112
trust_mri_ai	2.578455
trust_xray_ai	2.578455
current_education	1.640400
ai_accuracy_check	1.067937

Fig. 11. Variable importance in the decision tree model

Based on the results shown in Figure 11, it can be concluded that the variable trust (36.93), which refers to the level of trust in the accuracy of the diagnosis provided by ChatGPT, has the greatest importance in predicting the variable ai_self_diagosis. In addition to this, significant influence is also observed for ai_self_diagnosis_experience (14.19) and ai_self_diagnosis_frequency (12.96), which relate to previous experience and the frequency of using artificial intelligence for making diagnoses. Furthermore, among the more important variables are year of birth (11.43) and age_group (10.99), which indicates the influence of demographic factors. However, the variable-importance values represent raw importance scores generated by the decision tree algorithm, and these values should not be interpreted as percentages or as direct measures of explained variance, but should be interpreted as relative indicators of how much each variable contributed to the splitting process and classification performance within this specific model.

5.3 Factor analysis

In order to examine the structure of variables related to trust in the use of artificial intelligence in medical diagnostics, and to test the H6, it was necessary to conduct a factor analysis. The reliability of the scale was assessed using the Cronbach alpha coefficient. Within this analysis, three variables were used that relate to respondents' trust in the application of artificial intelligence in different medical diagnostic methods, namely trust in MRI analysis, trust in X-ray analysis, and trust in CT analysis. This analysis was conducted to examine whether the type of medical technology has an impact on the level of trust that respondents have, or whether it reflects their general attitude toward artificial intelligence in medical diagnostics.

```
raw_alpha std.alpha G6(smc) average_r S/N   ase mean sd
median_r

      0.96      0.96    0.94      0.89  24 0.0056 2.1 0.6     0.9
```

Fig. 12. Reliability analysis results (Cronbach's alpha)

A reliability analysis was conducted, and its results can be seen in Figure 12. Based on these results, it can be concluded that the scale has exceptionally high internal consistency, given that the value of Cronbach $\alpha = 0.96$, which indicates a strong relationship between the variables and the fact that they measure the same construct.

```
Standardized loadings (pattern matrix) based upon corre-
lation matrix

          MR1   h2  u2  com

mri_ai    0.93 0.87 0.128  1
xray_ai   0.92 0.84 0.156  1
ct_ai     0.98 0.95 0.049  1
```

Fig. 13. Interpretation of reliability analysis results

```
          MR1

SS loadings  2.67
Proportion Var 0.89
```

Fig. 14. Results of factor analysis (standardized loadings)

The results of the factor analysis are shown in Figure 13, and they indicate that there is one dominant factor. All observed variables have exceptionally high factor loadings, which further confirms their strong connection with a common dimension.

One factor explains as much as 89% of the total variability shown in the Figure 14. This is evidence of a high level of homogeneity and indicates that most of the differences in respondents' answers can be explained by a single latent construct, meaning that it relates to general trust in the application of artificial intelligence in medical diagnostics. This shows that respondents do not significantly distinguish between different diagnostic methods and technologies but rather form a unified attitude toward artificial intelligence as a whole when it comes to medical diagnostics. (see Figure 15)

```
Parallel analysis suggests that the number of factors =
1  and the number of components =  1
```

Fig. 15. Results of parallel analysis for determining the number of factors

The results of the factor analysis are significant because they enable simplification of the analysis and confirm that a single construct related to trust is sufficient. It can help

in more precise modeling of the factors that influence respondents' attitudes toward using artificial intelligence for self-diagnosis.

5.4 Mediation analysis

In order to understand the relationship between respondents' previous experience and their willingness to use artificial intelligence for self-diagnosis, and to test H7, a mediation analysis was conducted. The main goal of this analysis was to examine whether trust in artificial intelligence represents a mediating variable that could explain the way in which previous experience influences respondents' behavior. It was examined whether previous experience affects the level of trust, and then whether that level of trust affects the willingness of respondents to use artificial intelligence tools.

```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)     3.657      0.3000  12.193  < 2e-16 ***
ai_experience  -0.859      0.1787  -5.014  1.42e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1
‘ ’ 1
```

Fig. 16. Regression results for the effect of experience on trust

The results of the linear model show that previous experience in using artificial intelligence tools has a statistically significant impact on the level of trust of respondents, which is shown in Figure 16. This result indicates the existence of a significant relationship between experience and trust, where differences in experience lead to differences in the level of trust.

```
                     Estimate 95% CI Lower 95% CI Upper p-value
ACME (control)       0.135289     0.058550      0.225017
<2e-16 ***
ACME (treated)       0.192894     0.130830      0.252181
<2e-16 ***
ADE (control)        0.085214      0.013588      0.140331
0.028 *
ADE (treated)        0.142819     0.018153       0.216371
0.028 *
Total Effect         0.278108     0.203091      0.335538
<2e-16 ***
Prop. Mediated (control) 0.486463  0.237281  0.908706
<2e-16 ***
Prop. Mediated (treated) 0.693592 0.504720  0.949490
<2e-16 ***
ACME (average)       0.164091     0.101005      0.235689
<2e-16 ***
```

```
ADE (average)         0.114071     0.016850      0.173164
0.028 *
Prop. Mediated (average) 0.590028 0.406900  0.927213
<2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1
‘ ’ 1


Sample Size Used: 160
```

Fig. 17. Mediation analysis results showing indirect and direct effects

The results of the mediation analysis, shown in Figure 17, indicate that there is a statistically significant indirect effect of previous experience on respondents' willingness to use artificial intelligence tools through trust, as evidenced by the value ACME = 0.164. The direct effect of experience is also statistically significant, as shown by the value ADE = 0.114. The total effect is 0.278, while the analysis of the proportion mediated indicates that trust explains approximately 59% of the total effect of experience on respondents' willingness to use artificial intelligence in medical diagnostics.

6 Conclusion

The aim of this paper was to identify the key factors that influence trust and respondents' willingness to use artificial intelligence in medical diagnostics, which was examined through the application of statistical and advanced analytical methods. The results indicate that sociodemographic factors, primarily age and level of education, have a significant impact on the way artificial intelligence tools are used and perceived in medical diagnostics, as well as in everyday life, while gender has an impact only when it comes to trusting the AI usage in medical diagnostics, where the conclusion is that males are more willing to use it than females.

A special contribution of this paper is the integration and combination of different approaches, including cluster analysis, machine learning models, and mediation analysis, which made it possible to gain a deeper understanding of respondents' behavior. Cluster analysis identified three clearly defined groups of respondents with different levels of trust and use of AI tools, while predictive models identified trust as the most important factor in decision-making regarding the use of artificial intelligence. In addition to its central role in shaping respondents' attitudes, trust also acts as a mediator between previous experience and willingness to use artificial intelligence. Based on this, it was concluded that user behavior cannot be explained solely by the direct influence of experience, but also by the way trust in technology is formed.

The scientific contribution of this paper lies in combining statistical and advanced analytical methods to better understand user behavior in the context of artificial intelligence in medical diagnostics, with a particular focus on the role of trust. The results

can serve as a starting point for future research focused on modeling the adoption of AI systems, as well as for developing approaches aimed at increasing user trust and acceptance. In practice, these findings can be useful in the design and implementation of AI solutions in healthcare, especially in improving transparency and adapting systems to user needs. Future studies could expand this research by including additional variables, larger samples, and longitudinal data, to gain a deeper understanding of how user attitudes change over time.

Disclosure of Interests. The author has no competing interests to declare that are relevant to the content of this article.

References

1. Ahmad, H., Milne, M., Buchlak, Q., Ektas, N., Sanderson, G., Chamtie, H., Karunasena, S., Chiang, J., Holt, X., Tang, C., Seah, J., Bottrell, G., Esmaili, N., Brotchie, P., Jones, C. (2023). Machine learning augmented interpretation of chest X-rays: A systematic review. *Diagnostics, 13*. https://doi.org/10.3390/diagnostics13040743
2. Alasmari, A. A., Alruwaili, R. F., Alotaibi, R. F., Youssef, I. K., & Asklany, S. (2025). Demographic influences on trust in artificial intelligence across cognitive domains: A statistical perspective. PLOS One, 20. https://doi.org/10.1371/journal.pone.0331003
3. Ali, S., Akhlaq, F., Imran, A., Kastrati, Z., Daudpota, S. M., & Moosa, M. (2023). The enlightening role of explainable artificial intelligence in medical & healthcare domains: A systematic literature review. Computers in biology and medicine, 166, 107555. https://doi.org/10.1016/j.compbiomed.2023.107555
4. Alowais, S., Alghamdi, S., Alsuhebany, N., Alqahtani, T., Alshaya, A., Almohareb, S., Aldairem, A., Alrashed, M., Saleh, K., Badreldin, H., Yami, M., Harbi, S., Albekairy, A. (2023). Revolutionizing healthcare: The role of artificial intelligence in clinical practice. *BMC Medical Education, 23*. https://doi.org/10.1186/s12909-023-04698-z
5. Al-Qaness, M., Zhu, J., Al-Alimi, D., Dahou, A., Alsamhi, S., Elaziz, M., Ewees, A. (2024). Chest X-ray images for lung disease detection using deep learning techniques: A comprehensive survey. *Archives of Computational Methods in Engineering, 31*, 3267–3301. https://doi.org/10.1007/s11831-024-10081-y
6. Anderson, P. G., Tarder-Stoll, H., Alpaslan, M., Keathley, N., Levin, D. L., Venkatesh, S., Bartel, E., Sicular, S., Howell, S., Lindsey, R. V., Jones, R. M. (2024). Deep learning improves physician accuracy in the comprehensive detection of abnormalities on chest X-rays. *Scientific Reports, 14*, 25151. https://doi.org/10.1038/s41598-024-76608-2
7. Aras, S., Drakos, C., Manimangalam, V., Nasir, M. A., Burns, C., Smith, D., & Equils, O. (2026). Influencing public acceptance of artificial intelligence (AI) in healthcare delivery. Frontiers in Digital Health, 7. https://doi.org/10.3389/fdgth.2025.1664345
8. Astobiza, A. M., Alonso, M., & Lozano, R. O. (2025). Trust and AI in healthcare: a systematic review. Monash Bioethics Review. https://doi.org/10.1007/s40592-025-00272-z
9. Briganti, G., & Le Moine, O. (2020). *Artificial intelligence in medicine: Today and tomorrow*. Frontiers in Medicine, **7**, 27. https://doi.org/10.3389/fmed.2020.00027
10. Busch, F., Hoffmann, L., Xu, L., Zhang, L. J., Hu, B., García-Juárez, I., Toapanta-Yanchapaxi, L. N., Gorelik, N., Gorelik, V., Rodriguez-Granillo, G. A., Ferrarotti, C., Cương, N. N., Thi, C. A., Tuncel, M., Kaya, G., Solís-Barquero, S. M., Avila, M. M. C., Ivanova, N. G., Kitamura, F. C., . . . Zhang, S. (2025). Multinational Attitudes Toward AI in Health Care

and Diagnostics Among Hospital Patients. JAMA Network Open, 8. https://doi.org/10.1001/jamanetworkopen.2025.14452
11. Carini, C., & Seyhan, A. A. (2024). *Tribulations and future opportunities for artificial intelligence in precision medicine*. Journal of Translational Medicine, 22(1), 411. https://doi.org/10.1186/s12967-024-05067-0
12. Cflow. (2025, November 20). AI in clinical workflows: How healthcare providers can improve efficiency and patient care. https://www.cflowapps.com/ai-for-clinical-workflows/
13. David-Olawade, A. C., Olawade, D. B., Vanderbloemen, L., Rotifa, O. B., Fidelis, S. C., Egbon, E., Akpan, A. O., Adeleke, S., Ghose, A., & Boussios, S. (2025). *AI-driven advances in low-dose imaging and enhancement—A review*. Diagnostics, 15(6), 689. https://doi.org/10.3390/diagnostics15060689
14. Dingel, J., Kleine, A., Cecil, J., Sigl, A., Lermer, E., & Gaube, S. (2024). Predictors of Health Care Practitioners' Intention to Use AI-Enabled Clinical Decision Support Systems: Meta-Analysis Based on the Unified Theory of Acceptance and Use of Technology. Journal of Medical Internet Research, 26. https://doi.org/10.2196/57224
15. Fotaki, A., Puyol-Antón, E., Chiribiri, A., Botnar, R., Pushparajah, K., & Prieto, C. (2022). *Artificial intelligence in cardiac MRI: Is clinical adoption forthcoming?* Frontiers in Cardiovascular Medicine, 8, 818765. https://doi.org/10.3389/fcvm.2021.818765
16. Goktas, P., & Grzybowski, A. E. (2025). Shaping the Future of Healthcare: Ethical Clinical Challenges and Pathways to Trustworthy AI. Journal of Clinical Medicine, 14. https://doi.org/10.3390/jcm14051605
17. Google Forms. (2025). Survey on the application of artificial intelligence in medical diagnostics. https://docs.google.com/forms/d/e/1FAIpQLScP35GAW-vuK1N-T_Q_gSg2w19ZfeCJu4mZcWYfCxwCFWxOqQ/viewform
18. Harada, Y., Suzuki, T., Harada, T., Sakamoto, T., Ishizuka, K., Miyagami, T., Kawamura, R., Kunitomo, K., Nagano, H., Shimizu, T., Watari, T. (2024). Performance evaluation of ChatGPT in detecting diagnostic errors and their contributing factors: An analysis of 545 case reports of diagnostic errors. *BMJ Open Quality, 13*, e002654. https://doi.org/10.1136/bmjoq-2023-002654
19. Haupt, M., Maurer, M. H., & Thomas, R. P. (2025). *Explainable artificial intelligence in radiological cardiovascular imaging—A systematic review*. Diagnostics, 15(11), 1399. https://doi.org/10.3390/diagnostics15111399
20. Hirani, R., Noruzi, K., Khuram, H., Hussaini, A. S., Aifuwa, E. I., Ely, K. E., Lewis, J. M., Gabr, A. E., Smiley, A., Tiwari, R. K., & Etienne, M. (2024). *Artificial intelligence and healthcare: A journey through history, present innovations, and future possibilities*. Life, 14(5), 557. https://doi.org/10.3390/life14050557
21. Hoppe, J. M., Auer, M. K., Strüven, A., Massberg, S., & Stremmel, C. (2024). *ChatGPT with GPT-4 outperforms emergency department physicians in diagnostic accuracy: Retrospective analysis*. Journal of Medical Internet Research, 26, e56110. https://doi.org/10.2196/56110
22. Hung, M., Yevseyevich, D., Khazana, M., Schwartz, C., & Lipsky, M. S. (2025). *Charting new territory: AI applications in dental caries detection from panoramic imaging*. Dentistry Journal, 13(8), 366. https://doi.org/10.3390/dj13080366
23. Ibrahim, F., Münscher, J.-C., Daseking, M., & Telle, N. (2025). The technology acceptance model and adopter type analysis in the context of artificial intelligence. Frontiers in Artificial Intelligence, 7. https://doi.org/10.3389/frai.2024.1496518
24. Javaid, M., Haleem, A., Singh, R., Suman, R., & Rab, S. (2022). Significance of machine learning in healthcare: Features, pillars and applications. *International Journal of Intelligent Networks, 3*, 58–73. https://doi.org/10.1016/j.ijin.2022.05.002

25. Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology, 2*, 230–243. https://doi.org/10.1136/svn-2017-000101
26. Jiang, Y., Edwards, A. V., & Newstead, G. M. (2021). *Artificial intelligence applied to breast MRI for improved diagnosis*. Radiology, 298(1), 38–46. https://doi.org/10.1148/radiol.2020200292
27. Kauttonen, J., Rousi, R. A., & Alamäki, A. (2025). Trust and Acceptance Challenges in the Adoption of AI Applications in Health Care: Quantitative Survey Analysis. Journal of Medical Internet Research, 27. https://doi.org/10.2196/65567
28. Khalid, N., Qayyum, A., Bilal, M., Al-Fuqaha, A., & Qadir, J. (2023). Privacy-preserving artificial intelligence in healthcare: Techniques and applications. *Computers in Biology and Medicine, 158*, 106848. https://doi.org/10.1016/j.compbiomed.2023.106848
29. Khan, M. M., Shah, N., Shaikh, N., Thabet, A., Alrabayah, T., & Belkhair, S. (2024). Towards secure and trusted AI in healthcare: A systematic review of emerging innovations and ethical challenges. International journal of medical informatics, 195, 105780. https://doi.org/10.1016/j.ijmedinf.2024.105780
30. Khalifa, M., & Albadawy, M. (2024). *AI in diagnostic imaging: Revolutionising accuracy and efficiency*. Computer Methods and Programs in Biomedicine Update, 5, 100146. https://doi.org/10.1016/j.cmpbup.2024.100146
31. Kraus, M., Anteby, R., Konen, E., Eshed, I., & Klang, E. (2023). Artificial intelligence for X-ray scaphoid fracture detection: A systematic review and diagnostic test accuracy meta-analysis. *European Radiology, 34*, 4341–4351. https://doi.org/10.1007/s00330-023-10473-x
32. Kumar, Y., Koul, A., Singla, R., & Ijaz, M. (2022). Artificial intelligence in disease diagnosis: A systematic literature review, synthesizing framework and future research agenda. *Journal of Ambient Intelligence and Humanized Computing, 14*, 8459–8486. https://doi.org/10.1007/s12652-021-03612-z
33. Loh, H., Ooi, C., Seoni, S., Barua, P., Molinari, F., & Acharya, U. (2022). Application of explainable artificial intelligence for healthcare: A systematic review of the last decade (2011–2022). *Computer Methods and Programs in Biomedicine, 226*, 107161. https://doi.org/10.1016/j.cmpb.2022.107161
34. Mahamud, E., Fahad, N., Assaduzzaman, M., Zain, S. M., Goh, K. O. M., & Morol, M. K. (2024). *An explainable artificial intelligence model for multiple lung diseases classification from chest X-ray images using fine-tuned transfer learning*. Decision Analytics Journal, 12, 100499. https://doi.org/10.1016/j.dajour.2024.100499
35. Méndez-Suárez, M., Monfort, A., & Hervás-Oliver, J.-L. (2023). Are you adopting artificial intelligence products? Social-demographic factors to explain customer acceptance. European Research on Management and Business Economics. https://doi.org/10.1016/j.iedeen.2023.100223
36. Moravec, V., Hynek, N., Gavurová, B., & Kubák, M. (2024). Everyday artificial intelligence unveiled: Societal awareness of technological transformation. Oeconomia Copernicana. https://doi.org/10.24136/oc.2961
37. Nair, A., Ong, W. Y., Lee, A. Y. Y., Leow, N. W., Makmur, A., Ting, Y. H., Lee, Y. S., Ong, S. H., Tan, J. L., Kumar, N., & Hallinan, J. T. P. D. (2025). *Enhancing radiologist productivity with artificial intelligence in magnetic resonance imaging (MRI): A narrative review*. Diagnostics, 15(9), 1146. https://doi.org/10.3390/diagnostics15091146
38. Nazir, S., Dickson, D., & Akram, M. (2023). Survey of explainable artificial intelligence techniques for biomedical imaging with deep neural networks. Computers in biology and medicine, 156, 106668. https://doi.org/10.1016/j.compbiomed.2023.106668

39. Nia, N. G., Kaplanoğlu, E., & Nasab, A. (2023). *Evaluation of artificial intelligence techniques in disease diagnosis and prediction*. Discover Artificial Intelligence, 3, 5. https://doi.org/10.1007/s44163-023-00049-5
40. Nouis, S. C., Uren, V., & Jariwala, S. (2025). Evaluating accountability, transparency, and bias in AI-assisted healthcare decision- making: a qualitative study of healthcare professionals' perspectives in the UK. BMC Medical Ethics, 26. https://doi.org/10.1186/s12910-025-01243-z
41. Pesapane, F., Giambersio, E., Capetti, B., Monzani, D., Grasso, R., Nicosia, L., Rotili, A., Sorce, A., Meneghetti, L., Carriero, S., Santicchia, S., Carrafiello, G., Pravettoni, G., & Cassano, E. (2024). Patients' Perceptions and Attitudes to the Use of Artificial Intelligence in Breast Cancer Diagnosis: A Narrative Review. Life, 14. https://doi.org/10.3390/life14040454
42. Rodler, S., Kopliku, R., Ulrich, D., Kaltenhauser, A., Casuscelli, J., Eismann, L., Waidelich, R., Buchner, A., Butz, A., Cacciamani, G., Stief, C., & Westhofen, T. (2023). Patients' Trust in Artificial Intelligence-based Decision-making for Localized Prostate Cancer: Results from a Prospective Trial.. European urology focus. https://doi.org/10.1016/j.euf.2023.10.020
43. Saha, A., Bosma, J. S., Twilt, J. J., et al. (2024). *Artificial intelligence and radiologists in prostate cancer detection on MRI (PI-CAI): An international, paired, non-inferiority, confirmatory study*. The Lancet Oncology, 25(7), 879–887. https://doi.org/10.1016/S1470-2045(24)00220-1
44. Shafi, I., Fatima, A., Afzal, H., Díez, I., Lipari, V., Breñosa, J., & Ashraf, I. (2023). *A comprehensive review of recent advances in artificial intelligence for dentistry e-health*. Diagnostics, 13(13), 2196. https://doi.org/10.3390/diagnostics13132196
45. Sharma, S. (2023). *Artificial intelligence for fracture diagnosis in orthopedic X-rays: Current developments and future potential*. SICOT-J, 9, 21. https://doi.org/10.1051/sicotj/2023018
46. Sheth, D., & Giger, M. L. (2020). *Artificial intelligence in the interpretation of breast cancer on MRI*. Journal of Magnetic Resonance Imaging, 51(5), 1310–1324. https://doi.org/10.1002/jmri.26878
47. Thaldar, D., & Bottomley, D. (2025). Public trust of AI in healthcare in South Africa: results of a survey. BMC Medical Ethics, 26. https://doi.org/10.1186/s12910-025-01272-8
48. Tiwari, A., Mishra, S., & Kuo, T. R. (2025). *Current AI technologies in cancer diagnostics and treatment*. Molecular Cancer, 24(1), 159. https://doi.org/10.1186/s12943-025-02369-9
49. Zhang, J., & Zhang, Z.-M. (2023). Ethics and governance of trustworthy medical artificial intelligence. BMC Medical Informatics and Decision Making, 23. https://doi.org/10.1186/s12911-023-02103-9
50. Zuhair, V., Babar, A., Ali, R., Oduoye, M. O., Noor, Z., Chris, K., Okon, I. I., & Rehman, L. U. (2024). *Exploring the impact of artificial intelligence on global health and enhancing healthcare in developing nations*. Journal of Primary Care & Community Health, 15, 21501319241245847. https://doi.org/10.1177/21501319241245847

THEMATIC SECTION

Designing Resilient and Intelligent Organizations in Times of Permanent Disruption

Resilient and Intelligent Organizations
in Times of Permanent Disruption

SYMORG 2026 • CONFERENCE PROCEEDINGS

Original Research Article *https://doi.org/10.5281/zenodo.23475178*

Resilience as Organizing When Structure Fails

Ana Aleksić Mirić[1][0000-0002-9368-6501] and Maksim Vojvodić[2][0009-0001-5254-8583]

[1] University of Belgrade, Faculty of Economics and Business, 11000 Belgrade, Serbia
[2] Independent researcher
anaa@ekof.bg.ac.rs

Abstract. Organizational resilience is commonly conceptualized as a capability embedded in structures, routines, and organizational design. However, this perspective assumes that such structures retain their coordinating capacity under conditions of disruption. Drawing on empirical findings from a large-scale study of organizations during the COVID-19 pandemic, where organizational design and prior experience did not significantly explain resilience, we challenge this assumption.

We develop a process-based reconceptualization of organizational resilience as an organizing capability that emerges when disruption undermines the ongoing enactment of structure. Building on process and practice-based perspectives, we argue that organizational structures derive their coordinating capacity from recurrent enactment and that disruption interrupts these patterns, dissolving the stability of structure. Under such conditions, resilience emerges through the active reconstitution of organizing, specifically through the reconstruction of meaning, coordination, and action. By shifting attention from design to organizing, the paper offers a novel framework for understanding how organizations maintain functioning under conditions in which established structures no longer hold.

Keywords: Organizational Resilience, Organizing, Structural Enactment.

1 Introduction

Organizations increasingly operate in environments characterized by disruption, uncertainty, and rapid change. From global crises such as the COVID-19 pandemic to technological shifts and market volatility, the ability of organizations to maintain functioning under adverse conditions has become a more pressing concern in management research. This concern is typically captured through research on the concept of organizational resilience, which broadly refers to the capacity of organizations to withstand, adapt to, and potentially benefit from disruptive events (Linnenluecke, 2017; Vogus & Sutcliffe, 2007).

A dominant stream of research conceptualizes resilience as an organizational capability embedded in structures, routines, and human resource systems that enable organizations to absorb shocks and respond effectively to disruption (Lengnick-Hall, Beck, & Lengnick-Hall, 2011; Vogus & Sutcliffe, 2007). Within this perspective, resilience is treated as something organizations possess, and thus can be developed in advance through design, experience, and accumulated capabilities. However, this assumption becomes difficult to sustain when organizations demonstrate resilience precisely in situations where structure should matter most but appears not to.

Observations from large-scale disruptions, such as the COVID-19 pandemic, suggest that organizations can demonstrate relatively high levels of resilience even under abrupt and unprepared transitions to remote work. Unexpectedly, traditional organizational design variables, as well as prior experience with remote work, do not consistently exhibit a significant impact on resilience outcomes. Instead, resilience appears to be more strongly associated with factors operating at the level of individual and collective action (Aleksić Mirić et al., 2022). This empirical anomaly raises a more fundamental question than that typically addressed in resilience research. Rather than asking what factors predict resilience, we ask: what is organizational resilience when established structures and prior experience fail to explain it?

This question points to a deeper theoretical tension. If resilience is embedded in organizational design, why does it persist when the coordinating capacity of that design appears to break down? And how is it possible, in highly structured and increasingly digitalized organizational environments, for structure to exhibit such limited explanatory power precisely under conditions of large-scale disruption?

We argue that answering these questions requires reframing how organizational structure itself is understood. Rather than viewing structure as a stable, enduring property of organizations, we build on process and practice-based perspectives in organizational theory, which emphasize that organizational structures are not fixed entities but are continuously constituted through ongoing action (Giddens, 1984; Weick, 1979; Feldman & Pentland, 2003). From this standpoint, the effectiveness of structure depends on its sustained enactment over time (Orlikowski, 2000; Tsoukas & Chia, 2002). Disruption, therefore, does not simply challenge organizational structures; it interrupts the recurrent patterns of action through which those structures are reproduced. When these patterns are destabilized, structure does not disappear, but it loses its capacity to coordinate action. This shift is critical. If structure loses its coordinating power under disruption, then resilience cannot be found in structure itself.

This matters practically as much as it does theoretically. Much of managerial thinking on resilience has emphasized investment and allocation of resources into structural elements, backup systems, and stronger routines, all of which may themselves be undermined by a large-scale disruption (Lengnick-Hall, Beck, & Lengnick-Hall, 2011; Vogus & Sutcliffe, 2007). Such situations test not only whether an organization has a strong structure, but also how organizing is achieved. Overinvestment into structure as the backbone of resilience may leave organizing capabilities underdeveloped, thus making it more difficult for the organization to regain coordination during disruption.

We therefore reconceptualize organizational resilience not as a design-embedded capability, but as an organizing capability that emerges when the ongoing enactment of structure is disrupted. Specifically, we define organizational resilience as the capacity to reconstitute meaning, coordination, and action when the ongoing enactment of structure is disrupted. In this view, resilience is not something organizations have, but something they do through processes of sensemaking, interaction, and adaptive action when existing structures no longer provide reliable guidance (Weick, 1995).

This perspective also reframes what it means to describe an organization as "resilient." We do not argue that it is incorrect to say, "this is a resilient organization." Rather, we argue that such a label should not be understood as attributing a stable property embedded in organizational design, but rather as a shorthand for the organization's demonstrated capacity to organize effectively.

This paper makes three contributions. First, it challenges structure-centric explanations of resilience by demonstrating that organizational design and prior experience may not adequately account for resilience under conditions of disruption. Second, it advances an organizing perspective that locates resilience in the active reconfiguration of meaning, coordination, and action, rather than in pre-existing capabilities. Third, it introduces the disruption of structural enactment as a mechanism that explains how and why the coordinating capacity of structure breaks down, thereby creating the conditions under which resilience emerges.

By reframing resilience as an enacted organizing capability, this paper contributes to a more processual understanding of how organizations respond to disruption and provides a foundation for future research on the dynamics of organizing under conditions in which established structures no longer provide reliable guidance.

2 The Promise of Structure in Resilience Research

A central assumption in organizational resilience research is that resilience is rooted in the structural and experiential foundations of the organization. Organizational resilience has emerged as a central construct in management research, reflecting the growing importance of understanding how organizations respond to disruption, uncertainty, and adverse events. Broadly, resilience refers to the capacity of organizations to withstand, adapt to, and potentially benefit from challenging conditions (Vogus & Sutcliffe, 2007; Linnenluecke, 2017). However, despite increasing scholarly attention, the concept remains theoretically fragmented, with multiple and sometimes competing definitions across disciplines (Linnenluecke, 2017; Annarelli & Nonino, 2016).

Early conceptualizations of resilience emphasized recovery and stability, focusing on the ability of organizations to "bounce back" to prior states following disruption. More recent perspectives have shifted toward dynamic interpretations, emphasizing adaptation, transformation, and learning in response to changing conditions (Linnenluecke, 2017). Yet, even within these more dynamic perspectives, there is limited consensus regarding the underlying mechanisms through which resilience emerges in practice.

Within the dominant capability-based perspective, resilience is understood as rooted in the structural and experiential foundations of the organization (Lengnick-Hall et al., 2011; Linnenluecke, 2017). Specifically, several key antecedents of resilience are consistently emphasized in literature.

First, organizational design and structure are seen as primary carriers of resilience. Formalized roles, clearly defined authority relations, and coordinated processes are expected to provide stability under conditions of uncertainty, ensuring continuity of operations when external conditions become volatile. Structural arrangements are assumed to enable efficient information flows, coordinated decision-making, and reliable execution of tasks even under pressure (Vogus & Sutcliffe, 2007).

Second, organizational routines and practices are identified as critical mechanisms through which resilience is enacted. Established routines allow organizations to rely on previously validated patterns of action, reducing the need for improvisation in moments of disruption. Over time, these routines become repositories of organizational

knowledge, enabling consistent and predictable responses to recurring challenges (Feldman & Pentland, 2003). Third, human resource systems and behavioral capabilities play a central role in supporting resilience. Strategic HR practices are seen as cultivating cognitive flexibility, behavioral adaptability, and supportive organizational contexts, which together enable individuals and teams to respond effectively to changing conditions (Lengnick-Hall et al., 2011). Fourth, prior experience and organizational learning are assumed to strengthen resilience over time. Organizations that have previously encountered disruptions are expected to codify lessons learned, refine coordination mechanisms, and institutionalize adaptive practices. In this sense, resilience is understood not only as a function of current design but also as an outcome of accumulated experience and learning (Sutcliffe & Vogus, 2003).

Taken together, these factors form a coherent theoretical expectation: resilience is embedded in the design, routines, and experience of the organization. Organizations that are better structured, more routinized, and more experienced should, therefore, be better equipped to withstand and respond to disruption. This expectation becomes even stronger in contemporary organizational contexts characterized by increasing digitalization, formalization, and system integration. Advanced information systems, standardized processes, and technologically-enabled coordination mechanisms are designed precisely to ensure continuity of operations across time and space. Under such conditions, structure essentially serves as the primary carrier of organizational resilience. This capability-based perspective has been instrumental in shifting resilience beyond a purely reactive notion of recovery toward a more proactive understanding of preparedness and adaptation. However, it also implicitly locates resilience within pre-existing organizational arrangements, suggesting that resilience is something organizations possess prior to disruption. As a result, research has largely focused on identifying antecedents of resilience, such as slack resources, routines, and experience, rather than examining how resilience is enacted in real time under conditions of disruption (Linnenluecke, 2017; Annarelli & Nonino, 2016).

In contrast to capability-based views, process-oriented perspectives in organizational theory emphasize that organizations are ongoing accomplishments of organizing. From this standpoint, coordinated action depends on continuous processes of interpretation, interaction, and enactment (Weick, 1979, 1995). Sensemaking plays a central role in this process, as actors construct meaning and develop shared understandings that enable coordinated action under conditions of ambiguity and uncertainty (Weick, 1995).

Research on high-reliability organizations further highlights the importance of ongoing cognitive and relational processes, such as mindfulness, attention to failure, and continuous updating, in maintaining functioning under challenging conditions (Weick & Sutcliffe, 2001). Similarly, work on organizational responses to adversity emphasizes that resilience is not solely a function of pre-existing resources but involves dynamic processes of adjustment and coordination as organizations confront disruption (Williams et al., 2017).

However, even these process-oriented perspectives do not fully depart from structural assumptions. While they highlight the importance of ongoing organizing, they often stop short of specifying how disruption alters the conditions under which organizing occurs, particularly in relation to the role of organizational structure.

Insights from research on organizational routines and micro foundations further complicate the picture. Routines consist of both ostensive aspects (formal rules and understandings) and performative aspects (patterns of action in practice), with the latter continuously shaping and transforming the former (Feldman & Pentland, 2003). This implies that what appears to be a stable structure is, in fact, the outcome of ongoing enactment.

Micro foundational perspectives further reinforce this view by emphasizing that organizational outcomes are generated through the actions and interactions of individuals (Felin, Foss, & Ployhart, 2015). From this standpoint, structures do not act independently but are sustained through the behaviors and interactions of organizational members. Similarly, research on organizational learning highlights the tension between exploitation (reliance on established routines) and exploration (adaptation and experimentation), suggesting that disruption may render existing routines insufficient and necessitate shifts toward more adaptive forms of organizing (March, 1991).

These perspectives suggest that the effectiveness of organizational structure depends not only on its design but on its ongoing enactment in practice. Yet, despite these advances, existing research provides limited insight into how resilience emerges when established structures fail. This limitation becomes particularly salient when considering conditions of large-scale disruption. If resilience is embedded in organizational structures, routines, and experience, then such moments should reveal, rather than obscure, the effects of structure. When they do not, the issue is no longer one of boundary conditions or incomplete specification, it points to a different source of resilience altogether. The theoretical arguments developed in this paper are informed by observations from a large-scale disruption, the COVID-19 pandemic, which provide a revealing context for examining organizational resilience. Evidence from such contexts suggests that organizations can sustain relatively high levels of functioning even during abrupt and unprepared transitions to remote work (Aleksić Mirić et al., 2022).

3 Empirical Motivation: The COVID Study Anomaly

The empirical background of the study is based on a large-scale survey conducted in Serbia between March and May 2021, during a period in which many organizations had already experienced significant disruption caused by the COVID-19 pandemic and the shift to working from home. The sample included 2,171 respondents employed in organizations of different type, size, ownership structure, and age. Respondents were categorized according to key individual and work-related characteristics, including age, gender, marital status, education, occupation, work experience, and family status.

Organizational resilience was measured through a seven-item scale capturing both individual and organizational aspects of resilient behavior, including employees' ability to cope with unusual work situations, openness to new work methods, colleagues' problem-solving capacity, managerial support for resourcefulness, organizational flexibility, and freedom to deviate from established procedures when circumstances require it. The results indicate that respondents reported, on average, a relatively high level of resilience, with all items scoring above the neutral mean. The highest-rated aspect of resilience was the ability to apply new work methods when situations or requirements

changed (M = 4.12), suggesting a strong adaptive orientation among employees. The lowest-rated aspect was organizational flexibility as opposed to rigidity (M = 3.64), indicating that while individuals showed readiness to adapt, organizational structures and procedures were perceived as somewhat less flexible. These findings suggest that resilience during the pandemic in Serbia was strongly grounded in employees' adaptive capacity, while organizational flexibility remained a more challenging dimension of crisis response (Aleksić Mirić et al., 2022).

Two findings from the study proved particularly consequential for understanding organizational resilience. Firstly, organization design properties (job design, unit grouping, authority delegation, and coordination mechanisms) have shown no statistically significant effect on the resilience of that organization. Secondly, prior experience with remote work did not yield a statistically significant difference in resilience outcomes, contrary to established assumptions in organizational learning. Organizations and individuals with prior remote work experience were no more resilient than those working remotely for the first time.

These findings challenge popular theories on structure-based organizational resilience and learning, and raise two crucial questions: If structural preparedness and experience do not explain resilience under a large-scale external disruption, then what does? Secondly, why does structure show no effect on resilience when it should arguably matter the most?

4 When Structure Fails

Understanding how structure becomes ineffective under large-scale disruption requires reconsidering how organizational structure is conceptualized. Additionally, by focusing on structure and its resources (such as formal emergency procedures, backups, etc.) it becomes increasingly difficult to conceptualize resilience as anything but structurally embedded. Results of the COVID study invite a reevaluation of how the relationship between structure and resilience is conceptualized. In order to properly understand the results of the study, the view of organizational structure as a stable entity should be reframed. Organizational structure and its elements are continuously produced and reproduced over time (Giddens, 1984). Organizational structure is contingent on repeated, coordinated actions of its members. Structure shapes action, while these actions simultaneously reproduce structure through routines and procedures over time.

Indeed, organizational routines are continuously enacted and transformed in practice (Feldman & Pentland, 2003). Specific formal arrangements may exist, but without members to enact them through their continuous behavior, it is difficult for them to hold (Orlikowski, 2000). The stability of an organization's structure hinges on routines, task coordination, and repetition of its enactment by its members. In fact, organizational stability itself could be viewed as a temporary configuration that developed as a by-product of repeated ongoing enactment (Tsoukas & Chia, 2002). While the view of organizations and their structures as a continuous process of enactment is well-established, this view has been underutilized in analysis of disruption and its effect on structure.

During a large-scale external disruption, elements of organizational structure are often challenged. COVID caused a work-from-home mandate, thus disrupting the enactment of coordination, which was shifted through existing digital and communication structures. More subtly, however, disruption of such a scale challenged not only the structure, but the very enactment process that underpins organizational structure itself. Disruption of structural enactment affects three key segments: temporal, spatial, and procedural structuring.

Temporal structuring, according to Orlikowski & Yates, refers to the ongoing production of temporal patterns through organizational practice (2002). Patterns of time, such as schedules, deadlines, and daily rhythms, are continuously produced and reproduced through enactment. A disruption of a grand scale, such as COVID, directly challenges the current configuration of organization's temporal structuring. Project deadlines became uncertain, production experienced significant delays, and releases were pushed back. Without people present to enact the routine production of time, the temporal dimension of organizational structuring was significantly disrupted, especially in the early stages of pandemic, both in terms of internal time structure (i.e. project deadlines) and external (shipments, supply chain delays, etc.).

Just as temporal patterns are enacted rather than fixed (Orlikowski & Yates, 2002), spatial arrangements derive their organizing capacity from ongoing interaction and use (Lefebvre, 1991). Spatial structuring emphasizes that patterns of proximity, visibility, and interaction in organizations are not fixed properties of physical space but are continuously produced and reproduced through ongoing practice within that space (Lefebvre, 1991; Fayard & Weeks, 2007). In the context of COVID, organizational space was challenged most directly, through remote working arrangements and social distancing measures. The shift to remote work disrupted patterns of co-presence through which spatial coordination was enacted (Fayard & Weeks, 2007). By creating a literal physical boundary between an organization's space and its members, disruption weakened the ability of organization's members to reproduce organizing capacity associated with organization's space.

Finally, both organization's time and its space require the members to interact within the confines of both aspects and reproduce them in real time. Procedural structuring depends on the alignment of temporal and spatial patterns, and disruption across these dimensions undermines the enactment of routines. Procedural structuring highlights that routines, roles, and procedures are held up through ongoing action (Feldman & Pentland, 2003; Orlikowski, 2000). With both temporal and spatial configurations incapacitated, COVID further disabled the enactment of regular procedures and routines that are reinforced by both organization's time and space.

Taken together, these perspectives suggest that organizational structure derives its coordinating capacity from the sustained enactment of temporal, spatial, and procedural patterns, whose alignment is necessary for coordinated action. When disruption undermines this enactment, structure loses its ability to guide action, and established arrangements become insufficient for coordination. This shifts attention away from structure as an explanatory mechanism and toward the processes through which coordination is actively reconstructed, raising a central question: how do organizations maintain functioning when the ongoing enactment (and, by extension, the structure itself) is disrupted?

It is important to note, however, that not all disruptions are created equal. The structural conditions that govern organizational enactment of structure are most likely to break down when disruption is severe enough to affect all three key elements (temporal, spatial, and procedural). More routine, anticipated, or smaller-scale disruptions are likely to still be navigated effectively through existing structural elements. The enactment breakdown view developed, therefore, is most applicable to disruptions that destabilize existing rhythms, procedures, and coordination, particularly in situations of high task interdependence and high coordination demands.

5 Resilience as Organizing

The disruption of temporal, spatial, and procedural structuring does not eliminate the need for coordination, but shifts it from structurally embedded patterns to actively enacted organizing processes. Thus, organizational resilience, as we propose, is not a property embedded in organizational design, but an emergent outcome of organizing under conditions of disruption.

This view implies a fundamental shift in how resilience is understood. To describe an organization as "resilient" is not to attribute a stable characteristic embedded in its structure, but to recognize its capacity to organize effectively when established patterns of coordination are no longer viable. Resilience, therefore, reflects what organizations become capable of doing when structure ceases to function as a coordinating mechanism.

To specify this argument, we conceptualize organizational structure as an institutionalized expression of rules (both formal and informal) that guide action under particular conditions. These rules reflect historically accumulated assumptions about stable environments, predictable interactions, and recurrent patterns of coordination. In this sense, organizational design represents a contingent fit between rules and the conditions under which they have proven effective.

Large-scale disruptions fundamentally alter these conditions. What previously functioned as reliable guidance becomes insufficient for action. Under such conditions, coordination can no longer rely on the application of established rules. Instead, it depends on the capacity of organizational actors to actively reconstruct organizing in real time. Resilience, therefore, emerges through the ability to generate new patterns of meaning, coordination, and action when the rule-based foundations of structure weaken.

Building on this insight, we conceptualize organizing capacity for resilience as a configuration of three interrelated dimensions: orientation, orchestration, and openness (Fig. 1).

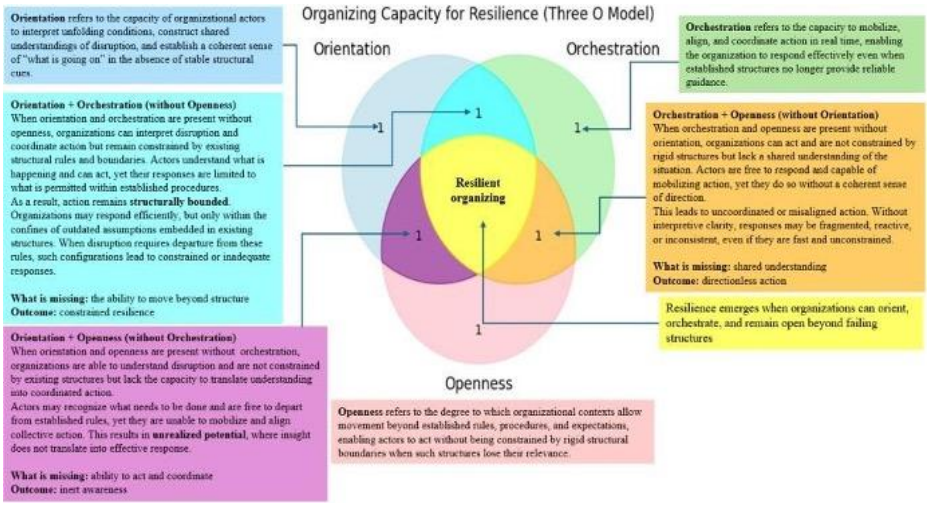


Fig. 1. The Three O Model. This figure conceptualizes resilience as emerging from the interdependent processes of orientation, orchestration, and openness. The overlap represents "resilient organizing," in which actors can interpret disruption, coordinate action, and adapt beyond established structures. Partial combinations illustrate constrained responses, where the absence of a key organizing process limits effective adaptation.

Orientation refers to the ability of individuals and organizations to make sense of disruption, interpret changing conditions, construct shared understandings, and establish direction for coordinated action. This dimension draws on sensemaking research, which emphasizes how meaning is constructed under conditions of ambiguity and uncertainty (Weick, 1995; Weick et al., 2005).

Orchestration captures the capacity to mobilize and coordinate action in real time, generating and aligning responses even in the absence of stable structural guidance. This dimension is grounded in coordination and routine-based perspectives, which highlight how organized action emerges through the alignment of interdependent tasks and actors (Malone & Crowston, 1994; Feldman & Pentland, 2003; Okhuysen & Bechky, 2009).

Openness reflects the extent to which organizational contexts allow movement beyond established rules, procedures, and expectations, particularly when existing structures no longer provide adequate guidance. This dimension builds on research in organizational learning and resilience, which emphasizes the importance of adaptation and exploration under conditions of uncertainty (March, 1991; Sutcliffe & Vogus, 2003; Lengnick-Hall et al., 2011).

The Three O Model acts as a substitute for the functions that structural elements normally play in organizing. In normal conditions, structure provides coordination,

guidance, role clarity, and procedural direction. However, once disruption disables the spatial, temporal, and procedural enactment of structure, organizing functions are carried by three Os. Under those conditions, orientation serves as a substitute for interpretive direction and allows organization's members to make sense of what is happening: orchestration aids in aligning actors and tasks in real time, while openness serves as a permission for members to adapt and consider novel ways of acting in a crisis beyond the existing rules. Therefore, the model does not replace structure permanently, nor does it negate its importance in a broader sense. Instead, it explains how organizations reconstruct the core organizing functions of structure when structure itself loses coordinating effectiveness.

Additionally, these dimensions do not operate independently but are mutually conditioning. Orientation enables orchestration by providing interpretive clarity, while orchestration translates understanding into coordinated action. At the same time, both orientation and orchestration are shaped by the degree of openness. In organizational contexts characterized by high rigidity, strong control orientation, and fear of deviation, the capacity to reinterpret situations and mobilize action is significantly constrained.

Conversely, higher levels of openness create conditions in which actors can move beyond established rules and engage in adaptive organizing. This view builds on organizational learning, emphasizing exploration and adaptation when existing routines become insufficient (March, 1991), and existing resilience research that highlights adaptability and flexibility under adversity (Lengnick-Hall et al., 2011; Vogus & Sutcliffe, 2007; Williams et al., 2017). In this sense, openness does not directly produce resilience but enables the enactment of orientation and orchestration by reducing the constraints imposed by existing structures, thus enabling actors to speak up, suggest novel approaches, and deviate from existing patterns without the fear of sanction (Edmondson, 1999).

Resilience, therefore, does not emerge from any single dimension, but from their configuration. Specifically, resilience arises when organizations can orient themselves within disruption, orchestrate coordinated action, and remain open enough to move beyond structural constraints. When these three dimensions intersect, organizations can reconstruct organizing even when existing structures lose their coordinating effectiveness.

6 Limitations and Suggestions for Further Research

This study has several limitations that point to important avenues for future research. First, the empirical context of the COVID-19 pandemic represents a large-scale, highly disruptive, and relatively unprecedented event. While this context provides a powerful setting for examining the breakdown of structural coordination, it also raises questions regarding the generalizability of our findings. Future research should examine whether similar dynamics occur in more localized, gradual, or domain-specific disruptions, where the conditions of structural enactment may be challenged to a lesser extent or in different ways.

Second, our theoretical framework conceptualizes resilience as an emergent organizing capability grounded in the interplay between orientation, orchestration, and openness. However, these dimensions have not yet been empirically operationalized or

tested. Further research is needed to develop reliable measures of these constructs and to examine how they interact over time in shaping resilience outcomes. In particular, longitudinal and process-oriented research designs would be well-suited to capturing the dynamic interplay between these dimensions as organizations respond to disruption.

Third, our argument rests on a process-based reconceptualization of organizational structure as enacted rather than fixed. While this perspective is well-supported in organizational theory, our study does not directly observe the micro-level processes through which enactment breaks down. Future research could build on this work by employing qualitative or mixed method approaches to examine how temporal, spatial, and procedural patterns are disrupted in practice, and how actors respond to these disruptions in real time. Another promising extension of enactment breakdown research could attempt to understand the temporal sequencing of the breakdown. Understanding how the breakdown unfolds in real time, as well as how different aspects of enactment (temporal, spatial, procedural) interact as they weaken, would potentially offer novel insights into effects of disruption on structure.

Finally, while we focus on the breakdown of structural coordination as a key mechanism underlying resilience, we do not fully specify the boundary conditions under which structure may retain or regain its coordinating capacity. It is possible that certain forms of structure, such as highly redundant, flexible, or digitally mediated arrangements, may be more resilient to disruption. Future research should explore how different structural configurations shape the extent to which organizing can be sustained or reconstructed under varying conditions of disruption.

7 Conclusion

This paper reconceptualizes organizational resilience as an enacted organizing capability that emerges under conditions in which the coordinating capacity of structure is disrupted. Drawing on empirical insights from the COVID-19 pandemic and building on process and practice-based perspectives, we show that organizational structures do not function as stable carriers of resilience. Instead, their effectiveness depends on the ongoing enactment of temporal, spatial, and procedural patterns. When disruption interrupts these patterns, structure remains present but loses its ability to reliably guide coordinated action.

Under such conditions, resilience cannot be fully explained through organizational design or accumulated experience. Instead, it emerges through the active reconstitution of organizing. We conceptualize this process through three interrelated dimensions: orientation, orchestration, and openness. Together, these dimensions enable organizations to reconstruct meaning, align action, and move beyond the constraints of established structures. In this view, resilience reflects the capacity of organizations to sustain coordinated action when existing arrangements no longer provide reliable guidance.

By demonstrating that the coordinating capacity of structure depends on its continuous enactment, this paper challenges structure-centric explanations of resilience and highlights the conditions under which such explanations may fail. More broadly, this perspective highlights the limits of structure-centric explanations in contexts of large-scale disruption and underscores the importance of examining how organizing unfolds

when the assumptions underpinning established structures are destabilized. In doing so, it opens new avenues for research on the dynamics of organizing under disruption and provides a foundation for understanding how organizations sustain functioning in increasingly uncertain and volatile environments.

Disclosure of AI use. AI-assisted tools were used during the drafting and revision process for language editing, structural feedback, and reference formatting. The authors thoroughly reviewed and revised all outputs, and take full responsibility for the content, citations, conclusions, and the final manuscript.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Aleksić Mirić, A., Bogićević Milikić, B., Janićijević, N.: Organisational resilience during the crisis caused by the COVID-19 pandemic in the Republic of Serbia: Experiences and recommendations for the future. In: Zarotiadis, G., Stojanović, Ž., Molnar, D. (eds.) *Post-COVID19 in SEE and Black Sea Region: Responses towards SDGs*, pp. 46–68. Centar za izdavačku delatnost Ekonomskog fakulteta u Beogradu, Belgrade (2022)
2. Annarelli, A., Nonino, F.: Strategic and operational management of organizational resilience: Current state of research and future directions. *Omega* 62, 1–18 (2016). https://doi.org/10.1016/j.omega.2015.08.004
3. Edmondson, A.: Psychological safety and learning behavior in work teams. *Administrative Science Quarterly* 44(2), 350–383 (1999). https://doi.org/10.2307/2666999
4. Fayard, A.-L., Weeks, J.: Photocopiers and water-coolers: The affordances of informal interaction. *Organization Studies* 28(5), 605–634 (2007). https://doi.org/10.1177/0170840606068310
5. Feldman, M.S., Pentland, B.T.: Reconceptualizing organizational routines as a source of flexibility and change. *Administrative Science Quarterly* 48(1), 94–118 (2003). https://doi.org/10.2307/3556620
6. Felin, T., Foss, N.J., Ployhart, R.E.: The microfoundations movement in strategy and organization theory. *Academy of Management Annals* 9(1), 575–632 (2015). https://doi.org/10.1080/19416520.2015.1007651
7. Giddens, A.: *The Constitution of Society: Outline of the Theory of Structuration*. University of California Press, Berkeley (1984)
8. Lengnick-Hall, C.A., Beck, T.E., Lengnick-Hall, M.L.: Developing a capacity for organizational resilience through strategic human resource management. *Human Resource Management Review* 21(3), 243–255 (2011). https://doi.org/10.1016/j.hrmr.2010.07.001
9. Lefebvre, H.: *The Production of Space*. Nicholson-Smith, D. (trans.). Blackwell, Oxford (1991)
10. Linnenluecke, M.K.: Resilience in business and management research: A review of influential publications and a research agenda. *International Journal of Management Reviews* 19(1), 4–30 (2017). https://doi.org/10.1111/ijmr.12076

11. Malone, T.W., Crowston, K.: The interdisciplinary study of coordination. *ACM Computing Surveys* 26(1), 87–119 (1994). https://doi.org/10.1145/174666.174668
12. March, J.G.: Exploration and exploitation in organizational learning. *Organization Science* 2(1), 71–87 (1991). https://doi.org/10.1287/orsc.2.1.71
13. Okhuysen, G.A., Bechky, B.A.: Coordination in organizations: An integrative perspective. *Academy of Management Annals* 3(1), 463–502 (2009). https://doi.org/10.1080/19416520903047533
14. Orlikowski, W.J.: Using technology and constituting structures: A practice lens for studying technology in organizations. *Organization Science* 11(4), 404–428 (2000). https://doi.org/10.1287/orsc.11.4.404.14600
15. Orlikowski, W.J., Yates, J.: It's about time: Temporal structuring in organizations. *Organization Science* 13(6), 684–700 (2002). https://doi.org/10.1287/orsc.13.6.684.501
16. Sutcliffe, K.M., Vogus, T.J.: Organizing for resilience. In: Cameron, K.S., Dutton, J.E., Quinn, R.E. (eds.) *Positive Organizational Scholarship: Foundations of a New Discipline*, pp. 94–110. Berrett-Koehler, San Francisco (2003)
17. Tsoukas, H., Chia, R.: On organizational becoming: Rethinking organizational change. *Organization Science* 13(5), 567–582 (2002). https://doi.org/10.1287/orsc.13.5.567.7810
18. Vogus, T.J., Sutcliffe, K.M.: Organizational resilience: Towards a theory and research agenda. In: *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, pp. 3418–3422. IEEE, Piscataway (2007). https://doi.org/10.1109/ICSMC.2007.4414160
19. Weick, K.E.: *The Social Psychology of Organizing*. 2nd edn. McGraw-Hill, New York (1979)
20. Weick, K.E.: *Sensemaking in Organizations*. Sage Publications, Thousand Oaks (1995)
21. Weick, K.E., Sutcliffe, K.M.: *Managing the Unexpected: Assuring High Performance in an Age of Complexity*. Jossey-Bass, San Francisco (2001)
22. Weick, K.E., Sutcliffe, K.M., Obstfeld, D.: Organizing and the process of sensemaking. *Organization Science* 16(4), 409–421 (2005). https://doi.org/10.1287/orsc.1050.0133
23. Williams, T.A., Gruber, D.A., Sutcliffe, K.M., Shepherd, D.A., Zhao, E.Y.: Organizational response to adversity: Fusing crisis management and resilience research streams. *Academy of Management Annals* 11(2), 733–769 (2017). https://doi.org/10.5465/annals.2015.0134

Original Research Article *https://doi.org/10.5281/zenodo.23475218*

Specific Characteristics of the Consulting Process in SMEs: Conceptual Framework

Jovan Krivokapić [1] [0000-0001-9010-7802] and Milan Jovanović [2] [0009-0004-7740-6217]

[1,2] University of Belgrade, Faculty of Organizational Sciences, Jove Ilića 154, 11000 Belgrade, Serbia
jovan.krivokapic@fon.bg.ac.rs

Abstract. Small and medium-sized enterprises (SMEs) are the dominant form of business in most economies, but they often lack sufficient internal capacities to overcome business challenges and solve complex problems, which is why they decide to seek assistance from consultants. This paper analyzes the key features of consulting projects in SMEs, considering that they are characterized by limited resources, a less rigid organizational structure, and centralized owner decision-making, as well as more flexible cooperation based on a higher level of mutual trust. Although the process of providing consulting services includes the same basic phases and steps as in large organizations, in SMEs it is generally shorter, less formal, and focused on achieving rapid effects and practically applicable solutions.

Keywords: Management consulting, Small and medium-sized enterprises, Consulting approach.

1 The concept and role of consulting

Management consulting emerged practically alongside the development of society, but it has changed its form in accordance with shifts in the business environment. The focus of consulting services has also evolved, thereby transforming its very essence. For this reason, consulting can be viewed quite broadly, and in that sense, consultants may be considered all those who attempt to change or improve business situations, whether they are employed within the organization or come from outside [1]. On the other hand, nowadays consulting is generally viewed as a narrower and more specific concept, referring to the provision of specialized services by independent external professionals that enable the management of the client organization to take action in complex managerial situations [2]. However, since the problems faced by client organizations may differ in form, complexity, and scope, the role of consultants is adapted to specific tasks and their characteristics. Therefore, consultants are expected to help with [3]:

- achieving organizational purposes and objectives,
- solving management and business problems,
- identifying and seizing new opportunities,
- enhancing learning,
- implementing changes.

Practice shows that clients most often decide to hire consultants when problems arise, which is logical, since hiring consultants requires certain financial investments, and many organizations first attempt to resolve challenges independently [4]. These problems can be viewed from several perspectives and may be [5]:

- the result of various external influences or poor use of internal potential,
- strategic or operational in nature,
- caused by issues in structure, processes or employees,
- consequences of poor past decisions, inadequate current management, or warnings for the future if changes are not properly monitored.

With this in mind, it can be concluded that the key motive for using consulting services lies in obtaining knowledge and expertise that the client cannot provide through internal capacities, and which should serve as an effective tool for managing change and developing the organization in the desired direction. At the same time, consultants are expected to apply an analytical approach to the problems faced by the client in order to create the conditions for successful cooperation and effective project outcomes [6].

2 The consulting process

Cooperation on consulting projects is based on a series of steps and activities that should enable the adequate achievement of the goals that the client seeks, that is, to reduce or eliminate the problems the client is facing. On the other hand, the specifics of each project will affect where the focus will be placed, but practical experience and a review of relevant literature suggest that it is possible to notice some common steps or phases that appear to a certain extent in each of them.

Table 1. The steps of the practical development of a management consulting project [7]

	Seven steps of Kolb and Frohman (1970)	*The three forms of formulation of the change of Jones and Lawrie (1994)*	*The management model of McLarty and Robinson (1998)*	*Oliveira (2006) consultant intervention stages*
Stages of the practical development of a consulting project	• Scouting • Entry • Diagnosis • Planning • Action • Evaluation • Termination	• Agenda • Exploratory action • Reflection	• Defining questions • Analysis • Diagnosis • Problems solution • Recommendations and implementation	• Identification • Entry • Position audit • Planning • Action • Monitoring and evaluation • Conclusion

The previous table shows how different authors perceive the consulting process in different ways, so it may have, at least formally, different steps, but a deeper analysis of them shows that the activities carried out within them are relatively similar. Having this in mind, it can be said that the consulting process may be viewed through several general phases, and these are [3], [8]:

1. Initiating cooperation – This is the phase in which the first contacts with the clients are established with the aim of determining whether and in what way consultants can help them. Accordingly, this is a preparatory and planning phase in which a preliminary diagnosis of the problem is established, and the following steps will, to a significant extent, depend on the quality of the activities carried out here. It often happens that the client contacts a larger number of potential partners in order to determine which one of them might help in the best way.
2. Establishing the diagnosis – During the establishment of the diagnosis, the consultant and the client cooperate intensively in order to define the precise scope of the problem that should be solved through the consulting project, but also to identify the factors that lead to it and/or influence it. In this step it is necessary to provide the consultant with data that are important for identifying specific characteristics, problems and challenges, and it is interesting that sometimes, due to the presence of consultants, changes in employee behavior can already be noticed at this stage.
3. Planning activities – Finding a solution to the problem implies the development of alternatives whose implementation should help clients achieve their goals. Here, the consultant is expected, through their expertise, to recognize adequate solutions, but also to select those that are truly applicable in practice, taking into account the business environment in which the client operates, as well as the real effects that can be expected in the given circumstances. This phase does not imply only the selection of the appropriate alternative, but also the creation of an implementation strategy, because solutions often involve unpopular and drastic measures, so it is also necessary to make a plan for overcoming resistance to change among employees.

4. Implementation – In the implementation phase, the proposed solutions are applied in practice, so sometimes they are carried out with the direct help and intensive participation of consultants, while sometimes the client performs them independently because it is usually a significantly cheaper option. In this step, the quality of the realization of previous activities is actually seen best, because it becomes clear whether the problem was perceived from the right angles and whether the proposed solutions are truly adequate. It often happens that in this phase action plans are corrected to a certain extent, because it is not possible to fully foresee all the factors that influence the success of implementation.
5. Completion of cooperation – The last phase of the consulting process is the phase in which one consulting project formally comes to an end, but in a broader sense it does not necessarily imply the end of future cooperation. If the client is satisfied with the approach, the relationship, and most importantly the proposed solutions and accompanying effects, then it is realistic to expect that he might turn again to the consultant when new challenges arise. On the other hand, sometimes, especially in the case of more complex consulting projects, more time needs to pass until the effects appear, so a final judgment on the success of the cooperation can only be made afterwards.

It is important to emphasize that the described consulting process is not linear, but rather an iterative set of steps, because solutions are reached through several cycles [9], which depends not only on the complexity of the problem, but also on the maturity of the client organization and its readiness to accept changes [10].

3 Key Differences in the Consulting Approach to SMEs

When observing the structure of the economy in any business environment, special importance must be given to small and medium-sized enterprises (SMEs), because they represent the dominant form of business in most economies, and at the global level they make up about 90% of the total number of enterprises, employ more than 70% of all workers, and contribute 50% of GDP worldwide [11]. Serbia is no exception, where SMEs dominate with 99.5% of the total number of companies and employ 62.5% of all workers [12].

Although it is obvious that SMEs represent the foundation of the economy [13], practice shows that they often face a lack of internal capacities necessary to overcome business challenges and solve complex problems. This fact increases their need for external expertise, so they often choose to hire consultants who are expected to help them, whether in determining strategic courses of action or in using potential for improvement of operational efficiency. In this way, SMEs gain access to knowledge and tools that they lack, which enables them to keep pace with large organizations.

On the other hand, differences in size, internal organization and the availability and accessibility of resources between SMEs and large systems lead to the need for certain differences in the consulting approach. This practically means that the basic principles on which the cooperation between consultants and clients is based will remain relatively similar, but that their application will be adapted to the specifics of the organizational

context and the particularities of the organization itself. In this way, preconditions are created for the cooperation to be successful, that is, primarily to meet the goals for which SMEs decide to seek assistance from consultants. The key differences in the approach that consultants have in SMEs are conditioned by the limited resources in SMEs, the lack of specialized knowledge and formal management structures, the greater flexibility of SMEs compared to large companies, as well as the specific relationship that consultants establish with key decision-makers [14].

3.1 Limited resources (financial, human, technological)

Compared to large enterprises, SMEs generally have limited financial resources, which can be a significant obstacle when considering the option of hiring consultants [15]. It is also interesting that even in those situations where a decision is made to establish cooperation with external experts, the scope of consulting projects is often limited, or cooperation is agreed upon for a shorter period, precisely for reasons of a financial nature. Consulting engagements in SMEs are mostly concrete and focused on clear or visible problems, while in large companies a broader and long-term perspective is more often considered. Consequently, there is greater pressure in SMEs to achieve results and effects as early as possible, meaning that a rapid return on invested funds is expected.

In addition to financial resources, SMEs sometimes also face problems with human resource capacities, because employees are engaged in a large number of functions, which takes a significant amount of their time and reduces their availability to work with consultants. Furthermore, the technological base that SMEs use in carrying out everyday activities may also be a limitation for the implementation of more complex solutions, so the scope for consultants' creativity in finding adequate alternatives is narrower than in larger companies [16], [17].

3.2 Lack of specialized knowledge and formal management structures

Practice indicates that in SMEs the organization is often less formalized, with more improvisation in carrying out daily activities, and with unclear boundaries of responsibility between employees and managers [18]. This practically means that the focus of employees in higher positions is directed toward several different areas, which makes their development in specific areas more difficult. Therefore, when they are expected to react quickly, they tend to respond intuitively, based on previous experience and without a deeper analysis of relevant data. This fact can represent an additional challenge for consultants because sometimes they are expected to propose solutions that do not differ greatly from the established way of working and thinking, but sometimes a new, more formal system for overcoming operational and management challenges is needed, and that may cause resistance in implementation of the proposed solutions. With that in mind, it can be concluded that the organizational aspect of cooperation on SME consulting projects is extremely sensitive.

3.3 Greater flexibility and focus on practical, quickly applicable solutions

Although they often have significant resource limitations compared to large companies, SMEs are generally more flexible and more open to change, which is a consequence of the aforementioned lower level of formalization and the easier adaptability of their structure and processes. On the other hand, dependence on limited resources makes them extremely sensitive to any major disruptions, so their flexibility must be carefully directed and managed through the analysis of external tendencies and trends. When designing solutions, consultants must create an adequate balance between using the advantages of flexible organization on the one hand and risk management on the other, especially because changes in SMEs have a faster effect, which can be positive, but may also threaten their stability in case the solutions are not appropriate and effective [19].

Since these limitations make it difficult for SMEs to decide on contracting complex or long-term projects that do not produce a clear result, the focus of these agreements is more often on operational improvements than on undertakings that imply major strategic shifts. Consequently, consultants are expected to provide recommendations that are clear, concrete and immediately applicable. However, it is also important to emphasize the fact that in consulting projects in SMEs it is easier to perceive the mutual relationships and direct and indirect influences between different parts of the organization, which represents both an opportunity and a challenge for consultants at the same time [18]. This cause-and-effect chain allows their solutions to be more comprehensive and useful for a larger part of the organization, but it also requires a different focus, as well as specific expertise and knowledge of different processes. On the other hand, the application of their solutions can generate significant and more visible effects than in large organizations, which is a particular motivation for them and leads to satisfaction when it is shown that their solutions "really work". This is more difficult to accomplish in large organizations because they generally need more time to achieve possible improvements, while such improvements are largely conditioned by other factors that are not directly related to the reasons for which consultants were engaged.

3.4 The strong role of the owner in decision-making and the greater importance of trust and a long-term relationship with the consultant

A special challenge for consultants is the significant and often exclusive role played by the SME owners in making all important decisions in the organization. The owner often bases these decisions on his own views, previous experience and personal judgment, so it is sometimes difficult for him to accept suggestions from external associates, especially if they are not in line with his vision and expectations. That is why the consultant has a special task to manage their relationship and cooperation carefully, as he aims to demonstrate expertise and thereby gain the owner’s trust in the proposed solutions. On the other hand, if that relationship is built in an adequate way and if the owner is satisfied with working with consultants, then the barriers to the implementation of the solutions are lower than in large companies where it is necessary for more decision-makers to accept the consultant's proposals and be ready to support the changes that those proposals bring [20].

Due to mistrust and skepticism regarding the expected effects, SME owners often have difficulty deciding whether to hire consultants, and when they do, they usually hire them based on personal recommendations or reputation. On the other hand, although that beginning can be challenging, in situations where the success of consulting projects is achieved, a deeper and stronger connection is established between the owner and the consultants, and the cooperation becomes long-term rather than one-time. The owner and the consultant develop a relationship that is closer and less formal, which certainly contributes to overcoming the initial tensions and creates preconditions for mutual satisfaction with the final outcome [21].

This relationship between consultants and owners must be based on trust and transparency in every phase of cooperation, which implies that the goals, expectations and responsibilities of all parties involved are clearly defined at the very beginning. The owner must be ready for active engagement and exchange of relevant information during the entire process, but also to accept responsibility for making final decisions, because the key role of the consultant is to be an advisor and expert who defines the directions of future work and development, but he is the one who must manage the organization even when the consulting projects are completed.

Considering all of the above, it can be concluded that the differences between SMEs and large corporations have a significant impact on the organization and implementation of consulting projects, as well as on the very nature of cooperation between consultants and clients. Accordingly, the following table summarizes the presented conclusions and provides a comparative overview of the differences across the dimensions of consulting projects discussed in this paper.

Table 2. Comparative overview of consulting process characteristics in SMEs and large corporations

Dimension	SMEs	Large corporations
Organizational structure	Simple, flat structure, low formalization, overlapping roles	Complex, multi-layered hierarchy, high formalization and specialization
Resource availability	Limited financial, human and technological resources	Abundant financial, human and technological resources
Decision-making process	Centralized owner/manager-driven decision-making	Decentralized and participative decision-making, multiple approval levels
Consulting time horizon	Short-term perspective, focus on quick and visible results	Long-term perspective, strategic orientation, projects lasting months or years
Nature of consulting projects	Narrower scope, focused and concrete projects	Broad scope, complex and integrated projects, multiple workstreams
Consultant–client relationship	Informal, close and personal relationship, high level of trust	Formal and role-based relationship, professional distance
Process formalization	Less formal and flexible processes, adapted in real time	Highly structured processes, clear procedures and documentation
Change implementation	Faster implementation with fewer stakeholders	Gradual implementation involving many stakeholders
Failure risk and impact	Higher sensitivity to failure due to limited resources	Lower operational risk due to resource buffers and organizational stability
Employee roles and expertise	Employees perform multiple functions simultaneously	Strong specialization and division of responsibilities
Implementation of solutions	Solutions can be implemented more quickly and directly	Solutions often require extensive coordination and approval
Visibility of consulting effects	Effects are more direct, visible and measurable	Effects of consulting projects are more difficult to isolate and measure

The findings presented in the table indicate that differences in the size and characteristics of client organizations require an adaptation of the very logic of consulting engagement, since the success of consulting projects largely depends on the alignment of the approach with the specific characteristics of the client and the business context in which it operates.

4 Conclusion

Hiring consultants in SMEs involves the same phases and steps as in large companies, but the cooperation between consultants and clients differs in several key features. In large companies the procedures are clearly defined and the hierarchy is rigid, while SMEs are usually characterized by a lower level of formalization and more flexible processes, so consultants are expected to have a less bureaucratic and more customized approach to the analysis and development of solutions. When it comes to the availability of resources, in large organizations the budget for the potential engagement of consultants is higher, while SMEs usually have limited financial resources, as well as human capacities that need to

carry the burden of changes that are the outcome of consulting projects. Finally, in large companies decisions are made at different levels of the organization following more detailed prior analyses, whereas in SMEs decision-making is centralized and often intuitive, which places consultants in an ambivalent position – a smaller number of decision-makers facilitates operational cooperation and the potential application of solutions, but on the other hand, it requires a complex and more personal relationship with them, in order to achieve the necessary level of trust, which is a prerequisite for the possible success of the consulting project.

Having this in mind, the consulting process in SMEs is often shorter and less formal than in large systems, while the focus is on quick solutions that produce measurable results and concrete effects in a short period of time. Some steps in this process are combined, and consultants are expected to have an extremely active role not only in problem analysis and development of solutions, but also in the implementation of selected alternatives and action plans. Finally, it can be concluded that the key factors of successful cooperation are in fact the speed of decision-making, the client's willingness to take action and the operational applicability of the selected solutions, since a consultant who works with SMEs is often better able to understand the entire organization and the relations between its key elements, which places him in a stronger position to design solutions that may produce a broader impact than would typically be the case in large business systems.

The main limitation of this paper lies in the fact that it is primarily based on a review of relevant literature and the professional experience of the authors, which is why future research should be directed toward the empirical examination of the applicability of the presented conceptual framework in different organizational and industry contexts, thereby making it possible to further validate and extend the presented conceptual findings.

References

1. Kipping, M., Clark, T.: Researching Management Consulting: An Introduction to the Handbook. In: The Oxford Handbook of Management Consulting, New York, Oxford University Press, pp. 1-28. (2012)
2. Kriechbaum, F.: Successful Business Relationships in Management Consulting: Driven by Soft Skills?. Examicus Publishing (2012)
3. Kubr, M.: Management consulting: A guide to the profession. International Labour Office, Geneva (2002)
4. Salvador, E., El Asraoui, H., Akbaraly, M.: The difficult relationship between the consultancy market and SMEs: inspiring insights from the case of Drôme. Revue de l'Entrepreneuriat/Review of Entrepreneurship 18(3), 133-163 (2019)
5. Krivokapić, J.: Uticaj konsaltinga na efikasnost organizacionog restrukturiranja. University of Belgrade, Belgrade (2016)
6. Phillips, J.J., Trotter, W.D., Phillips, P.P.: Maximizing the Value of Consulting: A Guide for Internal and External Consultants, John Wiley & Sons, Inc., Hoboken (2015)
7. Lopes da Costa, R., Pereira, L.F, Dias, Á., Gonçalves, R., Vinhas da Silva, R., Teixeira, N.: Life cycle development and approach of management. International Journal of Applied Systemic Studies 11(2), 97-120 (2024)
8. O'Mahoney, J., Markham, C.: Management Consultancy, Oxford University Press, Oxford (2013)

9. Andrii, K.: Assessing the Impact of An Integrated Consulting Methodology (Diagnostics, Planning, Implementation, Monitoring) On Improving the Productivity and Scalability of Smes in the United States. International Journal of Management and Business Development 3(3), 1-11 (2026)
10. Cawsey, T.F., Deszca, G., Ingols, C.: Organizational change: an action-oriented toolkit. SAGE, Los Angeles (2016)
11. International Council for Small Business: ICSB Annual Global Micro-,Small and Medium-Sized Enterprises Report. Washington, DC (2024)
12. The Serbian Business Registers Agency: Annual Report on Business Operations. Belgrade (2024)
13. Sagala, G.H., Őri, D.: Toward SMEs digital transformation success: a systematic literature review. Information Systems and E-Business Management 22, 667–719 (2024)
14. Galli-Debicella, A.: How SMEs Compete Against Global Giants Through Sustainable Competitive Advantages. Journal of Small Business Strategy 31(5), 13–21 (2021)
15. Abraham, F., Schmukler, S.L.: Addressing the SME Finance Problem. World Bank Policy Research Working Paper Series (2017)
16. Rokhim, R., Wulandari, P., Mayasari, I.: Small medium enterprises technology acceptance model: A conceptual review. International Journal of Business and Society 19, 689-699 (2018)
17. OECD: OECD SME and Entrepreneurship Outlook 2023. OECD Publishing, Paris (2023)
18. Harney, B., Gilman, M., Mayson, S., Raby, S.: Advancing understanding of HRM in small and medium-sized enterprises (SMEs): critical questions and future prospects. The International Journal of Human Resource Management 33(16), 3175–3196 (2022)
19. Massiera, P.: How do adaptive market experimentation, open marketing and strategic flexibility contribute to small firms' business model innovation in the context of a crisis?. Journal of Small Business and Enterprise Development 31(7), 1302-1326 (2024)
20. Mole, K., North, D., Baldock, R.: Which SMEs seek external support? Business characteristics, management behaviour and external influences in a contingency approach. Environment and Planning C: Politics and Space 35(3), 476-499 (2017)
21. Grabowski, W., Stawasz, E.: Business consulting, knowledge absorptive capacity, and innovativeness: A triangular model for micro and small enterprises in Poland. Journal of Entrepreneurship, Management and Innovation 19(1), 7–40 (2023)

Original Research Article *https://doi.org/10.5281/zenodo.23475247*

Research-Intensive Teams and Organizational Resilience: The Role of Cognitive Routines

Isidora Simeunović [1,2] [0009-0003-5720-8128]

[1] University of Belgrade, Faculty of Organizational Sciences,
Jove Ilića 154, 11000 Belgrade, Serbia
[2] Computer systems, Institute Mihajlo Pupin,
Volgina 15, 11060 Belgrade, Serbia
isidora.simeunovic@pupin.rs

Abstract. Organizations increasingly operate in conditions of permanent disruption shaped by emerging technologies, geopolitical shifts, environmental pressures, and continuous novelty. Previous studies link resilience to flexibility, learning, coordination, and adaptive capability, but give little attention to the recurring cognitive practices through which organizations interpret uncertainty and respond effectively. This issue is especially relevant in research-intensive teams, where goals shift, information is incomplete, and formal rules provide limited guidance.

This paper develops a conceptual framework positioning cognitive routines as organizationally embedded micro-mechanisms linking design conditions to resilience outcomes. Cognitive routines refer to recurring shared practices of interpretation, evidence review, reflection and learning. The framework highlights how design choices, such as onboarding, mentoring, review forums, communication architecture, and digital support tools may strengthen such routines. Resilience is viewed as an ongoing process of anticipation, coping, and adaptation.

The paper contributes by refining resilience theory, extending organizational design beyond formal structure, explaining why research-intensive teams provide a revealing context, and arguing that digital and AI-supported tools create value only when they strengthen rather than weaken human cognitive routines.

Instead of claiming an entirely new phenomenon, the paper synthesizes fragmented literatures into a design-oriented construct of recurring collective reasoning. It concludes that organizations facing persistent disruption survive not only through structure, but through how they repeatedly think, judge, and learn together.

Keywords: organizational resilience, organizational design, cognitive routines, team cognition, research-intensive teams, sociotechnical systems, intelligent organizations.

1 Introduction

Organizations now face a landscape of permanent disruption characterized by rapid technological change, geopolitical uncertainty, environmental crisis, regulatory

volatility, and changing stakeholder expectations. In such conditions, resilience is no longer relevant only during episodic crises; it has become a normal organizational requirement. Resilience alone is not enough if responses are slow, rigid, or poorly judged. Organizations today must be both resilient and intelligent: able to absorb shocks, adapt, and respond effectively to changing conditions [1]-[4].

Such conditions create a direct challenge for organizational design. Classical approaches often emphasized formal structure, hierarchy, reporting lines, specialization, and control mechanisms. But effective organizational design also considers roles, processes, incentives, technologies, communication styles, knowledge systems, and cultural conditions that shape how work is coordinated and understood. Longstanding insights in organizational theory have warned that structure alone does not constitute organization. In rapidly changing settings, designing organizations cannot be reduced to rearranging boxes on an organizational chart; it must also create conditions for judgment, learning, responsiveness, and shared problem solving [6]-[8], [17].

Existing literature associates organizational resilience with flexibility, resourcefulness, coordination, dynamic adaptation, and learning. These perspectives are valuable, but they often remain at a relatively macro level of explanation. They identify important organizational capacities without clearly specifying the recurring practices through which teams actually interpret ambiguous situations, question assumptions, evaluate evidence, and revise action under pressure. Resilience is therefore often described as an outcome or capability, but less explicitly as something realized in practice through disciplined collective cognition [1]-[5].

Because critical sensing, interpretation, experimentation, and adaptation often occur within teams, much of organizational resilience is enacted through team-level processes.

The limitation is particularly visible in research-intensive teams and other knowledge-intensive environments. Such teams frequently operate under conditions of persistent novelty, fluid problem definitions, incomplete information, and strong interdependence across specialized expertise. Success depends less on predefined procedures and more on the quality of interpretation, experimentation, reflection, and coordinated judgment. Research-intensive teams are used here as a theoretically revealing context because novelty, ambiguity, and distributed expertise make cognitive routines easier to observe and more consequential. Although the perspective may also apply to startups, hospitals, military units, consulting firms, and crisis-management teams, its mechanisms are especially visible in research-intensive settings [14]-[16], [18].

In response to this gap, the paper introduces the concept of cognitive routines as organizationally embedded micro-mechanisms of resilience. Cognitive routines are defined here as recurring, socially supported practices through which teams frame problems, detect weak signals, challenge assumptions, review evidence, reflect on decisions, and update interpretations over time. Unlike operational routines centered on repetitive task execution, cognitive routines concern recurring practices of reasoning, interpretation, and learning that become stabilized within organizational life through meetings, review forums, mentoring arrangements, onboarding processes, decision protocols, and shared artifacts [5], [9], [10], [19].

2 Theoretical Background

Organizational design has traditionally been associated with formal structure, including hierarchy, division of labor, reporting relationships, and coordination mechanisms. While these dimensions remain important, structure alone cannot explain organizational effectiveness, especially in turbulent environments. Organizations are also shaped by processes, incentives, technologies, communication flows, routines, and cultural norms that influence how work is interpreted and coordinated in practice [6]-[8], [17]. In turbulent conditions, effective design concerns not only formal arrangements, but also the creation of conditions that support learning, responsiveness, and distributed judgment.

Resilience is increasingly viewed not as a fixed trait, but as a dynamic process. Under permanent disruption, it includes three interrelated dimensions: anticipation, coping, and adaptation. These refer respectively to noticing emerging risks or opportunities, maintaining coordinated performance during disruption, and revising practices and capabilities during or after disruption [1]-[4]. Anticipation supports coping, while adaptation strengthens future anticipation through learning. Resilience can thus be viewed not only as an outcome, but also as a capability produced through recurring organizational practices [1]-[3], [20].

Sensemaking helps explain performance under ambiguity. When situations are unclear and information incomplete, actors must construct plausible interpretations that enable coordinated action. Organizational effectiveness under disruption depends not only on available information, but on how that information is interpreted, discussed, and translated into shared understanding [4], [14]. Related research on team cognition shows that performance in uncertain settings depends on shared mental models, distributed expertise, communication quality, and the ability to integrate fragmented knowledge across team members [14].

These perspectives connect closely with research on organizational routines. Routines have often been associated with stability, predictability, and efficient execution, but contemporary work also shows that routines may support flexibility, reflection, and learning [2], [5], [19]. Standard operational routines primarily organize task execution, whereas cognitive routines organize how teams repeatedly interpret, review, and adjust under uncertainty. They provide disciplined forums through which expertise can be surfaced, challenged, and recombined.

Research-intensive teams provide a particularly useful environment for examining these issues. Such teams often operate under novelty, evolving objectives, unclear pathways, and strong interdependence across specialized expertise. Examples include R&D units, interdisciplinary scientific collaborations, innovation labs, advanced engineering teams, and research institutes [14]-[16]. In such settings, predefined procedures remain relevant but are rarely sufficient. Teams must continuously frame problems, test hypotheses, integrate diverse knowledge, revise plans, and learn from failed experiments.

If cognitive routines are to become durable organizational capabilities, they must be embedded through deliberate design choices rather than left to chance. One important example is onboarding and organizational socialization. New members learn not only formal responsibilities, but also how the organization frames problems, escalates

concerns, evaluates evidence, handles disagreement, and reflects after decisions [9]-[13]. Mentoring, review forums, and cross-functional exposure can further strengthen these patterns by transferring experience-based judgment and reducing uncertainty during membership change.

Digital tools and AI-supported systems represent an additional contemporary design element. Shared dashboards, knowledge repositories, simulation tools, collaborative platforms, and AI-assisted analytics may strengthen collective sensemaking by improving visibility, memory, and analytical speed. They may also create risks through opacity, overreliance, information overload, or distorted confidence. Their value therefore depends on how they are integrated into human cognitive routines rather than on technology alone [21]-[24].

3 Research Gap and Paper Contribution

The literature reviewed above suggests that organizational resilience, design, sensemaking, team cognition, routines, onboarding, and digital support mechanisms all provide useful insights into how organizations survive and perform under changing conditions [1]-[8], [19]-[24].

However, these insights remain spread across only partially connected streams. Resilience research often emphasizes macro-level capacities or outcomes, including flexibility, preparedness, recovery, and renewal, while giving less explicit attention to the recurring practices through which teams interpret signals, challenge assumptions, integrate dispersed expertise, and revise action under pressure. Sensemaking and team cognition research address interpretive processes more directly, yet they are less frequently translated into actionable principles of organizational design. Research on routines contributes understanding of recurring patterns of coordination and learning, but less often distinguishes routines of task execution from routines of collective reasoning.

The central gap is not the absence of relevant literature, but the absence of a sufficiently explicit and design-oriented integration of that literature. Cognitive routines are not proposed as a replacement for sensemaking, team cognition, or organizational routines, but as an integrative construct that specifies how recurring practices of collective reasoning become embedded in work systems.

Although resilience is widely recognized as important, managers often struggle to operationalize it in practice [25]. This suggests a need for concrete organizational mechanisms through which resilience can be enacted and sustained.

To address this gap, the present paper develops a conceptual framework that positions cognitive routines between organizational design conditions and resilience outcomes. The model is propositional rather than deterministic. It does not claim universal linear causality, but offers a structured explanation of how design choices may enable recurring cognitive practices that, in turn, strengthen anticipation, coping, and adaptation.

The paper contributes in four ways. First, it refines resilience theory by shifting attention from broad organizational attributes toward explicit micro-level mechanisms through which resilient action is enacted in everyday work. Second, it extends to

organizational design research by arguing that effective design should include not only structures, incentives, and workflows, but also recurring systems for collective reasoning, reflection, and evidence-based adjustment. Third, it contributes to research on knowledge-intensive settings, especially research-intensive teams by clarifying why these contexts are especially revealing environments in which cognitive routines become visible and consequential. Fourth, it contributes to contemporary debates on intelligent organizations by arguing that digital and AI-supported tools do not automatically create intelligence. Their value depends on whether they strengthen or distort human cognitive routines related to interpretation, coordination, and learning [21]-[23].

4 Conceptual Framework

This section presents the proposed three-layer model explaining how resilience can be intentionally cultivated and strengthened in such teams.

Organizational Design Conditions ==(enables or constrains)==> Cognitive Routines ==(shapes)==> Resilience Outcomes

The model centers on the logic that design choices shape the conditions under which cognitive routines emerge and stabilize, while these routines influence how teams anticipate disruption, cope during disruption and adapt afterward.

4.1 Layer 1: Organizational Design Conditions

The first layer concerns the broader organizational conditions within which teams operate. These include formal and semi-formal elements of design that shape coordination, information flow, and the management of uncertainty. Relevant conditions include: (1) role clarity, (2) onboarding and mentoring systems, (3) review practices, (4) knowledge artifacts, (5) escalation paths, (6) cross-team interfaces, (7) communication systems or digital and AI-supported tooling.

These elements matter because they influence whether teams notice weak signals, exchange interpretations, challenge assumptions, escalate emerging risks, and preserve learning across time. Design conditions do not constitute cognitive routines themselves. Instead, they create structural, social and informational context within which such routines are more or less likely to emerge and stabilize. For example, role clarity reduces confusion during disruption, but excessive rigidity may suppress initiative. Mentoring arrangements may accelerate judgment transfer by helping less experienced members learn how uncertainty is handled in practice. Escalation paths may improve anticipation when they allow ambiguous concerns to surface early rather than remain locally trapped.

Design conditions do not automatically create resilience. They may enable useful cognitive routines or suppress them.

4.2 Layer 2: Cognitive Routines

The second layer consists of cognitive routines. These routines organize how teams repeatedly reason, interpret, review, and respond when conditions are unclear or changing.

Unlike standard operational routines that organize task execution and broader notions of sensemaking, which describe the interpretive process through which actors construct meaning under ambiguity, cognitive routines refer to recurring and organizationally reinforced practices that support and structure collective interpretation over time. In this sense, cognitive routines do not replace sensemaking; they provide the recurring organizational arrangements through which collective sensemaking becomes more consistent, shared, and reproducible across situations, teams, and time.

Core routines include (1) problem framing, (2) assumption surfacing, (3) weak-signal interpretation, (4) hypothesis formation, (5) evidence review, (6) structured reflection, and (7) decision logging. These routines are especially important in research-intensive work, where stable problem definitions and complete causal knowledge are rarely available. Such routines provide practical pathways through which collective sensemaking and adaptive action become more reliable.

Importantly, these routines are not treated here as purely individual habits or personality traits. Their organizational value lies in the fact that they can be reinforced through design choices, leadership expectations, shared forums, supporting artifacts, and everyday work practices.

Although cognitive routines are treated here primarily as collective and organizationally embedded practices, their emergence also depends partly on individual dispositions, communication habits, domain expertise, and interpersonal trust within teams. The framework therefore does not exclude micro-foundational influences; rather, it focuses on how such influences become stabilized, reinforced, and reproduced through recurring collective practices.

4.3 Layer 3: Resilience Outcomes

The third layer concerns resilience outcomes. Rather than treating resilience as a vague positive quality, the framework views it as a process expressed through three connected dimensions: (1) anticipation, (2) coping, (3) adaptation.

Anticipation concerns noticing emerging risks or opportunities. Coping refers to maintaining coordinated performance during disruption, while adaptation refers to revising practices, assumptions, or capabilities after new learning experiences.

These dimensions are analytically distinct but practically interconnected. Strong anticipation often improves coping by reducing surprise. Effective coping creates the stability required for adaptation. Adaptation, in turn, improves future anticipation by updating organizational knowledge.

In research-intensive teams, these outcomes may appear through reduced rework, faster escalation of critical issues, stronger knowledge continuity, improved cross-team learning, and better decision quality under pressure.

4.4 Linking the Three Layers

The central logic of the framework is that organizational design conditions influence resilience indirectly by enabling or constraining cognitive routines, which then shape anticipation, coping, and adaptation. For example, onboarding, mentoring, decision memory mechanisms, and AI-supported tools may improve resilience when they enhance collective sensemaking, but may undermine it when they create opacity, overconfidence, or cognitive overload [21]-[24].

From this perspective, resilient organizational design in research-intensive contexts should address not only speed, flexibility, or hierarchy, but also how teams sustain disciplined thinking, judgment, and learning under uncertainty and pressure.

4.5 Propositions

Proposition 1. Under high task novelty, organizations with established problem-framing and assumption-review routines will demonstrate stronger anticipatory resilience than organizations relying primarily on hierarchical reporting structures.

Proposition 2. Research-intensive teams using evidence-review and cross-functional interpretation routines will cope with disruption more effectively than teams relying on fragmented expert judgment.

Proposition 3. Organizations maintaining structured reflection and decision-trace routines will adapt faster after failure or external shocks than organizations relying on informal memory.

Proposition 4. Digital and AI-supported tools enhance resilience when they amplify collective reasoning quality, but weaken it when they substitute speed for judgment or create overconfidence.

4.6 Visual Summary of the Framework

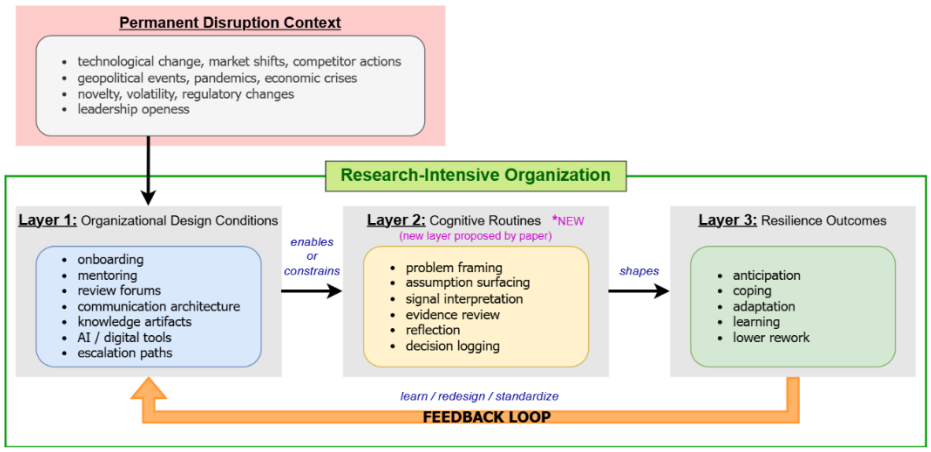


Fig. 1. Conceptual framework linking organizational design conditions, cognitive routines, and resilience outcomes under permanent disruption.

5 Discussion and Implications

5.1 Theoretical Implications

The framework reframes resilience as an enacted and design-sensitive process, rather than only structural property or broad organizational capability. Prior research has emphasized flexibility, communication resources, learning capacity, and adaptive capability as important sources of resilience [1]-[4], [20]. These remain valuable perspectives, but they do not always specify how resilience is produced in everyday organizational life.

The present model suggests that resilience depends partly on how organizations structure interpretation, evaluation, escalation, and learning. The framework focuses on cognitive processes, while recognizing that their effectiveness may be influenced by interpersonal dynamics and other pressures that accompany disruptive situations. In that sense, resilience is not only something organizations possess, but also something organizations repeatedly do. Cognitive routines represent an intermediate layer linking design conditions to adaptive outcomes.

More specifically, the framework suggests a pathway in which organizational design conditions support cognitive routines, cognitive routines facilitate collective sensemaking and learning, and these processes contribute to adaptive capacity and resilience over time.

For organizational design research, the implication is that design choices shape not only structure, coordination mechanisms, incentives and formal processes, but also how organizations interpret information, challenge assumptions, integrate dispersed expertise, and learn over time. Choices such as onboarding, review forums, mentoring, communication pathways, and decision-memory mechanisms may influence the quality and speed of collective judgment under uncertainty.

The paper is also relevant to current debates on intelligent organizations. Intelligence should not be equated with the mere presence of advanced analytics or AI systems. Organizational intelligence depends on whether technological tools strengthen or weaken human capacities for sensemaking, challenge, learning, and coordinated interpretation [21]-[24]. This view is consistent with Society 5.0 perspectives, which emphasize human-centered organizational development, collective intelligence, and the integration of technological and human capabilities instead of technology-driven transformation alone [26].

5.2 Practical Implications for R&D Organizations

First, onboarding should teach not only procedures and formal roles, but also how the organization frames problems, evaluates evidence, handles disagreement, escalates concerns, and reflects after decisions [9]-[13]. Second, mentoring systems should be viewed not only as talent development tools, but also as channels for transferring experience-based judgment. Third, organizations should create and sustain forums for collective reasoning. Examples include pre-mortems, post-project reviews, assumption-challenge meetings, evidence reviews, and cross-functional learning sessions. Such

routines can reduce avoidable rework, preserve learning from failure, and improve response quality during disruption. Fourth, leaders should evaluate organizational performance not only through efficiency metrics, but also through indicators of cognitive quality. These may include speed of issue escalation, diversity of perspectives considered, quality of post-project learning, decision transparency, and continuity of organizational memory.

5.3 Sociotechnical and AI-Aware Extension

The framework also highlights the sociotechnical nature of resilient organizations. Human cognition and technical systems should be treated as interdependent rather than separate domains [7], [8], [21]. Digital platforms and AI-supported systems may improve resilience by increasing visibility of emerging issues, preserving institutional memory, accelerating access to expertise, and supporting scenario analysis [21]-[24]. However, technological sophistication does not automatically generate organizational intelligence. AI may strengthen cognitive routines when it improves analytical reach while preserving human judgement, but may weaken them when it creates opacity, noise, superficial responsiveness, false certainty, or overconfidence.

The design question is not whether organizations should adopt AI, but how AI should be embedded within routines of interpretation, review, escalation, and learning.

6 Conclusion

This paper examined resilience in research-intensive teams operating under conditions of novelty, ambiguity, and non-routine work. It argued that resilience depends not only on formal structure or adaptive capability, but also on how teams collectively process uncertainty, interpret signals, evaluate evidence, coordinate judgment, and learn through disruption.

To address this issue, the paper proposed a three-layer conceptual framework linking organizational design conditions, cognitive routines, and resilience outcomes. Within this model, design choices influence how recurring cognitive routines emerge, stabilize, and support anticipation, coping, and adaptation over time.

Several limitations should be acknowledged. As a conceptual paper, the framework has not yet been empirically validated. In addition, the paper follows an integrative conceptual approach rather than a systematic review methodology, meaning that the selection and interpretation of literature streams may reflect analytical choices made by the author. The framework may also have practical limitations. The value of cognitive routines depends on how they are designed and enacted in practice. Poorly designed routines may encourage excessive deliberation, reinforce dominant viewpoints, or divert attention from timely action. Hence, their effectiveness is likely to depend on organizational context, task characteristics, and the balance between reflection and execution. Cognitive routines should therefore be viewed as enabling structures rather than universally beneficial practices under all conditions.

Future research should therefore examine how cognitive routines emerge, stabilize, and change across research-intensive teams through qualitative case studies, comparative research across R&D settings, and mixed-method designs. It may also operationalize cognitive routines through observable indicators such as shared diagnosis speed, escalation quality, cross-functional knowledge sharing intensity, decision documentation, repeated error reduction or avoidable rework, and adaptation speed after failed experiments. The framework is expected to be most relevant in environments characterized by novelty, ambiguity, distributed expertise, and non-routine problem solving. Its explanatory value may be lower in highly standardized settings where performance depends primarily on execution consistency and procedural stability. Future research could further examine these boundary conditions across contexts such as hospitals, military units, startups, consulting firms, public agencies.

When disruption becomes permanent, resilient organizations are distinguished not only by sound structure and strategy, but by their repeated capacity to think, judge, and learn together.

Disclosure of Interests. The author has no competing interests to declare that are relevant to the content of this article.

Acknowledgments. The author used AI-assisted language tools (ChatGPT, OpenAI) for copyediting purposes only, including grammar correction, language refinement, and improving clarity of expression. All conceptual development, arguments, interpretation, and final content are the sole responsibility of the author and are based on the author's own research and professional experience.

References

1. Ma, Z., Xiao, L., Yin, J.: Toward a dynamic model of organizational resilience. Nankai Business Review International 9(3), 246–263 (2018)
2. Jiang, Y., Ritchie, B., Verreynne, M.L.: Developing disaster resilience: A processual and reflective approach. Tourism Management 87, 104374 (2021)
3. Teo, W.L., Lee, M., Lim, W.S.: The relational activation of resilience model: How leadership activates resilience in an organizational crisis. Journal of Contingencies and Crisis Management 25(3), 136–147 (2017)
4. Tisch, D., Galbreath, J.: Building organizational resilience through sensemaking: The case of climate change and extreme weather events. Business Strategy and The Environment. 27(8), 1197–1208 (2018)
5. Savioja, P., Norros, L., Salo, L., Aaltonen, I.: Identifying resilience in proceduralised accident management activity of NPP operating crews. Safety Science 68, 258–274 (2014)
6. Al-Atwi, A.A., Amankwah-Amoah, J., Khan, Z.: Micro-foundations of organizational design and sustainability: The mediating role of learning ambidexterity. International Business Review 30(1), 101656 (2021)
7. Amir, S., Kant, V.: Sociotechnical resilience: A preliminary concept. Risk Analysis 38(1), 8–16 (2018)
8. Taysom, E., Crilly, N.: Resilience in sociotechnical systems: The perspectives of multiple stakeholders. She Ji: The Journal of Design, Economics, and Innovation 3(3), 165–182 (2017)

9. Michel, A.: A distributed cognition perspective on newcomers' change processes: The management of cognitive uncertainty in two investment banks. Administrative Science Quarterly 52(4), 507–557 (2007)
10. Lapointe, É., Vandenberghe, C., Boudrias, J.S.: Organizational socialization tactics and newcomer adjustment: The mediating role of role clarity and affect-based trust relationships. Journal of Occupational and Organizational Psychology 87(3), 599–624 (2014)
11. Kowtha, N.R.: Engineering the engineers: Socialization tactics and new engineer adjustment in organizations. IEEE Transactions on Engineering Management 55(1), 67–81 (2008)
12. Gallagher, E.B., Sias, P.M.: The New Employee as a Source of Uncertainty: Veteran Employee Information Seeking about New Hires. Western Journal of Communication 73(1), 23–46 (2009)
13. Korte, R.F., Brunhaver, S.R., Sheppard, S.D.: (Mis)interpretations of organizational socialization: The expectations and experiences of newcomers and managers. Human Resource Development Quarterly (2015), https://doi.org/10.1002/hrdq.21206
14. Ratzmann, M., Pesch, R., Bouncken, R., Martinez Climent, C.: The price of team spirit for sensemaking through task discourse in innovation teams. Group Decision and Negotiation 27(3), 321–341 (2018)
15. Acikgöz, A., Günsel, A., Kuzey, C., Zaim, H.: Team foresight in new product development projects. Group Decision and Negotiation 25(2), 289–323 (2016)
16. Beverland, M.B., Micheli, P., Farrelly, F.J.: Resourceful sensemaking: Overcoming barriers between marketing and design in NPD. Journal of Product Innovation Management 33(5), 628–648 (2016)
17. Coplien, J.O., Harrison, N.B.: Organizational Patterns of Agile Software Development. Prentice Hall, Upper Saddle River (2004)
18. Weick, K.E.: Sensemaking in Organizations. Sage Publications, Thousand Oaks (1995)
19. Feldman, M.S., Pentland, B.T.: Reconceptualizing organizational routines as a source of flexibility and change. Administrative Science Quarterly 48(1), 94–118 (2003)
20. Sutcliffe, K.M., Vogus, T.J.: Organizing for resilience. In: Cameron, K.S., Dutton, J.E., Quinn, R.E. (eds.) Positive Organizational Scholarship, pp. 94–110. Berrett-Koehler Publishers, San Francisco (2003)
21. Naikar, N., Brady, A., Moy, G., Kwok, H.W.: Designing human-AI systems for complex settings: ideas from distributed, joint, and self-organizing perspectives of sociotechnical systems and cognitive work analysis. Ergonomics 66(11), 1669–1694 (2023). https://doi.org/10.1080/00140139.2023.2281898
22. Schmutz, J.B., Outland, N., Kerstan, S., Georganta, E., Ulfert, A.S.: AI-teaming: Redefining collaboration in the digital era. Current Opinion in Psychology 58, 101837 (2024)
23. di Prisco, D., Dello Russo, S.: Sensemaking and AI: Unraveling individuals' reactions to the black box in a three-study investigation. Technological Forecasting and Social Change 226, 124491 (2026)
24. Shahzad, F., Ferreira, J.J.M.: Artificial intelligence driven dynamic capabilities: Unpacking micro-foundational mechanisms. Management Decision (2026). https://doi.org/10.1108/MD-07-2025-2141
25. Rakićević, Z., Anđelić, O., Popović, G., Branković, B.: The influence of COVID-19 on production operations planning. Management: Journal of Sustainable Business and Management Solutions in Emerging Economies 28(2), 23-36 (2023)
26. Tornjanski, V., Čudanov, M.: Towards Society 5.0 Era: Organizational Empowerment of the Sustainable Future. In: Proceedings of PaKSoM 2021, pp. 413-422 (2021)

THEMATIC SECTION

Digital Operations and Logistics Management

Resilient and Intelligent Organizations in Times of Permanent Disruption

SYMORG 2026 • CONFERENCE PROCEEDINGS

Original Research Article *https://doi.org/10.5281/zenodo.23475291*

Distribution strategies in e-commerce: Insights from Temu and AliExpress

Anđela Zdravić[1[0009-0009-8055-2289], Miljana Luković[2[0009-0002-8880-6182], Teodora Đokić[3[0009-0006-3180-8979] and Biljana Cvetić[4[0000-0002-2398-4327]

[1,2,4]University of Belgrade, Faculty of Organizational Sciences, Serbia
[3]Transfera Transport & Logistics, Belgrade, Serbia
andjela.zdravic@fon.bg.ac.rs, miljana.lukovic@fon.bg.ac.rs, djokicteodora6@gmail.com, biljana.cvetic@fon.bg.ac.rs

Abstract. Distribution is one of the most important processes in the functioning of e-commerce, as it directly affects supply chain efficiency, operational costs, and the level of customer satisfaction. The aim of this paper is to analyze and compare distribution strategies and logistics structures in e-commerce, with focus on identifying their key characteristics, advantages, and operational implications, as well as to provide practical insights into the logistics models of Temu and AliExpress. The paper analyses the business models of these two e-commerce platforms and examines how centralized and decentralized logistics structures influence operational control, fulfillment organization, and logistics coordination in cross-border e-commerce. The methodological approach includes a structured literature review and classification of relevant studies published between 2020 and 2026, as well as a comparative analysis of the selected case studies based on predefined criteria. The results indicate that different logistics structures and distribution strategies approaches involve a trade-off between control, cost efficiency, flexibility, and market reach. Centralized models enable higher efficiency, stronger operational control, and more standardized fulfillment processes, while decentralized models provide greater flexibility, scalability, and seller autonomy, but are associated with more variable service performance. The findings further suggest that distribution approaches operate at different distribution channel structures, fulfillment models, logistics capabilities, and platform-enabled logistics coordination.

Keywords: distribution, e-commerce, e-commerce distribution strategies, Temu, AliExpress.

1 Introduction

The rise of the Internet created a new channel of communication and business ventures. The most dominant type is the new sales domain known as e-commerce. E-commerce refers to a form of business activity in which the buying and selling of goods and services are conducted through digital platforms, most commonly via the Internet [42]. This business model enables direct interaction between companies and end users without the necessity of physical presence or reliance on traditional retail channels. The development of e-commerce over the past decade has been character-

ised by rapid expansion, further accelerated by ongoing digitalisation processes and shifts in consumer behaviour [13, 14]. Looking at recent statistical studies and data, the findings indicate that global digital adoption rates nearly doubled in recent years, leading to rapid, continued growth in digital revenues [12]. E-commerce platforms have emerged as the main competitors to traditional sales channels, no longer just an alternative channel, with the global e-commerce market value exceeding four trillion US dollars and projected to approach five trillion in the near future [26]. These findings support the rising trend of mass customer sales generated only from digital sales channels, which in turn, is reflected by the evident structural transformation in which digital channels increasingly act as the primary interface between supply and demand. As stated by [38], one of the key advantages of e-commerce lies in its ability to provide access to global markets while maintaining relatively low entry barriers. Digital sales platforms enable enterprises to reach a wider customer base at lower operational costs, and their easy scalability allows faster expansion without the proportional increase in expenses, creating a significant competitive advantage [49]. Small and medium enterprises are using e-commerce platforms to reduce or eliminate the role of intermediaries, gaining greater control over pricing and customer experience while also improving margins by generating larger profits and lower costs [5]. The increasing share of online sales in total retail, having reached 20% of global sales, confirms the long-term relevance of e-commerce as a core component of the modern economic system [39].

This study aims to address two research questions:

RQ1: What distribution strategies are most commonly used in e-commerce?

RQ2: How do centralised and decentralised logistics structures differ in terms of operational control, fulfillment organization, and logistics coordination in cross-border e-commerce platforms?

To address these research questions, a structured literature review was conducted to identify distribution strategies or fulfillment approaches in e-commerce. A comprehensive answer to the second question was provided through a comparative case study of Temu and AliExpress, which examined how different logistics structures are implemented in practice. The paper combines theoretical and industrial sources in order to answer these two research questions.

The paper is structured into six sections. After the introductory section, Section 2 provides a theoretical background on distribution strategies in e-commerce. Section 3 presents the methodological approach applied in this paper. Further, in Section 4, the results of the analysis of distribution strategies in e-commerce are provided. Section 5 presents the distribution strategies of Temu and AliExpress, with the aim of collecting and synthesising data on each enterprise's distribution strategies to conduct and present the results of their comparative analysis. The paper concludes with a final section summarising the main findings of the research.

2 Theoretical background: Distribution in e-commerce

In this section, fundamental terms regarding most commonly used business models will be defined and explained. Following this segment, the next part of this section showcases the most known distribution strategies and fulfillment approaches in e-commerce.

2.1 Business models

Today's most companies' business models which can be used in e-commerce are Business-to-Business, Business-to-Consumer, Customer-to-Business and Consumer-to-Consumer. Further, each model is briefly presented.

Business-to-Business (B2B) models involve the exchange of goods, services and information between organisations [44]. Their relationships are long, complex, and based on trust, shared value, and process integration. Digital tools facilitate collaboration, data exchange, and innovation within supply chains, transforming traditional industrial relationships into platform-based ecosystems [9, 44].

Business-to-Consumer (B2C) is defined as a business model in which companies sell products or services directly to customers [44]. Retailers and e-commencers often use this model, where digital platforms enable them to reach a larger number of customers. The models' main characteristics include [45]: large customer base, relatively short sales cycles, and a strong focus on brand development and user experience.

The Customer-to-Business (C2B) model follows a reverse flow in which individual consumers provide products, services, or value to a company [56]. The rise of this model is linked to the growth of digital platforms, the freelancing economy, and the sharing economy. Individuals offer their skills, data, content, or feedback to businesses, which use these inputs for their business needs such as marketing, product development, or operational improvements [56].

Consumer-to-Consumer (C2C) is a business model in which individual consumers sell products or services directly to other consumers through digital platforms [7]. The model gained popularity due to e-commerce platforms, second-hand marketplaces, and the sharing economy. Digital platforms act as intermediaries, supporting communication, payment processing, logistics, and trust between customers. This model features low transaction costs, direct interaction, and heavy dependence on reputation systems and trust mechanisms [7].

Based on these four main models, new business models have emerged, combining base models such as Business-to-Business-to-Consumer, used by most e-commerce platforms. The Business-to-Business-to-Consumer (B2B2C) model connects end consumers indirectly through intermediaries such as platforms or partner networks. This model is crucial in the digital economy, where platforms integrate B2B and B2C processes within a single ecosystem. Value is created through multiple layers of interaction between producers, platforms, and end users. These systems allow companies to manage relationships with business partners and end consumers simultaneously, generating new value within integrated ecosystems. B2B2C models are key to digital

transformation in traditional industries and support scalable platform-based solutions [35].

Modern e-commerce increasingly functions through platform business models. Platforms such as Shein, Amazon, Temu and AliExpress act as B2B2C intermediaries, connecting manufacturers or independent sellers with end consumers through digitally managed marketplace systems. The platform not only enables transactions but also plays an active role in logistics coordination, fulfilment organisation, seller management and the overall customer experience. The level of control a platform has over logistics operations varies between centralised and decentralised e-commerce models, making the platform a key factor in selecting distribution strategies.

2.2 Distribution and fulfillment strategies

In the context of e-commerce, distribution refers to the set of activities involved in storing, processing, and delivering products from sellers to end customers through digital order fulfilment systems [8]. Unlike traditional retail systems, e-commerce distribution relies on digitally coordinated logistics networks that may operate through either centralised or decentralised fulfilment structures, depending on the platform's business and logistics model. Contemporary e-commerce distribution systems are highly dependent on real-time information exchange, inventory visibility, logistics coordination, and data-driven decision-making processes. As a result, distribution is no longer perceived solely as a supporting logistics activity, but as a strategic component of value creation and customer experience management within e-commerce ecosystems [32].

Nowadays, fast delivery options, such as same-day delivery offered by local distributors, have increased consumer expectations, so companies are expected to provide these services. However, previous studies indicate that a significant segment of consumers is willing to tolerate longer delivery times if deliveries remain reliable, predictable, and associated with lower shipping costs [25]. The majority of enterprises recognise that distribution has become a key competitive factor in e-commerce. Companies that optimise their logistics networks and improve delivery performance are better positioned to meet customer expectations and enhance overall operational efficiency [16]. Distribution strategy has become a key factor in gaining competitive advantage, as companies now compete more through fast, reliable delivery and overall customer experience rather than only through product range or pricing. Selecting the appropriate distribution strategy is a complex decision, as it depends on multiple connected factors such as market structure, product features, customer expectations, logistics capabilities, and cost constraints [8]. Some of the main challenges companies face are deciding whether to rely on direct or indirect distribution channels, and determining the required level of warehousing capacity and the warehousing costs that make sense for their business logic.

Selecting an appropriate distribution strategy is a complex managerial decision influenced by the firm's business model, as shown in the previous section. Product portfolio characteristics and the desired level of control over the sales and delivery process. In this context, manufacturers must consider various factors, including the na-

ture of their business activities, the business model applied, the availability of intermediaries (such as distributors, wholesalers, and retailers), product-specific characteristics, and customer needs and expectations [37].

Distribution strategies according to the flow of goods are direct and indirect distribution. Direct distribution is a configuration in which manufacturers deliver products directly to end customers without intermediaries, enabling greater control over pricing, customer experience, and demand visibility [47]. Indirect distribution is a strategy that relies on intermediaries such as wholesalers, distributors, and retailers to reach final customers, allowing broader market access but reducing control over pricing and customer interaction [17].

The distribution strategies cross-docking and omnichannel are also used in e-commerce. Cross-docking allows the direct transfer of items from inbound storage to outbound transportation vehicles, with minimal or no storage, improving efficiency but requiring precise coordination and advanced information systems [54].

Omnichannel combines multiple sales and delivery channels into one unified system, ensuring a seamless customer experience and improved inventory coordination [1]. This strategy can include drop-shipping and click and collect. Drop-shipping is a modern fulfilment model that allows retailers to avoid holding inventory in their warehouses and to have items shipped directly to the customer's address. The orders travel directly from manufacturers to suppliers, who handle storage and delivery, thereby reducing operational costs. The main challenge is the effect on delivery times and lower quality control [36]. Click-and-collect supposes that customers place orders online and pick them up from a physical store or designated pickup location [18].

Last-mile delivery is the final stage of the distribution process, in which goods are delivered to the end customer. It is considered as a critical logistical challenge or problem. Last-mile delivery is often associated with high costs, operational inefficiencies, and increased complexity in e-commerce distribution systems. This problem is often affected by issues such as traffic congestion, failed deliveries and higher labour demands, which reduce overall efficiency. At the same time, it plays a key role in customer satisfaction because it represents the final point of contact between the company and the customer, making speed and reliability essential performance factors [45].

3 Methodology

In this section, a short literature screening and classification procedure review is presented. The first part of this section outlines the steps for creating and replicating the filtering and selection process, while the second part presents the synthesised findings. This section's purpose and goal are to create a theoretical foundation for the rest of the paper in the domain of distribution and fulfilment in e-commerce. The literature search was conducted in the Scopus database using a predefined Boolean search string (Figure 1).

```
TITLE-ABS-KEY( ("e-commerce" OR "cross-border e-commerce") AND ("distribution
  strategy" OR "fulfillment strategy" OR "e-commerce logistics") AND ("cross-border" OR
  fulfillment OR logistics OR distribution) ) AND PUBYEAR > 2019
```

Fig. 1. Boolean string used in Scopus, dated 20. May 2026

In Figure 2. the step-by-step process is shown. The initial search using the Boolean string identified 726 studies. The next step was carried out using the exclusion options provided by the Scopus database search engine. The first round of filters were: language (English), subject area (Business, Management and Accounting, Decision Sciences, Economics, Econometrics and Finance, and Engineering), and document type (articles, conference papers, and book chapters). This reduced the dataset to 330 studies. The next step in reducing the database was to limit the keywords using the inclusion criteria: “e-commerce”, “e-commerces”, “e-commerce logistics”, “marketplace”, “logistics”, “fulfilment”, and " distribution ". As well as applying filters for excluding keywords: “optimisation route”, “vehicle routing problem”, “transportation routes”, “machine learning”, and "path optimisation”.

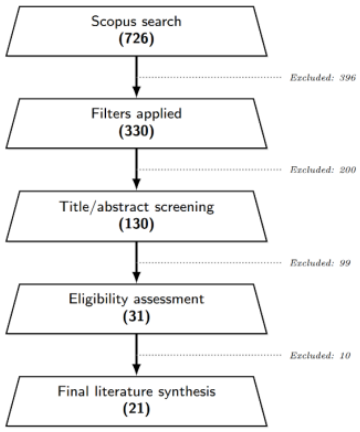


Fig. 2. Methodology filtering and selection process

This resulted in 130 potentially relevant studies, for which the authors manually reviewed the abstracts and applied their own filtering, resulting in 31 studies that underwent a detailed eligibility assessment. Studies were then selected for full review; these 21 studies were retained for detailed literature synthesis and thematic classification.

4 Analysis of distribution strategies in e-commerce

Following the application of the defined multi-stage research methodology, the selected studies were analysed to identify and systematize the main distribution and logistics strategies in e-commerce. As shown in Table 1., the selected papers have been grouped by their research direction based on the distribution strategy or fulfilment method mentioned. This also includes papers that highlighted other key distribution and logistics aspects that are relevant to the e-commerce marketspace. The 21 selected studies retained for detailed reading and analysis were distributed into the following five group categories: distribution channel structures, omnichannel fulfilment models, organization of logistics process flow, logistics process capabilities and platform-enabled logistics coordination.

Table 1. Classification of distribution and fulfilment-related approaches identified in the reviewed literature

Category	Approach	Description	Key characteristics	Typical application in e-commerce	Reference
Distribution channel structures	Direct distribution	Products are delivered directly from manufacturers or retailers to end customers without intermediaries	High control; direct customer relationship; higher operational responsibility	Brand-owned web shops; direct-to-consumer (B2C) models	[16, 23, 31]
	Indirect distribution	Products reach customers through one or more intermediaries	Broader market reach; reduced control; dependence on intermediaries	Online marketplaces and platform-based retailing	[24, 53]
Omnichannel fulfillment models	Drop-shipping	Products are shipped directly from suppliers to customers	No inventory ownership; low capital requirements; supplier-dependent fulfillment	Marketplace platforms and small online retailers	[14, 20, 28, 36]
	Click-and-collect	Customers purchase online and collect products at designated pickup locations	Channel integration; customer convenience; reduced delivery costs	Grocery retailing and omnichannel commerce	[2, 6, 26]
Organization of logistics process flow	Cross-docking	Products are transferred through distribution facilities with minimal storage time	Minimal warehousing; rapid throughput; reduced inventory holding	High-volume e-commerce fulfillment networks	[29, 40]

Category	Approach	Description	Key characteristics	Typical application in e-commerce	Reference
Logistics process capabilities	Last-mile delivery	Final stage of the delivery process from distribution centre to customer	High customer impact; high operational cost; delivery responsiveness	Home delivery and parcel distribution services	[13, 15, 30]
	First-mile and last-mile integration	Coordination of pickup and delivery activities	Improved visibility; network efficiency; delivery optimization	Integrated fulfillment and delivery networks	[34]
Platform-enabled logistics coordination	Centralized logistics structures	Logistics activities are coordinated and controlled by the platform	High visibility; standardized service; stronger operational control	Platform-managed fulfillment systems (e.g., Temu)	[23, 51]
	Decentralized logistics structures	Logistics activities are distributed among independent actors	Flexibility; scalability; lower platform responsibility	Marketplace ecosystems (e.g., AliExpress)	[24, 41, 43]

Direct distribution enables firms to retain full control over pricing, customer relationships, and overall experience, which strengthens customer loyalty and responsiveness, but requires substantial operational capabilities. In contrast, indirect distribution expands market coverage and reduces logistical complexity by leveraging intermediaries, albeit at the cost of diminished control and weaker customer engagement. Fulfillment model such as drop-shipping illustrates a shift toward flexibility and cost minimization, particularly for small and medium-sized enterprises, while maintaining acceptable service levels. There is a substantial number of papers exploring the last-mile delivery bottleneck, a logistics capability problem that is one of the main areas affecting customer satisfaction. The papers exploring the centralised and decentralised logistics structures of platforms have provided substantial theoretical understanding of the main differences between the two.

5 Case study analysis: Insights from Temu and AliExpress

In order to create a well-rounded and multi-dimensional overview of how distribution strategies are used in e-commerce, in this paper a short comparative analysis is conducted between two major players in today's e-commerce. These two enterprises, Temu and Aliexpress were chosen based on three major criteria: their market share in today's e-commerce domain, the portfolio of different products that can be sold and bought on them, as well as the main business model being B2B2C platforms. In the following subsections each enterprise is represented by their core values, differences as well as strategy indicators. The first subsection is focused on Temu's distribution

strategy, while the second one is focused Aliexspress distribution strategy. In the final subsection the results of comparative analysis of Temu and AliExpress are given.

5.1 Distribution strategy of Temu

The distribution strategy of Temu is based on a highly centralised platform controlled by e-commerce model that integrates elements of direct distribution and demand-driven coordination mechanisms. The platform operates through a digitally coordinated B2B2C marketplace model, which minimizes the need for traditional intermediaries and allows products to be offered at lower prices to end consumers. This approach significantly reduces costs and allows the platform to offer highly competitive prices [8, 52]. The business model combines elements of business-to- business (B2B) and business-to-consumer (B2C), allowing demand-driven production and closer alignment between consumer preferences and business processes [48]. This demand is oriented approach allows the platform to continuously adapt its product selection and inventory levels in response to evolving customer demand trends [21].

In terms of fulfillment, Temu relies on a structure minimizing warehousing costs while also improving inventory flexibility and enabling faster responses to customer demand. Although the platform has recently introduced external fulfillment centers in key markets, its core logistics model remains centered on direct shipping from manufactures to customers, which minimizes storage costs and enhances efficiency [33]. Temu's logistics system is centrally managed and tightly coordinated by the platform, ensuring close alignment between suppliers, logistics partners, and customers. At the same time, it relies heavily on third-party logistics or 3PL providers for operational execution, in order to ensure flexibility and scalability of their business. Although coordination is centralized, operational execution remains highly dependent on external logistics partners, which enable the platform to scale internationally quickly without owning a full logistics network [55].

This capability ensures continuous visibility of order status, thereby significantly improving the overall customer experience [52]. Despite its rapid growth and strong market penetration driven by aggressive pricing strategies, concerns remain regarding the long-term sustainability of this business model, particularly due to high operational costs and pressure on suppliers [22]. The model of Temu shows that centralized logistics coordination can improve delivery responsiveness and customer experience, but at the same time it increases operational complexity.

5.2 Distribution strategy of AliExpress

Company AliExpress use a decentralised e-commerce model in which independent sellers are primarily responsible for inventory management, fulfillment activities and logistics coordination [3]. The platform act as a global B2B2C marketplace which connects manufacturers and small to medium-sized businesses with customers in international markets.

AliExpress is strongly associated with the drop-shipping approach, enabling sellers to offer products without maintaining inventory. Drop-shipping should be seen as an

order fulfillment approach rather than a separate distribution strategy, since the platform largely depends on intermediary-based logistics coordination and third-party managed delivery systems. This reduces entry barriers for small and medium-sized enterprises but often results in longer delivery times and limited control over service quality [19]. In terms of logistics, the platform relies extensively on third-party logistics (3PL) providers and international shipping networks.

AliExpress integrates ordering, payment, and logistics services within a unified digital ecosystem, supported by solutions such as Alipay. AliExpress uses advance seller tools and the "AliExpress Labeling Service," in order to streamline their logistics operations and enhance, localized, and boost efficiency of their fulfillment processes [4]. To enhance last-mile delivery, AliExpress collaborates with international courier services and local delivery providers, improving accessibility and customer reach across global markets. The efficiency of its distribution system is still limited by longer lead times, variability in delivery performance and dependence on external logistics coordination. This structure enables scalability and extensive international reach, it also limits platform control over standardized delivery processes and consistent customer experience. The AliExpress model shows how decentralized logistics structures can support low-cost global operations, but often with expense of delivery speed and fulfillment consistency [55].

5.3 Comparative analysis: Temu vs AliExpress

Temu and AliExpress are selected as comparative case studies because they represent contrasting distribution systems within the same cross-border e-commerce sector, enabling a controlled comparison of centralised and decentralised logistics models. This allows for systematic evaluation of how different platform management structures affect logistics performance, delivery efficiency and customer experience under similar market conditions. The selection is justified by their comparable yet distinct positions in global e-commerce: Temu uses a centralised logistics system, while AliExpress operates under a decentralised, seller-led fulfilment model.

The comparative analysis is organised into three analytical levels: strategic, operational, and performance. The strategic level of analysis includes criteria such as the most commonly used transportation strategy, demand orientation, relationship with sellers, and the key competitive advantage obtained. At the operational level, factors such as logistics models, logistics management models, operational control, and fulfilment organisations are used as key criteria for logistics efficiency in cross-border e-commerce systems. The final analytical level groups criteria by performance level. The performance of each company was compared based on: delivery time window, customer satisfaction impact, service consistency and overall scalability.

This set of criteria ensures a comprehensive and comparable evaluation from operational, strategic, and customer-oriented perspectives. This classification ensures conceptual consistency by clearly separating structural, operational and outcome-based dimensions of analysis. The structure of the overall supply chain and the flexibility of their networks are analysed to compare differences between centralised and decentral-

ised distribution models and their effect on scalability. Table 2. presents the results of the comparative analysis between Temu and AliExpress.

Table 2. Results of comparative analysis [3, 4, 10, 11, 33, 39, 46, 50, 52, 55]: Temu vs AliExpress

Analytical level	Criterion	Temu	AliExpress
Strategic level	Transportation strategy	Predominantly air transport	Predominantly maritime transport
	Demand orientation	Demand-driven production and inventory planning	Greater reliance on existing seller inventories
	Relationship with sellers	Platform-managed approach with stronger coordination	Seller-managed approach with greater autonomy
	Key competitive advantage	Faster delivery and ultra-low prices	Global reach and extensive seller network
Organizational level	Logistics model	Centralized logistics supported by coordinated fulfillment operations and 3PL partners	Decentralized logistics supported by independent sellers and 3PL partners
	Logistics management model	Higher platform involvement in logistics coordination	Logistics activities coordinated by sellers and external partners
	Operational control	Higher platform control and visibility	Greater seller autonomy and flexibility
	Fulfillment organization	Standardized fulfillment processes	Seller-dependent fulfillment processes
Performance level	Delivery time performance	5–10 days	15–45 days
	Customer satisfaction impact	Higher satisfaction due to shorter and more predictable delivery times	More variable satisfaction due to heterogeneous fulfillment arrangements
	Service consistency	Higher	More variable
	Scalability	Moderate	High

AliExpress represents a more established platform with a decentralised, intermediary-driven model, while Temu, as a newer entrant, adopts a more centralised, vertically integrated approach with a strong emphasis on speed and supply chain optimisation. This structural contrast provides a suitable analytical basis for assessing the impact of distribution design on logistics performance and customer experience.

The Figure 3. shows how demographic and behavioural factors influence when users start adopting Temu and AliExpress. The results show that on AliExpress, the most important demographic factors, especially gender and middle age, which means that in the early phase of platform development exactly those characteristics of users had a key role. Demographic factors are of minor importance for Temu, but the strongest factor is early adoption, i.e., a short purchase interval. The development of digital platforms has changed the way user bases are formed, from a model based on demographics to one based on behaviour, where purchasing routines are very important in explaining when someone will start using a platform [27].

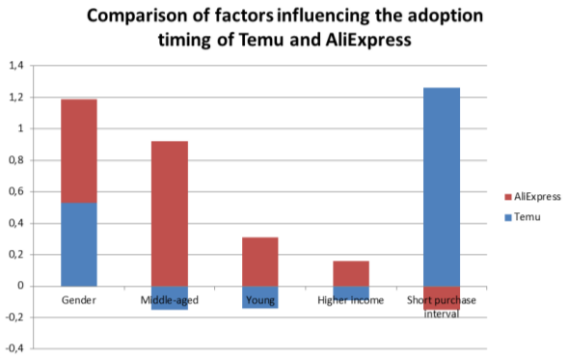


Fig. 3. Comparison of factors influencing the adoption timing of Temu and AliExpress

Temu is very dominant in the fashion industry, with 41%, almost double that of any single category on AliExpress (Figure 4.). This shows a younger, impulse-buying profile. AliExpress shows a more balanced allocation on the graph, with fashion and electronics sharing first place with 24% of the total score, which maintains a wide customer base (Figure 4.). AliExpress includes, of course, food at 1%, which Temu does not, and AliExpress has larger market shares than Temu in DIY and furniture [10, 11].

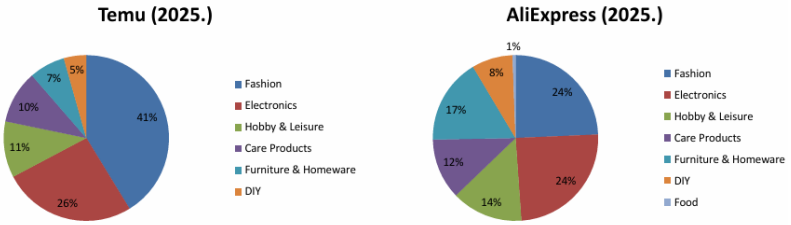


Fig. 4. Category breakdown for Temu and AliExpress

On Figure 5. Temu is the leader in revenue in both periods of observation, whether it is annual or monthly revenue. Temu's annual sales reached 92.5 million dollars in 2025. Temu is recording 38% higher monthly revenue than AliExpress, and Temu's revenue in April 2026 was 9.5 million dollars. This means Temu is growing. Both graphs show results, and it can be seen that Temu is better. However, the gap is growing, and this is a key indicator that Temu is much more aggressive in gaining market share [10, 11].

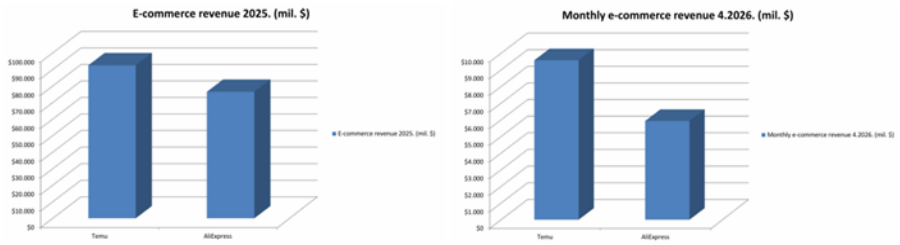


Fig. 5. Annual and monthly revenue of Temu and AliExpress

6 Conclusion

In the modern, highly challenging business environment, the distribution process is crucial to e-commerce operations. Given its direct connection to delivery, which affects service quality, operating costs, and market positioning, it is essential to ensure the unhindered execution of this process. The successful implementation of distribution in e-commerce includes the accurate delivery of shipments to end users, which significantly affects customer satisfaction. By selecting the right distribution strategies, or by combining them in the context of e-commerce shipment distribution, fast delivery to end customers is ensured, and consequently a high level of customer satisfaction. A properly structured distribution system enables companies to respond to market demands in real time, improve logistics flows, and minimise operational costs, while simultaneously increasing consumer satisfaction and loyalty.

The goal of this study was to address two research questions:

RQ1: What distribution strategies are most commonly used in e-commerce?

Using a literature screening and classification procedure methodology shown in Section 3 and filtering the Scopus database, the 21 relevant research papers are selected which were then analysed and classified in Section 4. The analysis revealed five main groups of paper types:

- Papers highlighting the fundamental distribution structures in e-commerce: direct and indirect distribution;
- Papers focusing on omnichannel fulfilment models most commonly used in e-commerce: drop-shipping and click and collect;
- Papers providing insight into how cross-docking is used in e-commerce as a change in the logistics flow process;
- Papers exploring logistics process capabilities with a focus on two major distribution problems known as last mile delivery and first mile delivery;
- Paper highlighting the two main types of platform-enabled logistics coordination: centralised and decentralised.

The answer to RQ1: In e-commerce, the two main focus points in terms of distribution are the direct distribution strategy and indirect distribution strategy. The next main focus in e-commerce is omnichannel fulfilment models: dropshipping and click-

and-collect. The short literature review also suggests that researchers are more focused on distribution and logistics problems than on the basic distribution types used in e-commerce.

RQ2: How do centralised and decentralised logistics structures differ in terms of operational control, fulfillment organization, and logistics coordination in cross-border e-commerce platforms?

The case study analysis of business practices and distribution models used by companies Temu and AliExpress provides a direct answer to RQ2. The findings are showcased in Section 5., provide practical data and examples of how business results for companies using centralised and decentralised logistics structures differ primarily in terms of operational control, fulfilment organisation and logistics coordination.

Therefore, the findings of the case study analysis suggest that centralised logistics structures benefit by prioritising control, consistency, and customer experience, whereas decentralised structures provide companies with greater flexibility, scalability and marketplace openness. There is no single answer to which structure model is better. The choice depends on the company's focus, business strategy, and main goal.

This paper has several limitations that should be acknowledged and considered when interpreting the authors' answers to the research question. The first type of limitation is focused on the literature review analysis. The selection process focused on papers that defined distribution strategy models used in e-commerce, excluding those focused on new conceptual models or theoretical approaches. This paper's case study analysis is limited to just two selected industry leaders chosen by the authors. Both e-commerce platforms can be classified as B2B2C models, with large product portfolios and multiple vendors. The authors acknowledge that, although Temu and AliExpress are highly relevant, they do not fully represent the diversity of global e-commerce platforms. The paper's limitations proved a good starting point for further research directions.

One future research direction can focus on providing a more detailed and insightful literature review of distribution strategies and new conceptual models under discussion in the research community. On the other hand, an empirical study can be conducted by collecting data via surveys or interviews with supply chain professionals. Future research could expand the scope of case study analysis by including additional e-commerce platforms, enabling a broader comparative perspective, and providing a deeper overview of how major players in e-commerce operate.

References

1. Akbariani, N.V., Setiyowati, H.: Digital Meets Distribution: Quantifying the Impact of Logistics Integration in Emerging Market Omnichannel Retail. Data J. Inf. Syst. Manag. 2, 4, 253–263 (2024). https://doi.org/10.61978/data.v2i4.920.
2. Al Oraini, B. et al.: The Future of Online Retail to Omnichannel Commerce in E-commerce Evolution. In: Reyad, S. and Hannoon, A. (eds.) Frontiers of Human Centricity in the Artificial Intelligence-Driven Society 5.0. pp. 377–392 Springer Nature Switzerland, Cham (2024). https://doi.org/10.1007/978-3-031-73545-5_32.

3. AliExpress: AliExpress Standard Shipping: Guide to Reliable Delivery - AliExpress Blog, https://www.aliexpress.com/p/blog/article/aliexpress_standard_shipping_guide_to_reliable_del-426018.html, last accessed 2026/04/22.
4. AliExpress: AliExpress Unveils Powerful New Tools for U.S. Sellers to Boost Growth and Efficiency, https://www.prnewswire.com/news-releases/aliexpress-unveils-powerful-new-tools-for-us-sellers-to-boost-growth-and-efficiency-302611764.html, last accessed 2026/04/22.
5. Aswar, K., Ermawati: E-Commerce Adoption by Small Medium Enterprises: An Extensive Literature Review, Information Management and Business Review, 12(4(I)):12-18 (2021). https://doi.org/10.22610/imbr.v12i4(I).3123.
6. Belvedere, V. et al.: Getting the most from E-commerce in the context of omnichannel strategies. Ital. J. Mark. 2021, 4, 331–349 (2021). https://doi.org/10.1007/s43039-021-00037-6.
7. Borwankar, S. et al.: Information Design in Consumer-to-Consumer (C2C) Marketplaces: The Roles of Anonymity and Product Quality Uncertainty. Prod. Oper. Manag. 10591478261434645 (2026). https://doi.org/10.1177/10591478261434645.
8. Do, A.D. et al.: The Impact of the Quality of Logistics Services in E-Commerce on the Satisfaction and Loyalty of Generation Z Customers. Sustainability. 15, 21, 15294 (2023). https://doi.org/10.3390/su152115294.
9. Dwivedi, Y.K. et al.: Social Media Adoption, Usage And Impact In Business-To-Business (B2B) Context: A State-Of-The-Art Literature Review. Inf. Syst. Front. 25, 3, 971–993 (2023). https://doi.org/10.1007/s10796-021-10106-y.
10. ECDB: E-Commerce Data & Analytics Based on Billions of Transactions, https://ecdb.com/resources/sample-data/retailer/aliexpress, last accessed 2026/05/31.
11. ECDB: Temu Company & Revenue 2015-2027, https://ecdb.com/resources/sample-data/retailer/temu, last accessed 2026/05/31.
12. Gosling, S. et al.: What do US consumers want from e-commerce deliveries? | McKinsey, https://www.mckinsey.com/industries/logistics/our-insights/what-do-us-consumers-want-from-e-commerce-deliveries, last accessed 2026/04/22.
13. Gruchmann, T. et al.: Do you feel guilty? A consumer-centric perspective on green nudging in last-mile deliveries. Int. J. Phys. Distrib. Logist. Manag. 55, 5, 540–566 (2025). https://doi.org/10.1108/IJPDLM-07-2024-0286.
14. Iglesias-Pradas, S., Acquila-Natale, E.: The Future of E-Commerce: Overview and Prospects of Multichannel and Omnichannel Retail. J. Theor. Appl. Electron. Commer. Res. 18, 1, 656–667 (2023). https://doi.org/10.3390/jtaer18010033.
15. Janjevic, M., Winkenbach, M.: Characterizing urban last-mile distribution strategies in mature and emerging e-commerce markets. Transp. Res. Part Policy Pract. 133, 164–196 (2020). https://doi.org/10.1016/j.tra.2020.01.003.
16. Kawa, A., Zdrenka, W.: Logistics value in e-commerce and its impact on customer satisfaction, loyalty and online retailers' performance. Int. J. Logist. Manag. 35, 2, 577–600 (2023). https://doi.org/10.1108/IJLM-02-2023-0078.
17. Kiessling, T. et al.: Supply Chain Disruptions and Need for Resilience: SMEs Direct/Indirect Exporting and Rapid Internationalization. IEEE Trans. Eng. Manag. 72, 115–133 (2025). https://doi.org/10.1109/TEM.2024.3514701.
18. Li, C., Shi, X.: Optimisation of a multilevel logistics network for prepositioned warehouses under an omni-channel retail model, Humanities and Social Sciences Communications, 11, 970 (2024). https://doi.org/10.1057/s41599-024-03496-2.

19. Li, D.: A Comparative Analysis of B2C Cross-border E-commerce Platforms: the AliExpress and the DHgate, International Conference on Education, Management, Computer and Society. (2016). 10.2991/emcs-16.2016.107.
20. Li, G. et al.: Reselling or drop shipping: Strategic analysis of E-commerce dual-channel structures. Electron. Commer. Res. 20, 3, 475–508 (2020). https://doi.org/10.1007/s10660-019-09382-3.
21. Liu, S.: Research on the growth mechanism of temu cross-border e-commerce under the low-profit, high-volume sales model. Glob. Media Soc. Sci. Res. J. 6, 375–381 (2025). https://doi.org/10.71465/gmssrj111.
22. Liu, T.W.: Temu Is Losing Millions of Dollars to Send You Cheap Socks | WIRED, https://www.wired.com/story/temu-is-losing-millions-of-dollars-to-send-you-cheap-socks/, last accessed 2026/04/22.
23. Mashalah, H.A. et al.: The impact of digital transformation on supply chains through e-commerce: Literature review and a conceptual framework. Transp. Res. Part E Logist. Transp. Rev. 165, 102837 (2022). https://doi.org/10.1016/j.tre.2022.102837.
24. Matsui, K.: Should competing suppliers with dual-channel supply chains adopt agency selling in an e-commerce platform? Eur. J. Oper. Res. 312, 2, 587–604 (2024). https://doi.org/10.1016/j.ejor.2023.06.030.
25. McKinsey: Moving past e-commerce to NeXT commerce | McKinsey, https://www.mckinsey.com/capabilities/growth-marketing-and-sales/our-insights/becoming-indispensable-moving-past-e-commerce-to-next-commerce, last accessed 2026/04/22.
26. Milioti, C. et al.: Modelling consumers' acceptance for the click and collect service. J. Retail. Consum. Serv. 56, 102149 (2020). https://doi.org/10.1016/j.jretconser.2020.102149.
27. Minjung, R., Hyunkyu, J.: Evolution of adoption determinants in cross-border e-commerce: comparing AliExpress and Temu. Asia Pac. J. Mark. Logist. 38, 4, 1217–1238 (2025). https://doi.org/10.1108/APJML-01-2025-0028.
28. Mohd Dali, N.R.S. et al.: e-Fulfilment Affecting e-Loyalty Mediated by e-Satisfaction and e-Trust in Digital Landscape – Jurnal Pengurusan. J. Pengur. 75, (2025). https://doi.org/https://doi.org/10.17576/pengurusan-2025-75-7.
29. Mohd, F. et al.: Global Research Trend in E-commerce Logistics: A Bibliometric Analysis. In: El Khoury, R. (ed.) Technology-Driven Business Innovation: Unleashing the Digital Advantage: Volume 2. pp. 291–301 Springer Nature Switzerland, Cham (2024). https://doi.org/10.1007/978-3-031-62656-2_26.
30. Nagadeepa, C. et al.: Last-Mile Delivery Innovations: The Future of E-commerce Logistics. In: El Khoury, R. (ed.) Technology-Driven Business Innovation: Unleashing the Digital Advantage, Volume 1. pp. 283–296 Springer Nature Switzerland, Cham (2024). https://doi.org/10.1007/978-3-031-51997-0_24.
31. Pan, Y.: Typical Cases of E-commerce Logistics. In: Handbook of E-commerce in China. pp. 491–535 Springer, Singapore (2025). https://doi.org/10.1007/978-981-96-7629-3_28.
32. Pramudita, A.S., Guslan, D.: The Impact of Logistics Service Quality on Consumer Satisfaction in E-Commerce Distribution Channels in Indonesia, Journal of Distribution Science 23(2):109-118, (2025). 10.15722/jds.23.02.202502.109
33. Raimkulov, A.: Fast, Cheap, Global: The Supply Chain of Temu, https://www.allthingssupplychain.com/fast-cheap-global-the-supply-chain-of-Temu/, last accessed 2026/04/22.
34. Ranathunga, M.I.D. et al.: Solution approaches for combining first-mile pickup and last-mile delivery in an e-commerce logistic network: A systematic literature review. In: 2021

International Research Conference on Smart Computing and Systems Engineering (SCSE). pp. 267–275 (2021). https://doi.org/10.1109/SCSE53661.2021.9568349.
35. Rejikumar, G., Asokan-Ajitha, A.: Business customer experience in B2B2C service settings: a scale development. J. Bus. Ind. Mark. 38, 4, 886–904 (2023). https://doi.org/10.1108/JBIM-07-2021-0342.
36. Risberg, A.: A systematic literature review on e-commerce logistics: towards an e-commerce and omni-channel decision framework. Int. Rev. Retail Distrib. Consum. Res. 33, 1, 67–91 (2023). https://doi.org/10.1080/09593969.2022.2089903.
37. Roshani, A. et al.: Designing resilient supply chain networks: a systematic literature review of mitigation strategies. Ann. Oper. Res. 341, 2–3, 1267–1332 (2024). https://doi.org/10.1007/s10479-024-06228-6.
38. Savitri, L. et al.: E-commerce and Financial Performance: Identification on Trends, Benefits and Challenges, ATESTASI Jurnal Ilmiah Akuntansi 6(1):528-543, (2023). 10.57178/atestasi.v6i1.931
39. SellersCommerce: 51 ECommerce Statistics In 2025 (Global And U.S. Data) | SellersCommerce, https://www.sellerscommerce.com/blog/ecommerce-statistics/, last accessed 2026/04/22.
40. Shamout, M.D. et al.: Streamlining E-Commerce and Last-Mile Delivery for Enhanced Customer Satisfaction: An In-Depth Analysis of Amazon's Strategies. In: 2024 2nd International Conference on Cyber Resilience (ICCR). pp. 1–8 (2024). https://doi.org/10.1109/ICCR61006.2024.10532856.
41. Shen, H., He, Y.: The geopolitics of infrastructuralized platforms: the case of Alibaba. Inf. Commun. Soc. 25, 16, 2363–2380 (2022). https://doi.org/10.1080/1369118X.2022.2128599.
42. Shi, B. et al.: E-commerce platform's customer-to-manufacturer mode choice: self-built or joint-development. Enterp. Inf. Syst. 18, 10, 2406015 (2024). https://doi.org/10.1080/17517575.2024.2406015.
43. Shi, T. et al.: Distribution Channel Choice and Divisional Conflict in Remanufacturing Operations. Prod. Oper. Manag. 29, 7, 1702–1719 (2020). https://doi.org/10.1111/poms.13185.
44. Skare, M. et al.: Innovation activity and the outcomes of B2C, B2B, and B2G E-Commerce in EU countries. J. Bus. Res. 163, 113874 (2023). https://doi.org/10.1016/j.jbusres.2023.113874.
45. Tahirov, N. et al.: Configuration of last-mile distribution networks for an encroaching manufacturer. Ann. Oper. Res. 344, 2–3, 679–720 (2025). https://doi.org/10.1007/s10479-024-06031-3.
46. The Global Statistics: Ecommerce Statistics 2025 | Online Shopping Stats – The Global Statistics, https://www.theglobalstatistics.com/ecommerce-stats/, last accessed 2026/04/22.
47. Tsantis, A. et al.: Trade shocks and direct shipping connections: causal insights into network adaptability and supply chain resilience. WMU J. Marit. Aff. (2026). https://doi.org/10.1007/s13437-025-00399-0.
48. Uchańska-Bieniusiewicz, A.: Innovations with sustainability potential within the ultra-fast business model. Corp. Gov. Int. J. Bus. Soc. (2024). https://doi.org/10.1108/CG-01-2024-0051.
49. Van Gelder, K.: eCommerce - Worldwide | Statista Market Forecast, https://www.statista.com/outlook/emo/ecommerce/worldwide, last accessed 2026/04/22.
50. Walsh, M.: Ecommerce Statistics (2025) — Key Facts and Figures, https://www.ecommercetrix.com/ecommerce-statistics/ecommerce-statistics/, last accessed 2026/04/22.

51. Wang, J. et al.: Effects of platform's blockchain strategy on brand manufacturer's distribution strategy in the presence of counterfeits. Comput. Ind. Eng. 177, 109028 (2023). https://doi.org/10.1016/j.cie.2023.109028.
52. Xiang, L.: Innovative Trends in Temu's Logistics Strategies, https://www.jusdaglobal.com/en/article/innovative-trends-temu-logistics-strategies/, last accessed 2026/04/22.
53. Yoon, Y.L. et al.: Buyer-supplier matching in online B2B marketplace: An empirical study of small- and medium-sized enterprises (SMEs). Ind. Mark. Manag. 93, 90–100 (2021). https://doi.org/10.1016/j.indmarman.2020.12.010.
54. Zhang, Z.: The Impact of Digital Transformation on Logistics Efficiency. Adv. Econ. Manag. Polit. Sci. 127, 1, 25–31 (2024). https://doi.org/10.54254/2754-1169/2024.OX18254.
55. Zhou, J.: Challenges and Countermeasures of Cross-Border E-commerce in International Trade – Based on the Perspective of Logistics. Adv. Econ. Manag. Polit. Sci. 161, 86–94 (2025). https://doi.org/10.54254/2754-1169/2025.19889.
56. Zhu, Y.: The Comparative Analysis of C2B and B2C. Int. J. Mark. Stud. 7, 5, p157 (2015). https://doi.org/10.5539/ijms.v7n5p157.

THEMATIC SECTION

E-Business Ecosystems and Technologies

Resilient and Intelligent Organizations
in Times of Permanent Disruption

SYMORG 2026 • CONFERENCE PROCEEDINGS

Original Research Article *https://doi.org/10.5281/zenodo.23475335*

Contemporary E-Commerce And AI Ecosystem – theoretical and practical issues

Sonja Vučenović 1[0000-0003-2375-5759] Nikola Macura 1[0009-0006-9917-6998] and Marija Pavličević 1[0009-0003-9066-9655]

1 University of Novi Sad, Faculty of economics in Subotica, Serbia
sonja.vucenovic@ef.uns.ac.rs

Abstract. This paper analyzes the transformation of the modern e-commerce ecosystem with a special focus on the role of artificial intelligence (AI) as a key driver of innovation and change in digital business. The development of digital technologies, changes in consumer behavior and the growth of platform-oriented models have led to a transition from traditional, centralized systems to more flexible and modular approaches, such as composable commerce. In such an environment, the e-commerce ecosystem is becoming increasingly data-driven, omnichannel and based on the integration of various actors and technologies.
The paper specifically considers the application of artificial intelligence in various segments of e-commerce, including personalized product recommendations, conversational artificial intelligence and chatbots, dynamic pricing, visual search and virtual product try-on, as well as supply chain management and digital content generation. The analysis of these technologies indicates their importance in improving the user experience, optimizing business processes and making decisions based on large amounts of data.
In addition to identifying key advantages, the paper also deals with the challenges that accompany the application of AI technologies, including issues of data privacy, transparency of algorithms and ethical responsibility. Through the presentation of practical examples, the research demonstrates the concrete effects of the implementation of artificial intelligence on business performance and the competitive position of organizations.
The paper emphasizes that the successful integration of artificial intelligence in e-commerce requires a strategic approach, continuous improvement of technological capacities and careful risk management, which ensures a sustainable competitive advantage in the digital economy.

Keywords: e-commerce, ecosystem, artificial intelligence.

1 Introduction

In the modern conditions of digital transformation, electronic commerce represents one of the most dynamic segments of the global economy, characterized by continuous technological innovations and changes in consumer behavior. The development of information and communication technologies, the growth of digital platforms and the increasing availability of data have led to the redefinition of traditional business models, whereby e-commerce has grown from a simple sales channel into a complex, integrated

digital ecosystem. In this context, organizations are faced with the need to continuously improve their digital capabilities in order to remain competitive in an environment characterized by high dynamics and intense market competition.

The modern e-commerce ecosystem is undergoing a significant transformation due to the accelerated development of digital technologies, changes in consumer behavior and the growing integration of platform business models. Traditional e-commerce models, which relied heavily on centralized and monolithic platforms, are gradually transforming towards more flexible and modular architectures known as composable commerce. This approach allows organizations to combine various digital components, services and application interfaces (APIs) to create scalable and adaptable e-commerce systems. According to Gartner (2023), composable commerce enables companies to innovate faster and adapt to changing market demands thanks to modular digital infrastructure.

One of the most significant drivers of change in the e-commerce ecosystem is the application of artificial intelligence (AI) and big data analytics. AI technologies enable highly personalized shopping experiences through recommender systems, automated communication with users, predictive analytics and price optimization.

The further development of the e-commerce ecosystem is characterized by the increasingly widespread use of augmented reality (AR) and virtual reality (VR) in online retail. These technologies allow consumers to visually try products in a digital environment before purchasing, thereby reducing uncertainty and increasing confidence in online shopping.

At the same time, the e-commerce ecosystem is becoming more and more platform-oriented, with digital platforms functioning as central nodes that connect different actors – manufacturers, distributors, logistics partners, wholesalers, retailers and end users. In this context, the concept of a digital ecosystem refers to a network of interconnected actors and technological resources that jointly create and exchange value.

Based on the above, the goal of this work is the analysis of the transformation of the e-commerce ecosystem with a special focus on the role and application of artificial intelligence. The paper seeks to identify key technological trends, explain the basic concepts of AI application in e-commerce, as well as to point out their effects on business models and user experience through practical examples. Additionally, the research aims to consider potential challenges and limitations in the application of these technologies, thereby contributing to a better understanding of the contemporary trends in the development of the digital economy. Based on the analysis of relevant scientific literature and the findings of previous research, the following research questions were formulated:

- How does digital transformation influence the development and restructuring of the modern e-commerce ecosystem?
- In which ways does artificial intelligence contribute to the optimization of operational processes in e-commerce organizations?
- How do AI-based technologies such as recommendation systems, conversational AI, dynamic pricing, visual search, and intelligent supply chain management influence consumer behavior and user satisfaction?

- What are the key benefits and challenges associated with the implementation of artificial intelligence in e-commerce ecosystems?
- How do issues related to data privacy, algorithmic transparency, and ethical responsibility affect the adoption and perception of AI technologies in e-commerce?
- To what extent can artificial intelligence contribute to the long-term sustainability and competitiveness in e-commerce?

Based on the established research questions, the main research hypothesis was formulated as follows:

H0: The application of artificial intelligence significantly transforms the e-commerce ecosystem by improving operational efficiency, customer experience, and organizational competitiveness.

2 Methodology

This research is based on a qualitative methodological approach aimed at examining the role and application of artificial intelligence within the contemporary e-commerce ecosystem. In order to achieve the defined research objectives, methods of literature review, content analysis, conceptual mapping, and case study analysis were employed.

The first phase of the research involved a systematic review and analysis of relevant scientific literature, and contemporary theoretical and empirical studies related to e-commerce ecosystems and the application of artificial intelligence technologies. The literature analysis provided the theoretical foundation for identifying key concepts, technological trends, opportunities, and challenges associated with AI implementation in e-commerce.

In addition, a qualitative content analysis was conducted in order to examine existing business practices and identify dominant patterns of AI application within electronic commerce. The analysis included examples of AI implementation by leading global retailers and major e-commerce companies recognized for the advanced integration of artificial intelligence technologies into their business.

Based on the findings obtained through the literature review and content analysis, a conceptual mind map was developed to systematize and visually present the relationships between the key elements of the e-commerce ecosystem, AI technologies, business processes, and user experience dimensions. The mind map served as an analytical framework for understanding the interconnections between AI and e-commerce.

Furthermore, the research incorporated a case study approach through the analysis of concrete examples of AI application in e-commerce practice. Particular attention was given to the implementation and effects of technologies such as recommendation systems, conversational artificial intelligence, dynamic pricing, visual search, and intelligent supply chain management. The selected case studies included examples of globally recognized retailers and e-commerce platforms that actively utilize AI technologies to improve operational efficiency and customer engagement.

The study also relied on secondary statistical data and existing market research findings concerning the effectiveness and business impact of AI implementation across dif-

ferent segments of e-commerce. The analysis of secondary data enabled a broader understanding of current adoption rates and consumer responses, and the strategic significance of AI technologies in the digital economy.

The applied methodological framework provides a comprehensive overview of the contemporary application of artificial intelligence in e-commerce and establishes a basis for future empirical and practical research focused on the effects, potentials, and limitations of AI integration in digital business ecosystem.

3 Results

In the modern digital environment, artificial intelligence (AI) is becoming one of the key drivers of the transformation of business models, especially in the domain of electronic commerce and digital platforms. The development of advanced algorithms for data processing, machine learning and computer vision has enabled companies to more efficiently analyze user behavior, optimize business processes and improve user experience. A particularly important role is played by technologies that enable personalization of the offer, automated communication with users, intelligent pricing, improved interaction with products and optimization of operational processes. In this context, various AI technologies — including personalized recommendation systems, conversational artificial intelligence, dynamic pricing, visual search and virtual product testing, as well as intelligent supply chain management — are becoming an integral part of modern digital ecosystems and are the subject of intense academic and practical research (Figure 1).

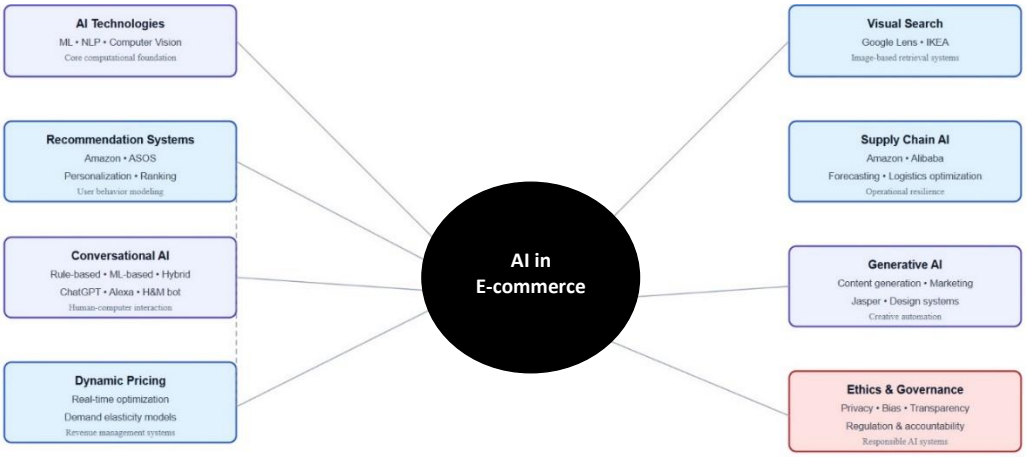


Fig. 1. Conceptual map of the application of artificial intelligence in electronic commerce

Personalized Product Recommendations represent a set of methodological and technological procedures by which users in a digital environment (e.g. e-commerce, streaming platforms, online advertising) are presented with products or services adapted to their individual preferences, needs and behavior patterns (Ricci, et al., 2011). Enhancing AI

algorithms is essential to support big data analytics and decision-making among users within an online digital space, supporting pattern identification, personalized suggestions, and experience, ensuring tailored customer experience and involvement (Jadiga, 2025). This concept relies on an interdisciplinary framework that includes fields such as information technology, machine learning, data analytics, marketing, and behavioral economics. The goal of personalized recommendations is to optimize the user experience and increase the relevance of the offer through algorithmic processing of user data.

Conversational AI and Chatbots represent technological systems based on artificial intelligence that enable automated, interactive and contextually adjusted communication between a human and a computer system through natural language. Fundamentally, chatbots are software applications designed to simulate human conversation (Izadi & Forouzanfar, 2024). In contemporary literature, Conversational AI encompasses a broader set of technologies that integrate natural language processing, machine learning, and dialogue management systems to achieve natural human-machine interaction (Adamopoulou & Moussiades, 2020).

AI-based Dynamic Pricing is a pricing strategy in which machine learning algorithms automatically optimize prices in real time or in short time intervals. While traditional dynamic pricing models rely on historical data to determine optimal prices, modern dynamic pricing algorithms leverage an even richer dataset in addition to the capabilities of artificial intelligence and machine learning to achieve optimised dynamic pricing and more accurate market trend forecasting (Thota, et al. 2024). This approach enables adaptive learning from large and heterogeneous data, thereby improving demand forecasting and revenue maximization (Chen, Mislove, & Wilson, 2016). The main goal of this approach can be defined as maximizing revenue and/or profit while simultaneously preserving competitiveness and customer satisfaction. The application of algorithms from the field of Machine Learning and Data Science allows systems to predict demand, identify patterns of purchasing behavior and automatically adjust prices in accordance with changes in the market.

Visual Search and Virtual Try-On represent modern technologies in the field of digital commerce and user interaction with products, based on computer vision and artificial intelligence. The ability to virtually try on products addresses common challenges associated with online shopping, such as uncertainty about fit and style, ultimately enhancing the overall customer experience and satisfaction (Islam, e al. 2024). This approach tends to reduce the cognitive load of the user and increase the precision of the search, enabling more efficient matching of user needs with appropriate products. The main goal of the Virtual Try-On system is to increase user confidence, reduce the rate of product returns and improve the overall user experience. Both technologies tend towards the personalization and digital transformation of the user experience, with the ultimate goal of increasing conversion, user satisfaction and the efficiency of e-commerce.

Supply Chain Management (SCM) and inventory management with the application of artificial intelligence (AI) are key components of modern business systems that integrate advanced analytical methods to optimize the flow of materials, information and resources. The application of AI enables more accurate demand forecasting, inventory optimization, cost reduction, and increased efficiency of the entire supply chain (Choi,

Wallace, & Wang, 2018; Ivanov et al., 2019). However, despite these promising developments, several barriers impede the full-scale adoption of AI in SCM: technological implementation obstacles, disparities in organizational digital maturity and concerns around ethics, transparency and cybersecurity (Daios, 2025).

Generative AI for creating content and visual materials is a group of algorithmic technologies based on artificial intelligence that enable the automatic generation of text, images, graphics, videos and other digital content. These systems rely on deep learning models, especially generative models like Generative Adversarial Networks (GANs) and transformers, which learn patterns from large data sets to create new, coherent and creative content (Goodfellow et al., 2014).

Based on the presented concepts, it can be concluded that the application of artificial intelligence is one of the key factors in the transformation of modern digital business systems. Technologies such as personalized recommendation systems, conversational artificial intelligence, dynamic pricing, visual search and virtual product try-on, as well as advanced solutions for supply chain management and digital content generation, enable organizations to more efficiently manage data, improve customer experience and optimize operational processes. The integration of these technologies creates a basis for making more accurate business decisions based on data, increasing competitive advantage and developing innovative business models. Therefore, further research in this area has a significant role in understanding the potential, but also the challenges of implementing artificial intelligence in a digital business environment.

4 Discusion

In the modern digital economy, artificial intelligence (AI) is becoming one of the key drivers of the transformation of electronic commerce and the way organizations manage interaction with users, business processes and market strategies. The development of advanced algorithms of machine learning, natural language processing and computer vision has enabled companies to analyze large amounts of data and, based on it, automate decision-making, improve the personalization of offers and optimize operational activities. The ever-widening application of artificial intelligence also opens up important questions related to the transparency of algorithms, the protection of data privacy and ethical responsibility in the use of automated decision-making systems, which makes this area an important subject of modern academic research.

The system of personalized recommendations works based on the collection and analysis of various types of data, such as: purchase history, searches and clicks, demographic data, implicit and explicit user preferences. Retailers like Amazon and ASOS use machine learning to analyze browsing/purchase history and suggesting items with higher conversion probability. From an economic point of view, personalized recommendations contribute to increasing sales, reducing information asymmetry and strengthening user loyalty. According to analyses conducted by McKinsey & Company (2021), recommendation systems can contribute between 15% and 30% of the total revenue of e-commerce platforms, while personalized recommendations significantly increase conversion rates and user engagement. However, their application also raises

ethical issues related to data privacy, transparency of algorithms and the possibility of creating so-called "bubble filter", i.e. limiting the user's exposure to various options. Personalized Product Recommendations represent the key mechanism of modern digital platforms, based on analytical data processing and algorithmic decision-making. Their purpose is to improve the user experience by increasing the relevance of the offer, while at the same time balancing between technological efficiency, market goals and ethical standards.

Table 1. Examples of AI personalization connected to product recommendations

AI personalization	Example	Disruption
Personalization tool for ecommerce product recommendations	eSelf.ai	A powerful tool from AI personalization eCommerce examples that provides users with real-time video product advice.
Email personalization	Omnisend	Specializes in multi-channel marketing automation for eCommerce brands.
Personalization tool for ecommerce user-generated content	Loox	Helps eCommerce brands build trust with photo and video reviews.
Tool for behavior-triggered messaging	Active-Campaign	Automates customer journeys with predictive content suggestions.
Tool for personalized search experiences	Algolia	AI-powered search personalization for eCommerce platforms.
AI tool for personalization	Dynamic Yield	AI-driven content and product recommendations.

In the literature, different categories of chatbots are distinguished according to the technological basis and the way of generating responses. Traditional chatbots are based on pre-defined rules and scripts (rule-based systems), whereby they function through deterministic response patterns (Dale, 2016). In contrast, modern AI chatbots use machine learning and data-driven models to recognize patterns in user queries and tailor responses (Adamopoulou & Moussiades, 2020). In addition, hybrid models have been developed that combine a rule-based approach and a learning-based approach, which achieves greater flexibility and accuracy of the system (Dale, 2016). AI agents (e.g., Zendesk, ChatGPT) handle customer service, 24/7, across platforms like WhatsApp and Instagram to resolve issues and provide recommendations.

Conversational AI and chatbots represent a significant area of modern research in the field of artificial intelligence and human-computer interaction. Global adoption of conversational AI systems has increased rapidly in recent years, with estimates suggesting that approximately 16% of the global population actively uses generative AI tools in daily activities, including chatbot-based systems (Microsoft AI Economy Institute, 2025). Their development and application are based on the integration of advanced techniques of natural language processing and machine learning, with the continuous improvement of models for understanding the context and generating responses (Jurafsky & Martin, 2023). These systems have a growing importance in various industrial and social contexts, especially in automating communication and improving user experience. Sephora uses chatbots based on Conversational AI to improve the user experience in online shopping, H&M chatbot provides fashion recommendations based on user preferences, while Amazon through the Alexa system enables voice interaction

and product purchase, and Uber uses a chatbot for customer support and problem solving.

The application of AI in dynamic pricing relies on revenue management, microeconomic theory and decision theory under conditions of uncertainty. Price is viewed as a decision variable in the optimization problem, with models taking into account demand elasticity and consumer willingness to pay (Phillips, 2020). Practical applications include: real-time pricing, personalized pricing and predictive seasonal pricing (Chen et al., 2016; Phillips, 2020). Although AI-dynamic pricing increases efficiency and revenue, it raises issues of transparency, fairness and ethics in the use of data (Bertsimas & Perakis, 2006; Chen et al., 2016). AI algorithms, such as those used by Amazon, optimize prices in real-time based on competitor, pricing, demand, and stock levels. AI-based dynamic pricing systems typically achieve moderate but consistent revenue improvements in the range of approximately 2–10%, demand elasticity, and implementation complexity (Chenavaz & Dimitrov, 2025; Aschersleben & Steiner, 2024).

Visual Search is a technology that allows users to search for products using visual information, usually photos or scanned images, instead of classic text queries. The system analyzes the visual characteristics of the image, such as shape, color, texture and patterns, and finds products with similar attributes in the database. Google is developing the Google Lens tool, which allows users to pre-search for information using photos (eg products, text, landmarks), while IKEA Creative uses advanced visual search and AR.

The application of AI in supply chain and inventory management enables the transformation of traditional methods based on heuristics or statistics into dynamic, adaptive systems that adapt in real time to changes in demand, inventory levels and supply chain disruptions (Ivanov et al., 2019). AI predicts demand, manages stock levels, and optimizes logistics to reduce errors, as seen with Amazon and Alibaba. The application of AI in supply chain and inventory management brings the following benefits: reduction of excess and shortage of inventory, optimization of transportation and storage costs, improved demand forecasting and production planning, increasing agility and resilience of the supply chain to disruptions (Ivanov et al., 2019).

Applying AI in SCM and inventory management requires careful data management, privacy protection and algorithmic transparency. In addition, the complexity of AI systems can increase a company's dependence on technology and require continuous training of employees (Choi et al., 2018). The integration of artificial intelligence into supply chain and inventory management represents a strategic approach to optimizing business processes and increasing competitive advantage. AI enables adaptive and predictive planning, reducing operating costs and increasing supply chain resilience, thus contributing to more efficient and sustainable business in the digital economy (Ivanov et al., 2019).

The application of Generative AI in the creation of content and visual materials enables: automation of creative processes and reduction of production time, generation of personalized and interactive content, innovations in marketing, education, entertainment and design, experimentation with new styles and formats without high production costs. Brands use AI for creating product descriptions (e.g., Jasper) and generating marketing images with models in different outfits (e.g., Zara). The application of generative

AI opens up significant ethical and regulatory challenges, including: the possibility of spreading misinformation and false content, copyright and intellectual property issues of generated materials, potential bias in generated content due to unbalanced data (Bender et al., 2021).

Generative AI for creating content and visual materials represents a revolutionary approach to digital production, integrated with advanced artificial intelligence models. Its application enables automation, creative innovation and personalization of content, but requires careful consideration of ethical, legal and social implications to ensure responsible and sustainable use of this technology (Bender et al., 2021).

Based on the analyzed concepts, it can be concluded that artificial intelligence plays a key role in the transformation of modern digital business models, especially in the field of electronic commerce. Technologies such as personalized recommendation systems, conversational artificial intelligence, dynamic pricing, visual search and the application of AI in supply chain management enable the improvement of user experience, optimization of business processes and decision-making based on data. At the same time, their application raises issues of data privacy, transparency of algorithms and ethical responsibility, which indicates the need for further research and responsible implementation of these technologies in digital business systems.

Despite providing a comprehensive overview of the role and application of artificial intelligence in the contemporary e-commerce ecosystem, this study is subject to several limitations. First, the research is primarily based on a qualitative methodological approach, including literature review, content analysis, and case study analysis. Although these methods enable an in-depth understanding of current trends, technologies, and business practices, they do not provide direct empirical measurement of causal relationships between the implementation of artificial intelligence and organizational performance outcomes. Consequently, the conclusions are predominantly interpretative and exploratory in nature. The study relies significantly on secondary data sources, such as scientific publications, industry reports, market analyses, and publicly available business information. The presented findings reflect the current stage of AI adoption in e-commerce and may require future updates and longitudinal examination.

Finally, the research does not include primary empirical investigation involving consumers, managers, or industry professionals through surveys, interviews, or experimental methods. Future studies could complement the present analysis by incorporating quantitative and mixed-method approaches in order to obtain more detailed empirical insights into user perceptions, organizational readiness, and the measurable impact of AI technologies in e-commerce. Future studies could examine the relationship between AI adoption and organizational performance indicators such as operational efficiency, customer satisfaction, consumer loyalty, and financial performance. Additionally, future research should investigate consumer perceptions, attitudes, and trust toward AI-based technologies in e-commerce environments, especially in relation to personalized recommendation systems, conversational AI, automated decision-making, and data privacy concerns.

Conclusion

The conducted analysis confirms that digital transformation has fundamentally influenced the development of the modern e-commerce ecosystem. The rapid advancement of information and communication technologies, together with the expansion of digital platforms and data-driven business models, has transformed e-commerce from a traditional online sales channel into a complex and interconnected digital ecosystem. In such an environment, organizations are required to continuously adapt their technological infrastructure, business processes, and strategic approaches in order to maintain competitiveness and respond effectively to rapidly changing market demands and consumer expectations.

The research further demonstrates that artificial intelligence plays a central role in optimizing operational processes and improving decision-making within e-commerce organizations. Through the application of AI technologies, companies are able to process and analyze large volumes of data more efficiently, automate repetitive tasks, predict market trends, and support strategic decision-making based on real-time insights. AI contributes significantly to improving operational efficiency, reducing costs, and enhancing the overall agility of organizations operating in digital markets.

Additionally, AI-based technologies such as personalized recommendation systems, conversational AI, dynamic pricing, visual search, and intelligent supply chain management have a substantial impact on consumer behavior and user satisfaction. Personalized recommendations improve customer engagement by offering products and services aligned with individual preferences, while conversational AI enhances communication and customer support availability. Dynamic pricing models allow organizations to respond more effectively to market fluctuations, and visual search technologies simplify the purchasing process and improve user experience. Intelligent supply chain management systems contribute to faster delivery, optimized inventory management, and increased reliability of services, all of which positively influence customer satisfaction and loyalty.

The findings also indicate that the implementation of artificial intelligence in e-commerce ecosystems provides numerous benefits, including higher operational efficiency, improved personalization, increased competitiveness, and enhanced innovation potential. However, alongside these advantages, organizations face significant challenges related to technological complexity, implementation costs, integration of AI systems, and the need for specialized knowledge and digital competencies. These challenges demonstrate that successful AI adoption requires not only technological investments but also organizational readiness and strategic management.

Particular attention must be given to issues concerning data privacy, algorithmic transparency, and ethical responsibility. The increasing reliance on AI technologies raises concerns regarding the collection and processing of personal data, potential algorithmic bias, and the lack of transparency in automated decision-making processes. These factors directly affect consumer trust and the broader acceptance of AI solutions in e-commerce. Therefore, organizations must establish clear ethical standards, ensure compliance with data protection regulations, and promote transparent and responsible use of artificial intelligence technologies.

Finally, the research confirms that artificial intelligence has significant potential to contribute to the long-term sustainability and competitiveness of digital business models in e-commerce. Organizations that successfully integrate AI into their digital ecosystems are better positioned to innovate, respond to market changes, and create superior customer experiences. Consequently, AI is not only a technological tool but also a strategic factor shaping the future development of the digital economy and modern e-commerce ecosystems.

Acknowledgments. The authors wish to thank the editor and the anonymous reviewers for their valuable and constructive comments, which greatly improved the manuscript.

Disclosure of Interests. The authors declares that they have no relevant or material financial interests that relate to the research described in this paper.

References

1. Adamopoulou, E., Moussiades, L.: Chatbots: History, technology, and applications. *Mach. Learn. Appl.* 2, 100006 (2020). https://doi.org/10.1016/j.mlwa.2020.100006
2. Aschersleben, P. & Steiner, W.: Dynamic pricing using flexible heterogeneous sales response models. *OR Spectrum* (Springer) (2024). https://doi.org/10.1007/s00291-024-00756-0 Bender, E.M., Gebru, T., McMillan-Major,
3. Shmitchell, S.: On the dangers of stochastic parrots: Can language models be too big? In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)*, pp. 610–623. ACM, New York, NY (2021). https://doi.org/10.1145/3442188.3445922
4. Bertsimas, D., Perakis, G.: Dynamic pricing: A learning approach. In: Lawphongpanich, S., Hearn, D.W., Smith, M.J. (eds.) *Mathematical and Computational Models for Congestion Charging*. Applied Optimization, vol. 101. Springer, Boston, MA (2006). https://doi.org/10.1007/0-387-29645-X_3
5. Chen, L., Mislove, A., Wilson, C.: Peeking beneath the hood of Uber. In: *Proceedings of the 2016 ACM Conference on Online Social Networks (COSN '16)*, pp. 1–11. ACM, New York, NY (2016). https://doi.org/10.1145/2815675.2815681
6. Choi, T.M., Wallace, S.W., Wang, Y.: Big data analytics in operations management. *Prod. Oper. Manag.* 27(10), 1868–1885 (2018). https://doi.org/10.1111/poms.12838
7. Daios, A.; Kladovasilakis, N.; Kelemis, A.; Kostavelis, I. AI Applications in Supply Chain Management: *A Survey. Appl. Sci.* 2025, 15, 2775. https://doi.org/10.3390/app15052775 Dale, R.: The return of the chatbots. *Nat. Lang. Eng.*22(5), 811–817 (2016) . https://doi.org/10.1017/S1351324916000243
8. Gartner: *Composable Commerce: The Future of Digital Commerce Architecture*. Gartner Re- search, Stamford, CT (2023).
9. Goodfellow, I., et al.: Generative adversarial nets. In: *Advances in Neural Information Pro- cessing Systems (NeurIPS 27)*, pp. 2672–2680. Curran Associates, Inc., Red Hook, NY (2014). https://papers.nips.cc/paper_files/paper/2014/hash/f033ed80deb0234979a61f95710dbe25-Abstract.html
10. Islam, T., Miron, A., Liu, X. & Li, Y.: Deep Learning in Virtual Try-On: A Comprehensive Survey. *IEEE Access* 12, 29475–29502 (2024). https://doi.org/10.1109/ACCESS.2024.3368612

11. Ivanov, D., Dolgui, A., Sokolov, B., Ivanova, M.: Literature review on disruption recovery in the supply chain: From resilience to viability. *Int. J. Prod. Res.* 57(23), 7435–7457 (2019). https://doi.org/10.1080/00207543.2017.1330572
12. Izadi, S., Forouzanfar, M.: Error Correction and Adaptation in Conversational AI: A Review of Techniques and Applications in Chatbots. *AI* 5(2), 803–841 (2024). https://doi.org/10.3390/ai5020041
13. Jadiga, S.: Understanding the role of AI in personalized recommendation systems, applica- tions, concepts, and algorithms. *International Journal of Computer Trends and Technology (IJCTT)* 73(1), 106–118 (2025). https://doi.org/10.14445/22312803/IJCTT-V73I1P113
14. Jurafsky, D., Martin, J.H.: *Speech and Language Processing*, 3rd edn. Prentice Hall, Upper Saddle River, NJ (2023).
15. McKinsey & Company: The value of getting personalization right—or wrong—is multiply- ing. McKinsey & Company (2021). Available at: https://www.mckinsey.com/capabili-ies/growth-marketing-and-sales/our-insights/the-value-of-getting-personalization-right-or- wrong-is-multiplying (accessed 19 May 2026).
16. Microsoft AI Economy Institute: *Global AI Adoption Report.* Microsoft (2025). Available at:https://www.microsoft.com/en-us/corporate-responsibility/topics/ai-economy-institute (ac cessed 19 May 2026).
17. Phillips, R.: *Pricing and revenue optimization*, 3rd edn. Stanford Univ. Press (2020)
18. Régis Y. Chenavaz & Dimitrov, S.: Artificial intelligence and dynamic pricing: a systematic literature review. *International Journal of Production Research* (2025). https://doi.org/10.1080/15140326.2025.2466140
19. Ricci, F., Rokach, L., Shapira, B., Kantor, P.B. (eds.): *Recommender Systems Handbook.* Springer, New York (2011). https://doi.org/10.1007/978-0-387-85820-3
20. Thota, D. V. , Kumar, D. N. M, D. K. K. M, D. S. , Momin U. & Mishra D. P.: E-Commerce Management and AI-Based Dynamic Pricing Revenue Optimization Strategies. *Migration Letters* 21(S4), 168–177 (2024). Available at: https://migrationletters.com/index.php/ml/arti- cle/view/7157 (accessed 19 May 2026).

Original Research Article *https://doi.org/10.5281/zenodo.23475391*

Service Catalogues and Metaregisters as Enablers of Health Service Bus Integration: The ALIMS Case Study

Tatjana Stojadinović[1][0000-0002-8952-2205], Ilija Antović[2][0000-0002-2673-6628], Svetlana Jovanović[2][0009-0000-2648-0389] and Tijana Mladenović[3][0009-0001-0377-9756]

[1] Medicines and medical devices agency of Serbia (ALIMS), Belgrade, Serbia
[2] University of Belgrade – Faculty of Organizational Sciences, Belgrade, Serbia
[3]Medicines and medical devices agency of Serbia (ALIMS), Belgrade, Serbia

Abstract. The digital transformation of healthcare systems requires a structured approach to integrating heterogeneous information systems, regulatory registries, and service interfaces. In the Republic of Serbia, the Medicines and Medical Devices Agency of Serbia (ALIMS) has developed multiple regulatory registries and REST API services; however, their effective integration into the national digital health ecosystem requires standardization, cataloguing, and alignment with interoperability frameworks.

This paper presents a methodology for the development of service catalogues and metaregisters as key enablers for integrating regulatory services into the Government and health service bus. The approach is based on system integration principles, API service management, and registry interoperability, enabling the systematic identification, classification, and exposure of ALIMS services and datasets.

The proposed methodology is validated through the ALIMS case study, where a comprehensive service and registry catalogue was established, and a set of REST API services was prepared for integration with the national service infrastructure. The results include the classification of services into open and secured categories, definition of integration requirements, and improved readiness for interoperability with other public administration and healthcare systems.

The main contribution of this paper is a practical and replicable framework for transforming regulatory information systems into interoperable, service-oriented architectures, supporting the development of integrated digital health ecosystems.

Keywords: Service Catalog, Metaregister, eHealth

1 Introduction

Public institutions generate and maintain large volumes of heterogeneous data. By ensuring their availability through open data initiatives and interoperable digital services, these datasets can be reused to create added value in the form of new economic activities, improved efficiency of public administration, enhanced public services, and increased transparency of government operations [1].

In the healthcare domain, these challenges are particularly pronounced due to the complexity of regulatory processes, the number of stakeholders involved, and the need for reliable and secure data exchange. In the Republic of Serbia, the Medicines and Medical

Devices Agency of Serbia (ALIMS) maintains a number of regulatory registries and exposes selected data through REST API services. However, the lack of a unified approach to service cataloguing, registry standardization, and integration with national interoperability infrastructure limits their full potential.
The aim of this paper is to present a methodology for the development of service catalogues and metaregisters as enablers of integration with the Government Service Bus and the future Health Service Bus within the Republic Integrated Health Information System (RIZIS). The paper addresses the following research question: how can regulatory registries and existing API services be systematically prepared for integration into a service-oriented, interoperable national digital health infrastructure?

This paper presents a methodology-driven case study based on the ALIMS experience in establishing service catalogues, metaregisters, and Government Service Bus integration as prerequisites for future Health Service Bus interoperability within the Republic Integrated Health Information System (RIZIS). The proposed methodology combines registry analysis, metadata standardization, service catalogue development, API governance, and migration of existing REST services to a service-oriented integration platform. The contribution of the paper is twofold: (1) it provides a structured methodology for aligning domain-specific registries and services with the State Administration Meta-Registry and Government Service Bus, and (2) it demonstrates its practical application through the migration and exposure of ALIMS services and registries. Although validated in the pharmaceutical regulatory domain, the proposed approach can be replicated in other healthcare and public administration systems facing similar interoperability and integration challenges.

1.1 Methodology

The proposed methodology is based on the application of system integration principles, API service management, and registry interoperability, with a particular focus on the registries maintained by ALIMS.

The methodology includes two main phases:

1. Establishment of the ALIMS Service and Registry Catalogue, aligned with the State Administration Meta-Registry, enabling structured identification, classification, and documentation of regulatory datasets and services;
2. Migration and exposition of existing ALIMS REST API services on the Government Service Bus (Enterprise Service Bus – ESB), as preparation for their future integration into the Health Service Bus within the RIZIS framework.

This approach ensures the standardization of service interfaces, improves data exchange capabilities, and provides a foundation for interoperability with other public administration and healthcare systems.

2 Previous research and legal aspects

This chapter presents the legal and conceptual framework of open data, interoperability, metaregisters, and data exchange.

Open data and content can be freely used, modified, and shared by anyone for any purpose [2]. Ensuring the availability of open data enables improved interoperability, more efficient data exchange, and supports the digital transformation of the healthcare system.

Interoperability is defined as the ability of two or more systems to exchange data and to use that data after the exchange [13]. It represents a prerequisite for connecting heterogeneous systems, such as healthcare registries operating on different technological platforms [14]. Interoperability is therefore a key enabler for the establishment of metaregisters.

The main objectives of interoperability in healthcare systems include:

- centralization and standardization of data,
- enabling data exchange between information systems,
- improving efficiency and accessibility of e-services,
- compliance with national eGovernment and EU standards (EIF, FHIR).

At the same time, it is necessary to ensure data security, clearly defined authorization mechanisms, and compliance with personal data protection regulations.

A service catalogue represents a central repository of available services within the system. It contains descriptions of e-services, API specifications, available registers, and data sources, enabling service discovery, interface definition, and standardized service usage. The concept is based on Service-Oriented Architecture (SOA), and includes elements such as service name, functionality description, input and output parameters, communication protocols, and versioning. These principles are well established in the literature on web services and service-oriented architectures [17, 18, 19].

A service catalogue must support metadata management, search and filtering, version control, access management, and API documentation publication.

All registry services can be organized within metaregisters [3]. Metaregisters represent structured catalogues of registries, datasets, and associated services, providing contextual information about data sources, purpose, governance, technical characteristics, and usage models [3,15]. This approach is aligned with established metadata management and business intelligence frameworks [15,16].

Unlike primary data registries, metaregisters do not store primary data but metadata describing the structure, meaning, and use of data. This enables consistent interpretation, flexibility, and easier integration across systems [3,16].

Service-oriented metaregisters enable controlled access to data through standardized interfaces, depending on user roles and permissions. They transform isolated data sources into interoperable services, supporting a user-centric approach to data access [3].

Metaregisters must support record management, linking with external registers, version tracking, validation and verification, and data export (XML, JSON, CSV). Their

development includes mapping key registries, aligning metadata with European standards, and defining a technical framework for data exchange. In this context, metaregisters represent a foundation for transparent and efficient digital governance.

The Health Service Bus represents an integration platform for data exchange, service orchestration, and API management. As a central communication layer based on SOA principles, it enables integration of heterogeneous systems, message transformation, authentication and authorization, monitoring, and transaction management.

Modern healthcare information systems face challenges in integrating a large number of heterogeneous applications, including hospital information systems, laboratory systems, electronic health records, and health insurance systems. The Health Service Bus addresses these challenges by enabling interoperability and efficient system communication.

The Health Service Bus enables:

- data exchange between healthcare systems,
- service orchestration,
- message transformation and routing,
- secure communication.

It represents a domain-specific implementation of Enterprise Service Bus architecture, characterized by loose coupling, service reuse, scalability, flexibility, and centralized communication control.

2.1 Legal aspects of health registries in the European, Serbian and global environment

A health registry is a data management system used for collecting health data to assess the outcomes of a population characterized by a certain disease, condition, or risk, and it serves a predetermined scientific, clinical, or administrative purpose [3,12].

Additionally, it should be clarified that the term “health registry” is often used simultaneously for IT or paper solutions and systems for collecting and storing data from the health system, as well as the health record that is thus created [3,12].

There are three main areas into which we can divide the reasons for the establishment of health registries – those that serve to monitor a) a medical product such as a drug, biomaterial, or medical device, b) a disease or health condition, c) the quality of health services or their combinations [3,12].

However, the quality and consistency of data are often not satisfactory, which leads to the impossibility of comparing data. Even in cases where the data are comparable, the cost of such data collection often outweighs the benefit [3,12].

Currently, there are a number of national registers in the world, as well as international, regional, and cross-border health meta-registers. The project activities of these meta-registers, as well as their products, often overlap, leading to unnecessary multiplication and redundancy of metadata. There is also the problem of insufficient interoperability regarding the technical standards used in the creation of metadata models, i.e., metadata schemas; disagreement on the legal basis for cross-border construction of metaregisters, data exchange, and guaranteed data privacy; multiple initiatives covering

the same area; and unclear criteria for providing incentives to meta-register owners as well as financial support to the meta-registers themselves. There is a need for a unified and systematically regulated meta-register at the European (and possibly global) level. A European health meta-register could serve all residents of the European Union regardless of specific national, cultural, or political factors [3].

As an example of successful health metaregistry practices, we highlight meta-registries in areas such as cancer (EUROCourse) [5], diabetes (EUBIROD) [4], arthroplasty (EAR-EFORT) [6], rare diseases (EPIRARE) [7] and (IRDiRC) [8], multiple sclerosis (EUReMS) [9], orphan drugs (EBE Europabio Joint Task Force (JTF) on Rare Diseases and Orphan Medicinal Products) [10], and RD-CONNECT [11].

The success of these registries has been ensured through long-term planning and influencing the targeted population of patients and registry owners by gradually producing political decisions and prevention programs. All data requests directed from targeted user populations can be viewed as services that the registry, as an information system, provides to the user [3]

Health registries in European Union countries produce large amounts of data that are not interchangeable and, due to a lack of interoperability, cannot reach the rightful users of that data – researchers, professionals, and patients. Data from health registries are viewed as isolated islands of information, which makes accessing that data a demanding task. The purpose of a metaregistres is to collect data from data providers on one hand, and on the other hand, to deliver a predefined service to the service user. Interoperability achieved through this approach overcomes traditional problems of data exchange by handling services intended for end-users, rather than the data itself. At the European Union level, one such metaregistres has the working title PARENT (Patient Registry Initiative) 'Registry of All Registries.' This meta-registry will attempt to unite all significant European patient registries into a single service-oriented model [3].

The European Health Registries PARENT is conceived as a web-interface-mediated repository of metadata about European registries, a European service meta-registry of health registries. The meta-registry database would be available online to all interested users (patients, doctors, researchers, etc.). The PARENT web platform would enable the registration and logging of registries into a single shared 'directory' of registries and subsequent editing of metadata depending on user permissions. The PARENT registry will also be used for creating new standards for the use of metadata, comparisons, evaluations, and improvement of registry practices (benchmarking) and provide a knowledge base on existing registry practices in healthcare. The platform will also include a set of IT tools for managing business processes and metadata, as well as an interface for creating new business rules. The PARENT registry is not just a list of existing registries but also a set of methodologies and information architectures, business rules, and legal recommendations that make up a knowledge system of metaregistres. As a system, it is modular, scalable, flexible, dynamic, and transparent. A service-oriented approach ensures a holistic and rational approach to metadata management and ensures the sustainability of the metaregistry system. Moreover, the application of business rules and business process management enables connections with various stakeholders and propagates norms that ensure interoperability. The PARENT metaregistry will be populated with metadata from several sources: a) the aforementioned existing

European metaregistres in healthcare, b) partner countries collaborating with the PARENT project, c) individual national registries [3].

In Serbia in accordance with the Law on Health Documentation and Records in the Field of Healthcare ("Official Gazette Serbia", No. 92/2023) - Article 45, ALIMS shall establish a software solution, technological infrastructure (network, software, and hardware components), organization, and procedures for collecting, storing, processing, preserving, transmitting, displaying, and using data and information on medicines and medical devices, as well as health documentation such as registers of medicines and medical devices, which it is responsible for maintaining in electronic form.

The data controller for medicines and medical devices, i.e., the register of medicines and medical devices, is the Agency for Medicines and Medical Devices (Law on Health Documentation and Records in the Field of Healthcare, Official Gazette No. 92/2023, Article 42).

The health service bus plays a key role in connecting: healthcare institutions (health centers, hospitals), laboratories, pharmacies, health insurance funds and drug agencies Enables real-time data exchange, unique access to patient data, reduction of administrative procedures and increased efficiency of healthcare services.

Organizations such as the World Health Organization emphasize the importance of interoperability as a key factor for the development of digital health.

In practice, a service directory allows health systems to: find the appropriate service (e.g., insurance verification), use standardized interfaces and reduce duplication of functionalities. Metaregisters and service directories are closely related to the service bus and represent its logical layer.

The role of the metaregistry in the bus enables: a unified interpretation of data, synchronization between systems, elimination of semantic inconsistencies and without a metaregistry, data exchange would be hindered due to different interpretations of the same information.

The role of the service catalog in the bus enables: registration of all available services, dynamic service discovery, versioning and availability management and standards such as OASIS UDDI allow the implementation of such catalogs.

The synergy of these components is as follows: the metaregistry defines the meaning of data, the service catalogue defines how data is accessed and used, and the service bus enables data exchange. These three elements form the foundation of an interoperable eHealth system.

3 Research approach

3.1 Metaregistry and Catalog services of ALIMS

The ALIMS metaregister system represents a structured metadata repository describing medicines, medical devices, licenses and regulatory decisions, manufacturers, and marketing authorization holders.

The ALIMS service catalogue and metaregisters are aligned with the State Administration Meta-Registry. The Ministry of Health of the Republic of Serbia, in cooperation with the Office for IT and eGovernment, coordinates the establishment of registries required for the development of the Republic Integrated Health Information System (RIZIS). These include registries related to medicines and medical devices, which are owned by the Republic of Serbia, while their usage and management are performed by competent institutions, including ALIMS.

The main objectives of the proposed approach include:

- establishing a metadata-driven governance model over ALIMS registers and services,
- ensuring standardized, secure, and controlled data exchange with other public administration bodies and healthcare information systems,
- compliance with the requirements of legal, semantic, and technical interoperability in accordance with the Law on e-Government and the Law on Health Documentation and Records,
- preparing ALIMS services for service orchestration within the future Health Service Bus.

The implementation of this system is intended to provide reliable, secure, and standardized data exchange in the healthcare system, while improving the operation of ALIMS and related institutions.

The efficient functioning of a health service bus depends not only on technical infrastructure, but also on the quality of data and service management. In this context, metaregisters and service catalogues represent key components that enable discovery, description, and management of system resources, allowing systems not only to exchange data but also to interpret it consistently.

Key functions of metaregisters include:

- data standardization,
- semantic interoperability,
- centralized management of code lists,
- support for data exchange between systems.

In the healthcare system, metaregisters typically include: diagnosis codes (e.g., ICD classifications), healthcare service codes and electronic medical record structures.

In the drug regulatory system, metaregisters typically include: drug codes (e.g., generic drug names), active substance codes, structures of marketing authorization holders for drugs and structures of drug manufacturers…

In eHealth systems, the integration of metaregisters and service catalogues enables key functionalities such as the exchange of electronic health records, access to laboratory results, electronic prescription issuance, and data exchange between healthcare institutions. Standards such as HL7 FHIR support structured and interoperable data

exchange between systems [20]. In the Republic of Serbia, the institutional framework for such integration is being developed with the support of the Office for IT and eGovernment, which implements centralized registries and integration platforms.

The integration of metaregisters and service catalogues with a service bus provides several benefits, including improved interoperability, standardized data exchange, simplified system maintenance, and increased efficiency of healthcare services.

Challenges and Limitations are: complexity of metadata management, the need for constant registry updates, harmonization of different standards and data security and protection.

The proposed solution to these problems lies in service access to health registers and the data stored in them.

The data thus leaves the framework of the registers within which they are collected and is transformed into pre-defined packages of services according to the requirements of the end users, whose interaction is made possible by the concept of the metaregistry as a catalogue of services. One such model of a service-oriented metaregistres as a catalogue of services considered as a real possibility and need [3].

3.2 Project phases

The project phases for such a complex system include both information technology and the regulatory component, from project initiation, analysis of the current state, to design and implementation as in state eHealth systems and interoperable systems, taking into account that ALIMS already uses registers and systems such as RIMS (Regulatory Information System) for medicines and ADIS (Agency Digital Information System) for medical devices. The project phases can be divided into:

1. Phase 1 - Analysis and establishment of the service and register catalogue
2. Phase 2 - Migration of REST API services to the state authorities' service bus

The first phase includes: a comprehensive analysis of existing registers, data, records, and information systems of ALIMS (ADIS, RIMS), identification of all participants and defining the regulatory framework, defining the model of the Service Catalogue and registers, aligned with the structure and metadata of the State Administration Meta-registry, ensuring semantic alignment of the Catalogue with the Metaregister, defining non-functional requirements such as security, performance, scalability, interoperability, defining documentation for each register and service: legal basis, functional scope and ownership of data, datasets and data domains with clear indication of source and origin of data, exchange conditions and access policies, institutions and roles with access rights (authorized consumers), processes of data entry, modification, and lifecycle management (data lifecycle management), terms of use and restrictions in accordance with regulations, perform domain clustering of registers into

three thematic units: medicines in human medicine, medicines in veterinary medicine, medical devices in human medicine, and medical devices in veterinary medicine.

The result of this phase represents a formally verified catalogue of services and registries ready for integration with the State Administration Meta-Registry, in accordance with the e-Government Act and the definition of the Meta-Registry as a unified public electronic register of all electronic registries established in the Republic of Serbia.

The second phase includes the migration of REST API services to the Service Bus of state authorities, namely:

1. technical analysis of existing REST API services related to the Register of Medicines and Medical Devices, namely: register of issued quality control certificates for veterinary medicines, register of issued quality control certificates for medicines used in human medicine, register of import permits for unregistered medicines, register of approved clinical trials, register of medicines for veterinary use, register of medicines for human use, register of medical devices for human use, register of veterinary medicines with withdrawal periods and types of animals, register of import permits for unregistered medicines, register of clinical trials of human medical devices, register of import permits for unregistered medical devices, register of foreign manufacturers of medical devices, register of representatives of medical devices, register of medicines for which advertising authorization has been issued, register of veterinary medical devices and register of generic medical devices in human use.
2. Standardization of API interfaces in accordance with the rules of API governance and onboarding procedures of the Service Bus of state authorities,
3. Migration of a total of 24 services (15 open and 9 protected), including drug registers, medical devices, import permits, quality certificates, and clinical trials,
4. Implementation of an authentication and authorization mechanism (API keys, authentication),
5. Ensuring encrypted communication (HTTPS/TLS),
6. Implementation of centralized logging, monitoring, auditability, and access traces (logging, monitoring, audit trail),
7. Development of a dispatcher – API Gateway service as a single access point to ALIMS services via the ESB,
8. Conducting integration testing, security tests, and interoperability tests,
9. Preparation of updated technical documentation (OpenAPI specifications, integration guidelines),
10. Introduction in a limited environment of several healthcare institutions as a pilot project, validation of data exchange, accuracy of catalogues and registries.

The result of this phase should ensure that all listed ALIMS REST API services are operational, controlled, and accessible through the Government Agencies Service Bus by applying SOA – API-first architectural principles. Access to services through the Government Service Bus will provide high availability, scalability, and fault tolerance of the system.

4 Results

A representative example of the achieved results is the availability of open datasets on registered medicines for human and veterinary use, provided by the Medicines and Medical Devices Agency of Serbia (ALIMS). As one of the early adopters of the open data concept in the public sector, ALIMS has contributed to improving transparency and enabling data reuse across multiple domains, including the economy, public administration, and the scientific community (see Fig. 1).

The registers of medicinal products are published in structured, machine-readable formats (e.g., CSV and JSON) and are exposed as electronic services through the national eGovernment portal (see Fig. 2). These datasets have enabled the development of third-party applications, including web and mobile solutions (e.g., Mediately), providing up-to-date information on registered medicines in Serbia.

Within the context of the proposed methodology, these open datasets represent an initial layer of service exposure, which can be further structured and integrated through service catalogues and metaregisters. The classification, standardization, and documentation of these datasets enable their transformation into interoperable services, suitable for integration with the Government Service Bus.

The application of the proposed approach contributes to improved interoperability, more efficient data exchange, and enhanced readiness for integration into the broader digital health ecosystem, including future integration with the Health Service Bus.

Open data strengthens entrepreneurship, influencing the development of innovative products and services, providing alternatives for decision-making in management, planning and science, and contributing to the creation of a knowledge-based economy. In addition, open data creates added value for authorities, including better public services, reducing the burden on public administration and increasing cooperation with citizens, businesses and organizations. A centralized provision that relates to open data provides an efficient profit for public sector bodies and improves the quality of data [1].

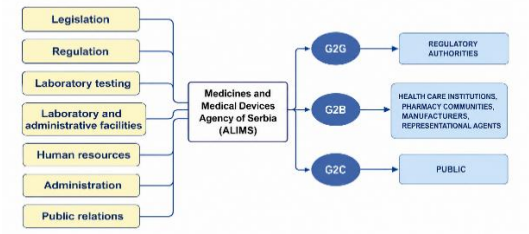


Fig. 1. E-business of ALIMS

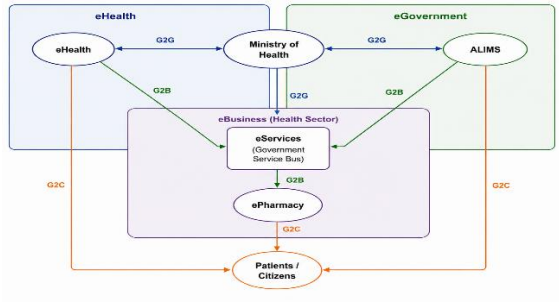


Fig. 2. E-Services of ALIMS and eHealth, ePharmacy

5 Conclusion

Metaregisters and service catalogues represent key enabling components for the integration of heterogeneous systems within the Health Service Bus, providing both technical and semantic interoperability as a foundation for modern eHealth systems. The results of the ALIMS case study demonstrate that the systematic cataloguing of registries and services, combined with the standardization and exposure of REST API services, enables their effective integration into the Government Service Bus and prepares them for future inclusion in the Health Service Bus within the RIZIS framework.

The proposed methodology contributes to the transformation of regulatory information systems into interoperable, service-oriented architectures, supporting improved data exchange, service orchestration, and integration with other public administration and healthcare systems. In this context, metaregisters ensure consistent interpretation of data, while service catalogues enable structured service discovery and management.

Open data guarantees greater transparency in the work of public authorities, stimulates efficiency in public administration, and enables citizens, businesses, and organizations to reuse public information for different purposes. It also supports the development of new products and services, improves decision-making processes, and contributes to the development of a knowledge-based economy [1]. However, its full potential is achieved only when integrated within a structured interoperability framework based on service catalogues, metaregisters, and service bus architecture.

Future developments will further emphasize the role of these components in supporting digital platforms, advanced data management, and artificial intelligence applications, contributing to the development of a fully integrated, data-driven healthcare ecosystem.

The presented approach can be replicated in other regulatory domains and public administration systems facing similar interoperability challenges.

References

1. T. Petrovic, I. Zivkovic, S. Jovanovic, “Used value of open data and open source in Serbia,” Proceedings of the XVIII International Symposium SymOrg 2022, June 11–14, 2022, pp. 170–172. Available: https://symorg.fon.bg.ac.rs/wp-content/uploads/2022/07/Book-of-Abstracts-Symorg2022.pdf
2. Open Knowledge Foundation, “Open Definition 2.1,” 2017. Available: http://opendefinition.org/od/2.1/en/
3. V. Pajić, I. Pristaš, M. Meglič, “Stvaranje registra europskih registara u zdravstvu – servisno orijentirani pristup,” Acta Medica Croatica, pp. 233-239, 2013. Available: https://www.academia.edu/116207950/Stvaranje_Registra_Europskih_Registara_U_Zdravstvu_Servisno_Orijentirani_Pristup
4. European Commission, “European Best Information through Regional Outcomes in Diabetes (EUBIROD),” 2008. Available: http://www.eubirod.eu/
5. European Commission, “European Cancer Registries: EUROCOURSE Project,” 2008. Available: http://www.eurocourse.org/
6. European Federation of National Associations of Orthopaedics and Traumatology (EFORT), “European Arthroplasty Register (EAR),” 2012. Available: http://www.ear.efort.org/
7. European Commission, “European Platform for Rare Diseases Registries (EPIRARE),” 2012. Available: http://www.epirare.eu/
8. European Commission, “International Rare Diseases Research Consortium (IRDiRC),” 2011. Available: http://ec.europa.eu/research/health/medical-research/rare-diseases/irdirc_en.html
9. European Commission, “European Register for Multiple Sclerosis (EUReMS),” 2011. Available: http://www.emsp.org/projects/eurems
10. European Commission, “EBE Europabio Joint Task Force on Rare Diseases and Orphan Medicinal Products,” 2012. Available: http://www.europabio.org/
11. European Commission, “RD-CONNECT: An integrated platform connecting databases, registries, biobanks and clinical bioinformatics,” 2012. Available: http://www.rd-connect.eu/
12. R. E. Gliklich, N. A. Dreyer, Registries for Evaluating Patient Outcomes: A User’s Guide, Rockville, MD: Agency for Healthcare Research and Quality, 2007.
13. Institute of Electrical and Electronics Engineers, IEEE Standard Computer Dictionary, New York: IEEE, 1991.
14. J. O’Brien, G. Marakas, Introduction to Information Systems, 13th ed., McGraw-Hill, 2005.
15. L. T. Moss, S. Atre, Business Intelligence Roadmap, Boston: Addison-Wesley, 2003.
16. D. Marco, M. Jennings, Universal Metadata Models, Wiley, 2004.
17. M. Papazoglou, Web Services and SOA: Principles and Technology, Pearson, 2007 and 2nd edition 2012, ISBN-13: 9780273732167 | Published 2012
18. K. Skarzyński et al., “SOA-based architecture and service discovery mechanisms,” 2017.
19. S. Mittal et al., “DEVS/SOA: A Cross-Platform Framework for Net-centric Modeling and Simulation in DEVS Unified Process,” pp.35, DOI:10.48550/arXiv.2407.03686, License CC BY-SA 4.0, 2024.
20. HL7 International, FHIR Release 4 (R4): Fast Healthcare Interoperability Resources, 2019. Available: https://www.hl7.org/fhir/

Original Research Article *https://doi.org/10.5281/zenodo.23475448*

From Digital Judiciary to AI-Augmented Justice: Applying the Berkana Two Loops Model to the Serbian Judiciary

Dušan Kuzmanović[1,2][0009-0005-6683-3050], Ana Đilas[1][0009-0008-6216-1375], and Zorica Bogdanović[2][0000-0003-4799-1588]

[1] Public Policy Secretariat, Belgrade, Serbia,
[2] University of Belgrade, Faculty of Organizational Sciences, Belgrade, Serbia

Abstract. This paper analyses the existing digital ecosystem of Serbian courts through the Berkana Two Loops model and the Theory of Change. Based on the level of development of ICT systems currently in production, the paper identifies the existing paradigm, or loop, of judicial digital transformation and examines its weaknesses. In outlining the construction of a new paradigm, the paper starts from the existing state of affairs and identifies the roles within the Berkana model, as well as the activities associated with them, that may provide potential for building a new AI-augmented judicial ecosystem. The Theory of Change is used to operationalize this transition. It serves to map technologies as possible solutions to the challenges of ICT systems and procedures and to connect these solutions with expected changes in judicial work.

Keywords: eJustice, Digital transformation, AI, Judicial reform, Berkana Two Loops.

1 Introduction

AI is already present in the judiciary of Europe and beyond, primarily in concrete court-related tasks such as case law search, anonymization, transcription, translation, workflow support, and information retrieval. The prevailing model is assistive rather than substitutive. AI supports courts in data-intensive and repetitive tasks, while final legal judgment remains with human actors [1, 2]. The region is pursuing these trends as well, through Serbian pilot projects, and case-law search in Montenegro.

The Serbian judiciary has already passed through an important phase of digital transformation. Over the last decade, courts and other judicial actors have gradually moved from a workflow dominated by paper carriers toward the use of electronic data, case management software, web applications, and electronic exchanges with state registers [3]. That process, as in other countries, did not merely introduce new tools. It also required procedural adaptation, legal interpretation, and the gradual translation of judicial routines toward programmable and interoperable operations [4, 5].

At the same time, the growing discussion on artificial intelligence in justice should not be treated as if it were starting from zero. In comparative practice, AI is already used for document search, anonymization and pseudonymization, triaging, workflow automation, transcription, translation, and different forms of decision support [1, 2]. In Serbia as well, pilot experiences [6] and analytical work indicate that the most realistic direction is not automated judging, but assistive use of AI in those parts of judicial work where repetitive review of large data volumes burdens judges and court staff. For this paper, the term "AI-augmented judiciary" refers to a mode of judicial work in which artificial intelligence does not replace human adjudication, but

operates as an assistive layer that supports activities such as search, classification, anonymization, transcription, information extraction, and workflow management. At the same time, final legal judgment and responsibility remain with human actors.

This paper argues that the relationship between these two developments can be better understood through the Berkana Two Loops model [7, 8]. Developed by Margaret Wheatley and Deborah Frieze within the Berkana Institute, the model describes system change as a period in which an old dominant system continues to function while a new paradigm emerges from connected experiments and gradually gains coherence and influence [8]. The Berkana Two Loops model is adopted because it moves beyond linear transformation to address the judiciary's systemic rigidity. By emphasizing 'hospicing' [9], the model provides a framework for the ethical retirement of legacy digital systems while fostering an AI-augmented 'second loop' emerging from existing data foundations. This approach is consistent with modern strategic methodologies used in high-stakes, hierarchical environments such as EU governance and military digital transitions [10, 11]. Applied to the Serbian judiciary, the model helps explain why AI does not arrive after digitalization is completed, but rather grows within and alongside the existing digital ecosystem. The central thesis of this paper is therefore that AI in the Serbian judiciary should be framed as a second loop emerging from the first loop of digital transformation, rather than as a separate or disruptive trajectory detached from prior institutional development.

While the Berkana Two Loops model provides the structural framework for understanding the systemic transition from digital to intelligent systems, the Theory of Change (ToC) methodology [12] serves as the operational roadmap. The ToC explicitly maps the causal pathway from the 'passive' digital inputs generated by the first loop to the 'active' AI-augmented interventions of the second loop. This approach ensures that AI implementation is not viewed as a mere technological upgrade, but as a strategic intervention designed to produce specific judicial outcomes, such as enhanced predictability and procedural efficiency [13].

Methodologically, the paper is conceptual and analytical rather than empirical. It is based on a secondary analysis of prior studies, pilot projects, and relevant normative and policy sources, which are synthesized to interpret the Serbian judiciary’s transition from digitalization to AI-augmented justice through the Berkana Two Loops framework. The purpose of the paper is therefore not to present new empirical results, but to organize existing evidence and policy-relevant developments into an analytical transition model supported by the Theory of Change framework, understood here as an established evaluation and planning methodology.

The paper is structured in seven parts. The first part presents the main features of digital transformation in the Serbian judiciary as the dominant institutional loop. The second part examines the limitations of that loop, including interoperability issues, data use, legal adaptation, operational rigidity, and informal digital workarounds. The third part considers the emerging second loop as an assistive AI-augmented phase of judicial work, provided that the necessary institutional, legal, and technical preconditions are in place. The fourth part applies the Berkana Two Loops model to explain why the key issue is not immediate AI adoption, but the preparation of a reliable digital and governance foundation for future AI-assisted justice. The fifth part develops a Theory of Change (ToC) framework by identifying key inputs, activities, outputs, outcomes, and expected impacts of the transition from digital judiciary to AI-

augmented justice. The paper closes by arguing that responsible AI augmentation depends on data quality, interoperability, legal safeguards, information security, and meaningful human control. Finally, the paper discusses key policy implications.

2 The First Loop: Digital Transformation as Judicial Stabilization

2.1 Digitalization as Infrastructure Building

Earlier work on the digital transformation of the Serbian judiciary defines this process as a transition from traditional, paper-based communication toward digitally governed procedures performed through electronic platforms [14] and supported by valid regulations. In that sense, digital transformation is hardly exhausted by the mere use of computers. It requires a legally recognized and operationally meaningful segment of procedure to be performed electronically, with visible effects on efficiency, transparency, accessibility, and legal certainty [15, 16].

In the Berkana Two Loops model, systemic transition is understood through several key roles and processes. The declining dominant system is not simply abandoned; it is "hospiced," meaning that its still-useful functions are preserved while obsolete, rigid, or harmful elements are gradually retired. At the same time, "pioneers" are actors, projects, or practices that experiment with new ways of working before these become mainstream. As such experiments connect and become more coherent, "bridges" are needed to transfer value from the old system to the emerging one, including knowledge, data, infrastructure, institutional legitimacy, and legal safeguards. "Stewards" are the institutions or communities that hold responsibility for guiding this transition, protecting core public values, and ensuring that the emerging system develops in a responsible and sustainable direction. Applied to the judiciary, these concepts make it possible to distinguish between maintaining essential digital infrastructures, retiring outdated procedural and technical arrangements, supporting early AI-assisted practices, and governing the transition toward a more mature AI-augmented justice system. The key terminology used here is derived primarily from Wheatley and Frieze's Berkana work on emergence, especially the movement from networks to communities of practice and systems of influence [7, 8]. The related language of "hospicing" the old system is used in later applications and interpretations of the Two Loops model to describe the careful retirement of declining institutional arrangements while preserving what remains valuable [9, 17].

The first loop of Serbian judicial modernization can therefore be described as a phase of stabilization through digital infrastructure. A particularly important element is the Judicial Information System (PIS/JIS), which serves as a portal for data exchange among courts, prosecution offices, public notaries, public enforcement agents, and other state registers. Prior analyses indicate that, between 2017 and 2023, more than 25 million electronic inquiries were performed through this system, replacing over 50 million paper correspondence actions. The practical effect was not only savings in material costs, but also real-time access to official data that could shorten certain procedures by three to six months [3].

The same first loop includes case management software, electronic public services, online filing-related functionalities in selected areas, digital communication channels, electronic court fees (ePayment), public e-auctions, and portals that expose replicated or operational data to authorized users. The digital ecosystem of courts in Serbia has

also been described through its major components: PIS (JIS), decentralized case management solutions such as AVP, SAPS, and SIPRES, and reporting systems that together support court administration and adjudicative workflows [3]. While not yet fully unified, this ecosystem already produces a large amount of electronic records, process data, and metadata.

Table 1. Main components of the current Serbian judicial digital ecosystem

Component	Main function	Role in the current ecosystem
Judicial Information System (PIS/JIS)	Electronic exchange of data between courts, prosecution offices, public notaries, public enforcement agents, and other state registers.	Core interoperability layer for access to official data and reduced reliance on paper correspondence.
Case-management systems (AVP, SAPS, SIPRES)	Registration and management of cases, procedural actions, documents, deadlines, hearings, and case metadata.	Operational backbone of everyday court work, although decentralized and heterogeneous across different courts.
Reporting and statistical systems	Collection and processing of data on performance, caseloads, backlogs, and duration of proceedings.	Support managerial and analytical functions, but reporting can still depend on manual aggregation and fragmented data.
eCourt/eSud services	Electronic filing and communication in selected proceedings where procedural rules allow electronic interaction.	Enable parties and professional users to interact with courts remotely in specific procedural contexts.
Electronic bulletin board / eBoard / eTabla	Electronic publication of court notices, procedural communications, and service-related information.	Replaces or supplements physical bulletin boards and supports more efficient notification and publication workflows.
eAuction	Electronic public sale platform in enforcement-related proceedings.	Mature service-oriented digital procedure with practical effects on transparency and procedural efficiency.
ePayment of court fees	Electronic payment of court fees, with integration into court-related digital services.	Reduces administrative burden and supports more efficient handling of fee-related procedural steps.
Case-law database and court portals	Publication, search, and access to judicial decisions and court-related information.	Supports transparency, legal information retrieval, and future advanced search and AI-assisted tools.

Digital powers of attorney and other specialized services	Electronic authorization and specific digital services connected to judicial or justice-sector procedures.	Extends the ecosystem beyond internal court administration toward user-facing digital interaction.

Table 1 is not intended to provide an exhaustive technical inventory, but to give a simplified overview of the main functional layers of the current ecosystem and their relevance for the transition toward AI-augmented justice.

Taken together, these tools show that the Serbian judicial ecosystem has moved beyond the mere digitization of records toward operational digital services that reshape specific procedural interactions.

2.2 The Structural Limits of the First Loop

Yet the first loop remains incomplete. The Serbian judicial ecosystem still reflects inherited decentralization, heterogeneous databases, and uneven technical maturity. The digital ecosystem study stresses that interoperability is not yet fully achieved at either the technical or semantic level and that centralization is yet a strategic direction for improving uniformity, sustainability, and data quality [18, 19]. Earlier analyses likewise emphasized that, in the absence of a centralized information system of the type used in more digitally advanced jurisdictions [20, 21], progress in Serbia often depended on individual information systems digitizing particular procedures rather than the judiciary as a single integrated environment.

The limits of the first loop are not only technical, but also legal and organizational. Recent Serbian work on territorial jurisdiction shows that centralized digital services can still be constrained by rules designed for territorially bounded paper procedures, creating avoidable routing, transfers, and delays even where the technological platform itself is already centralized [22, 23]. More broadly, the regulatory framework still does not fully use the possibilities created by digital technology: there is no clearly designated authority or procedure for deciding on access to court-held datasets, such as data on misdemeanor judgments, and certain services are still tied to the traditional court-counter model, since the centralized online issuance of certificates issued by commercial courts has not been fully provided independently of the physical court registry.

A further reason the first loop begins to decline can be described, in Berkana terms, as systemic failure or underperformance. The dominant digital infrastructure continues to function, yet it increasingly fails to satisfy operational needs at the point of use. When judges, clerks, and other users prepare reports manually outside the system, they implicitly confirm that the existing digital layer has become rigid in certain workflows. Rather than absorbing complexity, it generates additional administrative effort. In that situation, the institutional system remains formally present, but part of the real work migrates into workarounds and parallel channels.

This phenomenon is analytically important because it explains why informal digital practices emerge. The turn toward messaging applications, photographed documents, external LLM tools, and other ad hoc solutions should not be reduced to individual preference. Rather, these practices are weak signals that the first loop no longer provides sufficient speed, flexibility, and assistive capacity for parts of the workload.

This phenomenon is usually referred to as “Shadow IT” in information systems literature [24]. Without overlooking the significant information-security risks created by such practices, this analysis treats them primarily as weak points of the first loop: they reveal where the official digital infrastructure no longer provides sufficient usability, speed, integration, or assistive capacity for everyday judicial work. In Berkana terms, this is precisely the moment when the old curve begins to fall more sharply: the dominant system is no longer experienced only as support, but also as friction.

A key limitation of the existing digital ecosystem is that current judicial information systems are still not used for the automatic generation of regular annual reports on the work of courts and judges, nor as a foundation for public portals that would systematically inform the public about court performance statistics. This suggests that the existing systems, although valuable for case handling and selected operational functions, have not yet matured into a fully integrated data environment capable of supporting higher-order analytical, managerial, and transparency functions. As a result, reporting often remains dependent on separate administrative effort, manual aggregation, or parallel workflows, which reduces efficiency, weakens data reuse, and limits the larger public value that could be created from judicial data already generated within the system.

The first loop also appears inadequate relative to broader stakeholder expectations. Strategic responses to backlog reduction have mainly focused on electronic case management, enforcement digitalization, electronic delivery, statistical monitoring, and workflow standardization [25]. However, they have not yet fully incorporated a broader data-driven vision of judicial time management and caseflow optimization reflected in CEPEJ/SATURN instruments, such as systematic duration monitoring, public statistical information, procedural calendars, and ICT-supported planning [26, 27]. This shows that the limits of the dominant loop lie not only in technology availability, but also in the narrow framing of technology as support for existing workflows rather than as a basis for advanced managerial and analytical functions.

The report is particularly relevant because it shows that the first digital loop has not failed, but remains structurally constrained. These constraints include outdated equipment, the absence of standardized IT governance, weak information security, insufficiently regulated cooperation with vendors, unclear arrangements concerning the processing of personal data, and inadequate practical support for electronic documents [25].

These limitations are key for understanding the next phase. Artificial intelligence does not enter an institutional vacuum. It depends on whether digitalization has already produced reliable datasets, machine-readable metadata, and repeatable process patterns. For that reason, the first loop should be interpreted as the necessary substrate for any later AI-assisted development. Without digital records, structured fields, interoperable exchanges, and standardized procedural logic, AI remains either technically weak or normatively unsafe.

3 The Second Loop: Emergent AI Assistance in the Judiciary

3.1 AI as Augmentation Rather than Replacement

Recent work on the potential of AI in Serbia's judiciary points to a cautious but clear direction. The most promising uses of AI are not those that replace human judicial reasoning, but those that assist humans in handling volume, repetition, and complexity [26, 27]. Comparative practice reviewed in that paper includes document search and review, online dispute resolution, prediction-related tools, anonymization and pseudonymization, triaging and workflow automation, recording and transcription, translation, and information services [28–30].

The emerging second loop in the judiciary should not be understood as a vague turn toward "AI in law," but as the gradual consolidation of a new set of assistive capabilities within legal work. Recent scholarship in legal NLP helps make this more concrete by showing that the new paradigm is taking shape through tools for question answering, summarization, classification, entity recognition, argument mining, and other forms of structured support over legal texts and workflows [31, 32]. From that perspective, the second loop is not defined by the replacement of judges or legal professionals, but by the rise of increasingly specialized systems that can reduce repetitive cognitive labor, improve access to relevant materials, and support legally informed human work.

The most illustrative example is the OLGA [27, 33] system used in Germany for mass claims related to diesel emissions litigation. OLGA groups similar appeals, identifies relevant tags and facts within extensive pleadings, and supports judges in managing large sets of similar cases[1]. The key insight is that the tool is assistive. It does not decide the dispute, but it accelerates the identification of relevant similarities and improves consistency by reducing repetitive manual review. This makes OLGA especially useful as a conceptual proof of what the second loop looks like in judicial practice: not an automated court, but an AI-supported judicial workflow [33].

In Serbia, the same paper [27] notes several adjacent or preparatory possibilities: enhancement of case-law databases[2] through more sophisticated search, automatic anonymization of court decisions, speech-to-text transcription, translation support, and assistive tools for extracting key information from judicial documents. Earlier Serbian practice in template-based decision drafting for mass disputes also suggests that courts already work with forms of structured metadata retrieval from case files. AI extends this logic by enabling more flexible extraction from unstructured text [31]. The traffic violations machine learning pilot in Serbia provides an important empirical indication that AI can be meaningfully applied in judicial work where sufficiently comprehensive, accurate, and structured data already exist, while also confirming that such deployment must remain attentive to bias, interpretability, and human review [6].

[1] For further details, see: https://www.ibm.com/case-studies/blog/judicial-systems-are-turning-to-ai-to-help-manage-its-vast-quantities-of-data-and-expedite-case-resolution

[2] For further details, see: https://sudskapraksa.sud.rs/

3.2 Why the Second Loop Emerges from the First

The relationship between the two loops is therefore developmental rather than antagonistic. AI relies on what digitalization has already stabilized: electronic records, metadata, searchable texts, access to registers, and repeatable procedural patterns. The second loop does not begin only after the first loop has ended; it begins inside those parts of the first loop that have become sufficiently structured, reliable, and institutionally embedded to support more advanced forms of assistance.

For that reason, the rise of AI in the judiciary should be described as a layered transition: first, digitize; then, structure; then, interconnect; and only then, responsibly augment with AI.

Table 2. Berkana Two Loops model mapped to e-Justice and AI transition

Berkana element	Judicial application	Interpretation
Dominant system (Loop 1)	AVP, SAPS, SIPRES, PIS, PRONEP, eAuction, Case-law database, web	Systems focused on digitizing paper flows and official registers.
Bifurcation point (decline)	Administrative bottlenecks and territorial constraints, regulatory gaps in data governance, insufficient standardization, and the underutilization of data for automated report generation	The moment when the digital system becomes rigid and no longer fully resolves efficiency problems.
Pioneers (Loop 2)	Machine-learning pilot projects in misdemeanor cases; AI-enhanced case law database; RPA in eBoard and data cleansing process.	Innovations that introduce intelligence and assistance rather than mere electronic recording. Shift toward an assistive architecture that widens practical **access to justice** while preserving **Rule of Law** safeguards like fairness and accountability
Bridges	EUDI Wallet, RPA, Centralized CMS, Metadata standards, Open formats.	Technologies and institutional mechanisms that enable the transition from digital to intelligent justice. Paradigm shift from the mere recording of data toward real-time intelligent assistance
Hospicing (the declining system)	Ethical retirement of legacy digital systems, paper-based physical archives, and rigid territorial constraints.	Preserving essential functions while decommissioning obsolete or rigid procedural arrangements to free up institutional capacity for AI-augmented workflows.
Stewards	Ministry of Justice, Supreme Court, Sector Council for Information and Communication Technologies	Institutions and communities responsible for managing the retirement of obsolete processes and the growth of new ones.

4 Applying the Berkana Two Loops Model to the Serbian Judiciary

The Berkana Two Loops model offers a useful lens for organizing these findings. In the model, the dominant system does not instantly disappear when a new one emerges. The old paradigm continues to carry institutional weight even while its limits become more visible. At the same time, pioneers begin new practices on the

margins, connect them, and gradually form a stronger new normal [7, 8]. This description maps well onto the contemporary Serbian judiciary.

The first loop in Serbia is the digital judiciary as it currently exists: electronic case records, data exchange through PIS (JIS), specialized web systems, replicated databases, partial online services, and a still unfinished effort toward centralization and interoperability [20, 21]. This loop remains dominant because it still carries the everyday operation of courts. It also continues to be indispensable because legal certainty, procedural validity, and institutional memory are anchored in it [18, 19].

The second loop is the emerging AI-augmented judiciary. It appears first through modest but high-value use cases: clustering of similar claims, smarter search of case law, anonymization of judicial decisions, extraction of facts from large submissions, transcription of hearings [34], translation support, and workflow assistance[31, 33]. These are the pioneering practices of the new paradigm. They do not yet constitute a complete judicial system. Still, they point toward one in which human attention is reallocated from repetitive filtering and retrieval toward reasoning, evaluation of evidence, and legal justification.

The bridge between the two loops [12] can be understood through four interdependent conditions. The first is data quality: AI support is only as reliable as the completeness, accuracy, and consistency of the underlying judicial data. The second is interoperability: assistive tools become institutionally useful only when they can connect relevant systems and datasets without reproducing manual work in digital form. The third is legal and institutional safeguarding: AI-supported processes must comply with rules on data protection, information security [35], procedural fairness [36], and accountability [37]. The fourth is human oversight [1, 2]: the transition remains legitimate only if AI supports classification, retrieval, extraction, anonymization, or workflow management, while final legal judgment remains with human actors.

A recent Serbian pilot in misdemeanor traffic cases strengthens this argument from an empirical perspective. There, machine learning was trained on approximately 700,000 records from the SIPRES case-management environment, precisely because that domain offered a sufficiently large, historically relevant, and metadata-rich dataset [6]. The project demonstrated that where data quality and legal structuring are adequate, assistive AI can be meaningfully deployed to support greater consistency and faster handling of routine judicial work. At the same time, the pilot also exposed the importance of bias management[38], interpretability, secure pseudonymization, user training, and technical stability.

Figure 1 visualizes this transition in Berkana-inspired terms. The old loop is represented by the dominant digital judiciary built around AVP, eSud, and PIS. Beneath it, emergent informal digital practices ("shadow digitalization") are treated analytically not as desirable end-states, but as weak signals of unmet operational needs for speed, connectivity, and assistive intelligence. The bridging layer consists of exactly those institutional measures discussed above: cleansing dirty data, extracting metadata from legacy records, mapping decentralized applications to a more unified model, enabling interoperability, and supporting a phased migration to secure official applications. The new loop is therefore not automated adjudication, but an AI-augmented judiciary in which assistive functions operate on governed data and remain subject to human oversight.

The relevance of this transition extends beyond institutional efficiency. It is also connected to access to justice [15, 35]. Where digital and AI-enabled tools reduce notification failures, accelerate routine steps, improve retrieval of legal information, and lower procedural friction for users, they can make justice systems more usable and more accessible. In that sense, the second loop should be understood not as automated judging, but as an assistive architecture that may widen practical access to justice while preserving due process safeguards.

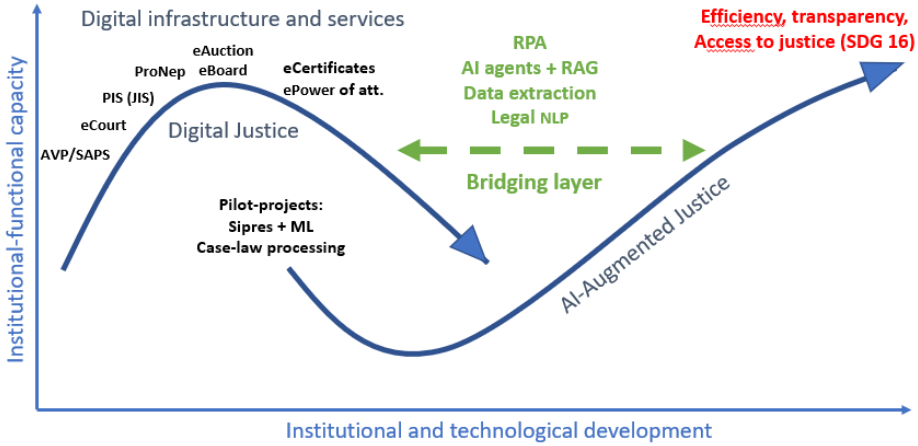


Fig. 1. Transition from the first loop of digital justice toward AI-augmented justice.

In the Serbian context, AI may also have a specific transitional function: supporting the move away from fragmented legacy systems. Natural language processing and machine learning can assist with identifying duplicate records, inconsistent naming patterns, missing values, incompatible field formats, and misclassified procedural data. They can also help extract structured metadata from unstructured judicial texts, including case type, procedural phase, dates, legal basis, and links between related cases. In this way, AI can contribute to making legacy data cleaner, more searchable, and more suitable for migration into a more unified judicial information architecture [31, 32, 39].

This function, however, must remain tightly governed. In the judiciary, probable matches and automated suggestions cannot be treated as authoritative where they affect party identity, case status, legal classification, or procedural history. AI-assisted cleansing and migration should therefore be auditable, rule-based where possible, and subject to human validation [26, 40].

Seen in this way, the Berkana model helps avoid two analytical mistakes. The first is technological rupture: the false assumption that AI replaces digitalization. The second is technological minimalism: the assumption that AI is merely another minor software feature. In reality, AI marks a new loop because it changes the function of digital infrastructures. Data no longer serve only storage, transmission, and retrieval; they become the basis for classification, recommendation, summarization, and

machine-assisted perception[31, 33]. Yet this new loop is still dependent on the institutional ecology created by the older one [7, 8, 12].

This perspective also clarifies why informal digital practices should be treated analytically, but not normalized institutionally. Serbian case study on the digital power of attorney[22] and comparative practice [24] demonstrates that when secure, legally recognized digital channels are absent, users may resort to ad hoc assistance, shared credentials, or other extra-systemic practices. These practices reveal demand for usability and remote participation, but in judicial settings they must be translated into auditable, interoperable, and legally structured forms of authorization and communication [26, 40].

The same logic applies to the growing use of commercial AI tools, chatbots, and privately developed case-law search solutions outside official judicial infrastructure. Such practices may offer immediate gains in speed, usability, and retrieval, especially where official systems remain functionally limited or insufficiently advanced. However, they also create a dual structural problem: first, the data, queries, and interaction patterns generated through such use do not remain within the judicial information ecosystem and therefore do not contribute to institutional learning, metadata enrichment, or future system improvement; second, they increase the risk that sensitive legal information, documents, or reasoning processes may move beyond the protected institutional environment. In that sense, these practices should be understood as signals of unmet functional demand, but not as a sustainable model for judicial digital transformation.

The second loop should not be constructed as an abstract programme of "introducing AI into the judiciary," but by identifying the networks through which new practices are already emerging, mapping their pioneers and operational needs, and translating their micro-level innovations into secure and institutionally governed solutions. In this sense, Berkana's notion of networks is methodologically useful because it directs attention to actual sites of transition rather than to formal system design alone. In our view, once such micro-tasks or atomic activities are identified, they should be systematically mapped to the classes of technologies most appropriate to their functional and legal characteristics[41, 42]; this approach is relevant not only for the application of AI, but also as a broader framework for identifying and introducing other technologies in the judiciary.

A comparable direction has also been publicly announced in Montenegro, where judicial authorities have presented a forthcoming AI-based system for case-law search, including intelligent retrieval and a chatbot operating on decisions stored in the judicial information system and judicial bulletins. For this article, the importance of that announcement lies less in evaluating the system ex ante and more in showing that the transition toward assistive, retrieval-based judicial AI is being articulated regionally, not only in Serbia.

5 Theory of Change: Components and Causal Logic

To understand the systemic transition described by the Berkana model, it is essential to define the operational logic behind the intervention. The Theory of Change (ToC) provides a causal framework that links technological activities to judicial outcomes through five fundamental components [12]. Table 2 summarizes the intervention logic

linking first-loop digital resources to expected judicial outcomes and long-term systemic impact.

Inputs represent the digital capital inherited from the first loop, including case-management data, eAuction records, legal databases, and digitized court materials. These resources provide the structured basis on which AI-assisted functions can operate.

Activities are the controlled second-loop interventions that transform passive data into active assistance, including machine-learning models, RAG-based research agents, and digital twins for process simulation [30].

Outputs are the direct technological products of these interventions, such as caseflow prediction prototypes, automated summaries, classification models, and advanced search tools.

Outcomes concern changes in judicial work, including reduced bottlenecks, improved consistency, faster access to precedents, and better use of metadata and case information.

Impact refers to the long-term goal of a faster, more predictable, and more accessible AI-augmented judiciary that strengthens legal certainty, rule-of-law performance, and access to justice [13, 15, 33].

Table 3. Theory of Change components mapped to AI-augmented justice

Component	Definition	Judicial application
Inputs	Foundation resources inherited from the first loop.	Case-management registries (AVP/SAPS), eAuction records, and legal databases.
Activities	AI-enabled interventions and controlled experiments.	RAG-based AI agents, Machine-learning support based on historical data, Judicial digital twins.
Outputs	Direct technological products of the interventions.	Predictive simulation reports, automated legal summaries, advanced search tool, data mining, data cleaning.
Outcomes	Systemic shifts and bridges toward the new paradigm.	Greater consistency in reasoning and reduced procedural bottlenecks, including metadata standardization.
Impact	Long-term goals and systemic maturity.	Enhanced legal certainty and strengthened rule-of-law performance, improved efficiency and access to justice with human-centered legal aid channels, in line with SDG 16 (Access to Justice) and the Rule of Law.

6 Safeguards for the Transition

The assistive model does not eliminate the risks of AI in justice. On the contrary, even non-decisional systems can affect legal work by shaping what information is retrieved, how documents are classified, which similarities are detected, and which procedural priorities become visible to users [28, 33, 44]. For that reason, safeguards must address not only fully automated decision-making, but also the more subtle effects of AI on judicial attention, reasoning, and institutional practice.

For Serbia, an important normative boundary is already visible. The analysis of AI in the Serbian judiciary points to the Law on Personal Data Protection as a general rule prohibiting decisions based solely on automated processing when such decisions produce legal effects or significantly affect an individual's status. The same paper also notes that AI systems used in justice under the logic of the EU AI Act fall into higher-risk categories and therefore require stronger safeguards related to robustness, governance, transparency, and human oversight [1, 40].

The transition toward AI-augmented justice should also be linked to the broader public-policy framework for impact assessment. Serbian planning-system legislation already requires ex-ante and ex-post assessment of the effects of public policies and regulations, including their social effects[45]. In the field of artificial intelligence, this logic is further reflected in the Ethical Guidelines for the Development, Implementation and Use of Trustworthy and Responsible Artificial Intelligence [26], adopted by the Government of Serbia in 2023. Although these guidelines are not legally binding, they provide an important soft-law framework for assessing AI systems through requirements such as human agency and oversight, technical robustness and safety, privacy and data governance, transparency, diversity, non-discrimination and equality, societal and environmental well-being, and accountability. For the judiciary, this means that the expected impact of AI should not be assessed only in terms of efficiency gains, but also through its effects on fairness, access to justice, data protection, non-discrimination, procedural safeguards, and human responsibility.

Data governance serves as a fundamental safeguard by regulating the judicial data life cycle from creation to long-term electronic archiving, and by ensuring that the second loop (AI) operates on legally compliant, high-quality inputs. In the Serbian context, this framework is anchored in the Law on Electronic Government and the Law on Electronic Document, which mandate strict metadata standards and secure interoperability. By functioning as a "quality gate," Data Governance mitigates the risks of "Garbage In, Garbage Out" and algorithmic bias. Consequently, it acts as a vital bridge between the two loops, ensuring that AI-augmented justice remains grounded in legal certainty and procedural integrity.

To operationalize the discussed safeguards, the transition must be guided by a structured risk-management approach. Table 4 synthesizes the primary risks identified in the Serbian judicial context ranging from algorithmic bias to the structural limitations of legacy data and maps them to specific mitigation strategies and their corresponding legal or ethical foundations. This matrix ensures that the assistive AI layer remains grounded in procedural integrity and international standards for high-risk AI systems.

Accordingly, the most defensible path for the Serbian judiciary is an assistive model of AI deployment. In that model, algorithms help classify, retrieve, extract, anonymize, or transcribe, but they do not render binding legal judgment. Such a model is also consistent with the conclusions of broader scholarship on AI and judicial professions, which argues that AI tools should remain aids rather than replacements for judicial reasoning, and that final authority and responsibility must remain with human judges and legal professionals [1, 2].

Table 4. Risk mitigation safeguards for AI-augmented justice

Risk	Mitigation (safeguard)	Legal/ethical basis
Algorithmic bias	Mandatory human-in-the-loop validation and regular bias audits.	Ethical Guidelines for Trustworthy and Responsible AI (Government of Serbia, 2023).
"Garbage In, Garbage Out"	Robust data governance and metadata cleansing of legacy records.	Law on Electronic Document and Law on Electronic Government.
Loss of transparency	Implementation of auditable design and AI explainability (interpretability).	EU AI Act (classification of judicial AI as high-risk).

7 Policy Implications

The Berkana Two Loops model suggests that the transition toward AI-augmented justice in Serbia should be operationalized gradually and institutionally, rather than through isolated technological experiments [7, 8]. Table 5 summarizes the main policy implications arising from the analysis.

Table 5. Policy implications for AI-augmented justice in Serbia

Policy area	Recommended action	Expected contribution
AI readiness	Conduct an AI-readiness assessment of existing judicial information systems, focusing on data quality, interoperability, metadata standards, information security, and legally suitable use cases [26, 45].	Identifies where AI can be responsibly introduced and where basic digital conditions must first be strengthened.
Priority use cases	Give priority to assistive applications such as anonymization, case-law search, transcription, classification, metadata extraction, backlog analysis, and workflow support [33, 44, 46].	Focuses AI deployment on low-risk, high-value functions that support rather than replace judicial work.
Stakeholder involvement	Include judges, court staff, prosecutors, lawyers, IT professionals, data-protection officers, academic experts, and civil society in identifying needs, testing tools, and defining safeguards.	Improves usability, reduces institutional resistance, and supports alignment with procedural rights and public trust.
Risk mitigation	Address fragmented legacy systems, uneven data quality, limited AI literacy, unclear accountability, and extra-systemic tools through controlled pilots, training, impact assessments, auditable design, secure official applications, and mandatory human review.	Reduces operational, legal, ethical, and security risks during the transition.

Evaluation indicators	Evaluate the transition through indicators such as data completeness, interoperability, reduction of manual work, processing time, consistency in repetitive cases, availability of anonymized case law, user satisfaction, public accessibility, validated AI outputs [47], and compliance with data-protection [35] and ethical standards [26].	Ensures that AI-augmented justice is assessed not only by speed, but also by fairness, accountability [16], and access to justice [15].

8 Conclusion

The Serbian judiciary has passed through the expected stages of introducing ICT tools into judicial work, including decentralized case-management applications, case-law databases, online portals, and applications supporting specific judicial and justice-sector procedures. However, a key missing element remains a centralized case-management system capable of serving as a reliable source of data for performance reporting, institutional monitoring, data exchange, and public information. AI tools may contribute to the improvement of these existing systems, especially by addressing some of their accumulated weaknesses. Their potential role is not limited to the introduction of new applications, but also includes data cleansing, migration support, better structuring of records, improved use of existing data, and assistance in preparing the institutional environment for a more integrated digital ecosystem. In this sense, AI should first be understood as a tool for strengthening the conditions under which digital justice can function more effectively.

The requirements of security, transparency, accountability, and human oversight are particularly important in this field, given the direct impact of judicial work on individual rights. For that reason, the use of AI in the judiciary cannot be treated in the same way as ordinary administrative automation. Higher standards are required, both at the level of system design and at the level of data governance.

The Berkana Two Loops model is useful in this context because it helps identify, at an early stage, those emerging practices and networks that may need to be institutionalized and digitalized through formal channels. It also points to the importance of supporting and maintaining legacy systems during the transition. Such "hospicing" of existing systems is not a passive preservation of outdated infrastructure, but a necessary condition for a smooth transition, continuity of work, and reliable access to previously accumulated data.

Further research should apply instruments derived from the Berkana Two Loops model and the Theory of Change to concrete judicial and administrative workflows, especially by breaking them down into atomic tasks and mapping them to appropriate AI and ICT tools.

Acknowledgments. The authors acknowledge the use of AI tools solely for copyediting purposes, including translation support, grammar and spelling checks, sentence-level stylistic refinement, and improvement of readability. AI tools were not used to generate the research idea, legal or technical analysis, methodological approach, findings, conclusions, or any substantive academic content of the manuscript. All substantive claims, legal interpretations,

references, examples, tables, and conclusions were independently checked and validated by the authors. The authors have reviewed, edited, and approved the final version of the manuscript and take full responsibility for its content.

References

1. Advisory Board On Artificial Intelligence: Reflections of the AIAB on the use of Artificial Intelligence in Judicial Systems. 11 (2025).
2. Advisory Board On Artificial Intelligence: 1st AIAB Report on the use of Artificial Intelligence (AI) in the Judiciary based on the information contained in the Resource Centre on Cyberjustice and AI. , Strasbourg (2025).
3. Đilas, A., Kuzmanović, D.: Digital Transformation in the Judiciary. E-Business Technologies Conference Proceedings. 3(1) (2023).
4. Brooks, C., Gherhes, C., Vorley, T.: Artificial intelligence in the legal sector: Pressures and challenges of transformation. Cambridge Journal of Regions, Economy and Society. 13, 135–152 (2020). https://doi.org/10.1093/cjres/rsz026.
5. Ministerio De Justicia Gobierno De Espana / Ministry of Justice of Kingdom of Spain: Digital Transformation Projects in the Justice Sector. (2023).
6. Kuzmanović, D., Đilas, A., Jevtić, I.: Machine Learning and Judicial Rulings Traffic Violations. E-Business Technologies Conference Proceedings. 4(1) (2025). https://www.ebt.rs/journals/index.php/conf-proc/article/view/235
7. Wheatley, M.J., Frieze, D.: Walk Out Walk On: A Learning Journey into Communities Daring to Live the Future Now. Berrett-Koehler Publishers, San Francisco (2011).
8. Wheatley, M., Frieze, D.: Using Emergence to Take Social Innovation to Scale. The Berkana Institute (2006). https://www.margaretwheatley.com/articles/using-emergence.pdf
9. Robinson, C.: Hospicing The Old, https://medium.com/thefarewellfund/hospicing-the-old-16e537396c4b, (2019).
10. Borchardt, S., Trane, M., Choucair, N., Fronza, V., Guerrieri, V., Musso, V., Vittuari, M., Marelli, L.: Transforming food systems – An exploratory research journey on the EU green transition. Vol. 1. Publications Office of the European Union (2026). https://doi.org/10.2760/9027136.
11. Harris, J.: Dinosaurs Can't Surf: Is the Australian Defence Organisation Change Resistant?, https://theforge.defence.gov.au/article/dinosaurs-cant-surf-australian-defence-organisation-change-resistant, (2026).
12. Weiss, C.H.: Nothing as Practical as Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families. In: Connell, J.P., Kubisch, A.C., Schorr, L.B., and Weiss, C.H. (eds.) New Approaches to Evaluating Community Initiatives: Concepts, Methods, and Contexts. Aspen Institute, Washington, DC (1995).
13. Amjad, B., Fatima, A., Mehmood, F., Amjad, A.: Towards Automation and Digital Justice: A Review of Digital Twin Technology for Automation in Justice System. Journal of Asian Development Studies. 14, 49–62 (2025). https://doi.org/10.62345/jads.2025.14.3.4.
14. Duzlevski, I.: Značaj digitalizacije u krivičnom pravu. In: Digitalization in Penal Law and Judiciary. pp. 59–69. Institut za uporedno pravo; Institut za kriminološka i sociološka istraživanja (2022). https://doi.org/10.56461/ZR_22.DUKPP.05.
15. Nastić, M.: Access to Justice in the Digital age. Teme. 235–249 (2025). https://doi.org/10.22190/TEME241104018N.
16. Mihajlović A.: Efficiency of justice – importance of the CEPEJ as the legacy of the Council of Europe. In: Jaskiernia Jerzy and Spryszak Kamil (eds.) XVI International Human Rights Conference "75 years of influence of the Council of Europe on shaping the European legal space in the areas of democracy, rule of law and protection of human rights." pp. 344–356. Adam Marszalek, Torun (2025).

17. Deron, M.: Sprouting technology otherwise, hospicing negative commons – Rethinking technology in the transition to sustainability-oriented futures, https://www.researchgate.net/publication/394427345_Sprouting_technology_otherwise_hospicing_negative_commons_–_Rethinking_technology_in_the_transition_to_sustainability-oriented_futures, (2025). https://doi.org/10.48550/arXiv.2508.05860.
18. European Commission for the Efficiency of Justice (CEPEJ): Evaluation of the Judicial Systems (2020–2022): Serbia, https://rm.coe.int/serbia-2020-en/1680a85c97, (2022).
19. Council of Europe. European Commission for the Efficiency of Justice (CEPEJ): Evaluation of the Judicial Systems 2024 (Data 2022): Serbia, https://rm.coe.int/serbia-2024-2022-/1680b1f6eb, (2024).
20. European Commission: Serbia 2024 Report, https://enlargement.ec.europa.eu/document/download/3c8c2d7f-bff7-44eb-b868-414730cc5902_en?filename=Serbia%20Report%202024.pdf, (2024).
21. European Commission, D.-G. for E., Neighbourhood, E.: Serbia Report 2025. , Brussels (2025).
22. Bosković D., Đilas A., Kuzmanović D.: Digitalizacija punomocja. In: YU INFO 32. NACIONALNA KONFERENCIJA u oblasti informaciono-komunikacionih tehnologija (2026).
23. Kuzmanović, D., Đilas, A.: Optimizing Judicial Service of Process via Centralized eBoard Publication: A Process-Oriented Case Study from Serbia. In: 25th International Symposium INFOTEH-JAHORINA, Sarajevo, Bosnia and Herzegovina, 18–20 March 2026 (2026).
24. Raković, L., Sakal, M., Matković, P., Marić, M.: Shadow IT – A Systematic Literature Review. Information Technology and Control. 49, 144–160 (2020). https://doi.org/10.5755/j01.itc.49.1.23801.
25. Državna revizorska institucija: Informacioni sistemi u pravosuđu: Izveštaj o reviziji svrsishodnosti poslovanja, https://dri.rs/storage/newaudits/2022-1-SV%20Informacioni%20sistemi%20u%20pravosudju.pdf, (2023).
26. Government of the Republic of Serbia: Ethical Guidelines for Development, Implementation and Use of Robust and Accountable Artificial Intelligence, https://www.ai.gov.rs/extfile/en/471/Ethical%20guidelines%20for%20development%20implementation%20and%20use%20of%20robust%20and%20accountable%20AI.pdf, (2023).
27. Kuzmanović, D., Đilas, A.: The Potential of AI in the Judiciary of Serbia. E-Business Technologies Conference Proceedings. 4(1) (2025). https://www.ebt.rs/journals/index.php/conf-proc/article/view/236
28. Yu, Y., Huang, S.-C., Shao, H.-L.: A Multi-Agent AI Framework for Legal Aid: Real-World Interactive Dialogue in Case Consultations. Presented at the December 4 (2025). https://doi.org/10.3233/faia251591.
29. Pour, V.: Packaging Thematic Analysis as an AI Workflow for Legal Research. Presented at the December 4 (2025). https://doi.org/10.3233/faia251628.
30. Markovich, R., Caro, L. Di, Rapp, A., Schifanella, C.: Legal Knowledge and Information Systems. (2025).
31. Farid Ariai, Joel Mackaenzie, Gianluca Demartini: Natural Language Processing for the Legal Domain: A Survey of Tasks, Datasets, Models and Challenges, https://arxiv.org/abs/2410.21306, (2025).
32. Rahman, M.M., Md, N.G., Rahman, M.M., Rahman, M.M., SD, M.K.S.: Natural language processing in legal document analysis software: A systematic review of current approaches, challenges, and opportunities. International Journal of Innovative Research and Scientific Studies. 8, 5026–5042 (2025). https://doi.org/10.53894/ijirss.v8i3.7702.
33. Getman, A.P., Yaroshenko, O.M., Shapoval, R. V., Prokopiev, R.Y., Demura, M.I.: The Impact of Artificial Intelligence on Legal Decision-Making. International Comparative Jurisprudence. 9, 155–169 (2023). https://doi.org/10.13165/j.icj.2023.12.001.

34. Marek Świerczyński, Alexandre Palanco: Selected national good practices on videoconferencing in judicial proceedings Complement to the CEPEJ Guidelines on videoconferencing in judicial proceedings. (2021).
35. Eltis, K.: The Judicial System in the Digital Age: Revisiting the Relationship between Privacy and Accessibility in the Cyber Context. McGill Law Journal. 56, 289–316 (2011). https://doi.org/10.7202/1002368ar.
36. Ambrisi, A., Aversa, R., Maurino, M., Venticinque, S.: A Methodology for Formal Modeling and Evaluation of the Judicial Process, https://www.accessogiustizia.it/., (2023). https://doi.org/10.1007/978-3-031-35734-3_32.
37. Satoh, K., Takeda, H., Goebel, R., Kano, Y., Kim, M.Y., Rabelo, J., Yoshioka, M.: Juris-Informatics: Law for AI and Law of AI. In: Proceedings - 2024 IEEE International Conference on Big Data, BigData 2024. pp. 4757–4762. Institute of Electrical and Electronics Engineers Inc. (2024). https://doi.org/10.1109/BigData62323.2024.10826058.
38. Kamiran, F., Karim, A., Verwer, S., Goudriaan, H.: Classifying Socially Sensitive Data Without Discrimination: An Analysis of a Crime Suspect Dataset. In: 2012 IEEE 12th International Conference on Data Mining Workshops. pp. 370–377. IEEE (2012). https://doi.org/10.1109/ICDMW.2012.117.
39. Kumar, A., Sangwan, S.R.: Transformative Natural Language Processing: Bridging Ambiguity in Healthcare, Legal, and Financial Applications. Springer Science+Business Media (2025). https://doi.org/10.1007/978-3-031-88988-2.
40. Office of the European Union L-, P., Luxembourg, L.: Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)Text with EEA relevance. (2024).
41. International Labour Organization: Artificial Intelligence Adoption and Its Impact on Jobs, https://www.ilo.org/sites/default/files/2025-07/G20_ILO_Artificial%20intelligence%20adoption%20and%20its%20impact%20on%20jobs.pdf, (2025).
42. Gmyrek, Pawel., Berg, Janine., Bescond, David.: Generative AI and jobs : a global analysis of potential effects on job quantity and quality. International Labour Organization (2023).
43. Новаковић, M.: A Review of the Efficiency of Justice and Other Elements of the 2022–2025 CEPEJ Action Plan: "Digitalisation for a Better Justice." (2022). https://doi.org/10.56461/ZR_22.DUKPP.15.
44. Matić-Bošković, M.: Impact of artificial intelligence on practicing judicial professions. Socioloski pregled. 58, 481–499 (2024). https://doi.org/10.5937/socpreg58-51385.
45. Mihajlović, A., Rajić Ćalić, J.: Analiza efekata primene veštačke inteligencije na društvo – strateški okvir u Republici Srbiji. In: Izazovi prilagođavanja pravnog okvira tehnološkim promenama u digitalnom društvu. pp. 29–44. Institut za uporedno pravo (2025). https://doi.org/10.56461/ZR_25.Digi.02.
46. Ilchyshyn, N.: Impact of digitalisation on the justice system: Challenges of administrative justice. Social and Legal Studios. 238–247 (2025). https://doi.org/10.32518/sals2.2025.238.
47. Gultekin, S., Rossi Reich, M., Galloni, F., Lagioia, F., Consiglio, E., Sartor, G., Bagnato, S.: A Two-Dimensional Evaluation Framework for Factual and Reasoning Assessment of LLMs in Legal Question Answering. Presented at the December 2 (2025). https://doi.org/10.3233/FAIA251579.

Original Research Article *https://doi.org/10.5281/zenodo.23475507*

Development of an AI-driven mobile service for smart cooking

Filip Stevanović[1][0009-0001-0479-1685] and Dušan Kostić[1][0009-0000-3432-6074]

[1] University of Belgrade, Faculty of Organizational Sciences, Belgrade, Serbia

Abstract. This paper analyzes the digital transformation of the culinary sector through the application of multimodal artificial intelligence. The focus of the research is the development of a mobile service "Sous Chef", which bridges the contextual gap between the visual identification of ingredients and their practical application in food preparation. The system uses advanced technologies, such as the EfficientNet-B4 and Gemini 1.5 Pro models, for automated detection of ingredients and dynamic synthesis of personalized recipes in real time. Technologically, the solution is based on the React Native and Supabase platforms, which achieves high efficiency on different operating systems with strict protection of user data. The developed prototype confirms the technical feasibility of the concept and shows how process automation can alleviate decision fatigue and encourage sustainable resource management in the household. It is concluded that this approach makes smart cooking widely accessible, eliminating the need to invest in expensive and specific IoT hardware. The main value of the work is reflected in the successful integration of multimodal intelligence into everyday kitchen processes.

Keywords: artificial intelligence, mobile service, smart cooking

1 Introduction

Global food waste is one of the most serious environmental and economic challenges of our time [1]. The fact that approximately a third of global production is wasted, with almost half of this waste generated directly in households due to inefficient meal planning, indicates an urgent need for systemic solutions [2]. While digital transformation has radically changed the commercial delivery sector, the process of preparing food at home has remained technologically neglected due to a lack of intuitive tools to facilitate daily grocery management [3], [4].

Current solutions in the smart kitchen domain primarily rely on specific and expensive IoT hardware, which is often inaccessible to the average consumer due to its high price [5]. Such systems additionally face technical rigidity, limited perspective of interior cameras and rapid software obsolescence of hardware components [6]. A clear research gap has been recognized in the literature: although modern smartphones possess exceptional sensor potential, it has not yet been adequately utilized for the purpose of wider availability of smart culinary solutions [5], [6].

The main goal of this work is to design and implement an architecture that transforms a mobile device into an active culinary assistant. The system uses multimodal artificial

intelligence models to automate the recognition of available resources and enable dynamic recipe generation in real time. This approach enables the transition from passive advice to proactive management of culinary knowledge, thereby directly promoting sustainability through maximum utilization of existing household supplies.

The paper is organized into six logical chapters. After introductory considerations, the second chapter defines the theoretical framework of digital culinary ecosystems and multimodal intelligence. The third chapter explains in detail the solution architecture based on current principles, while the fourth chapter presents the key functionalities and user flow of the developed application. The fifth chapter focuses on a discussion of technical contributions and guidelines for further development, while the sixth chapter summarizes the conclusions and effects of the proposed innovation.

2 Theoretical background

2.1 Digital culinary ecosystems

The modern digital transformation in the culinary sector is no longer a mere adaptation of information technologies, but a fundamental reconfiguration of the entire value chain [3][7]. Digital culinary ecosystems have evolved from static information portals to dynamic, data-driven platforms that manage the entire nutritional cycle - from raw material procurement to final distribution [8]. A key driver of this change has been the emergence of online food delivery (OFD) platforms such as Wolt, Glovo, Uber Eats and DoorDash. By integrating advanced logistics algorithms, artificial intelligence for demand forecasting and real-time GPS tracking, these platforms have redefined consumer expectations for transparency and speed of service [9]. Global online delivery market to reach $289 billion in 2024, becoming the primary interface between restaurants and end users [10].

The development of mobile culinary services and the growing popularity of delivery platforms have led to a new business model in the hospitality industry, known as "dark kitchens" or "ghost kitchens" [11]. This model eliminates the costs of maintaining restaurants and renting expensive locations, thereby achieving operating margins 15–20% higher than traditional formats [10]. Multiple virtual brands can operate in one space, which allows for better use of equipment and labor. For example, the same kitchen can prepare healthy meals under one name in the morning, and in the evening operate as a pizzeria or fast food restaurant with a different appearance in the application. Revenues per square meter in these facilities are up to three times higher [11].

Against high commercial efficiency, household users face the phenomenon of "hyper-choice" and the paradox of choice [4]. The problem of the 'paradox of choice' within unsystematized databases often results in the phenomenon of 'decision fatigue', which leads users to repetitive eating patterns and nutritionally suboptimal choices [4][12] In contrast to modern generative solutions, conventional digital chefs function exclusively as static stores of information that require a predefined outcome from the user. Given that poor organization within households causes about 48% of total food waste, this

technology gap has been identified as a critical point of systemic inefficiency in resource management [12].

2.2 Smart culinary environment

While mobile services have transformed external communication with the culinary sector, Smart Cooking represents an advanced phase that integrates artificial intelligence (AI) and the Internet of Things (IoT) directly into the processes of food preparation and storage [13][14]. This concept is defined as a synergy of hardware, software and data analytics with the aim of automating and personalizing the culinary cycle [14]. The IoT infrastructure enables various devices to function as connected nodes that collect data and optimize resources in real time [15]. In such an ecosystem, smart devices automatically download parameters from the cloud platform without the need for manual adjustment [13]. Fridges and ovens equipped with internal cameras, sensors and touch-sensitive interfaces stand out as the hub of a smart kitchen [7][16]. These devices allow users to remotely view inventory, track expiration dates, and automatically order groceries through integrated delivery platforms [7][17].

Despite the technological potential, mass deployment of IoT systems faces critical barriers:

- Economic restrictions: The high initial investment required for smart devices far exceeds the cost of traditional alternatives, making them less accessible to the average consumer [5].
- Software-induced obsolescence: Unlike traditional mechanics that last for decades, smart devices lose support after a few years. Through a process known as "software tethering," manufacturers retain control over device functionality, turning software support into a tool to force new purchase cycles [18][19].
- Technical rigidity and "2D limitations": The effectiveness of indoor cameras is limited by the inability to detect food that is hidden, wrapped in foil or placed in drawers [5][6]. Also, protocol fragmentation and lack of interoperability between different brands (vendor lock-in) make it difficult to create a coherent experience [5].

These challenges have led to a paradigm shift towards the democratization of technology [6] [20]. Instead of buying expensive, dedicated hardware, the focus is shifting to software solutions that make use of smartphones' already existing processing power and sensors [6][21]. Mobile devices, equipped with high-resolution cameras and multimodal AI models (like Google Gemini), take on the role of the "eyes" and "brains" of the kitchen. This approach bridges the gap between physical foods and digital analytics through computer vision and generative recipe synthesis, making smart cooking accessible to a wide range of users without additional financial investment[6]. In this context, 'Sous Chef' functions as a software-defined node within the wider smart home ecosystem. Although currently relying on a smartphone, the architecture is designed for future interoperability with other smart devices, positioning the mobile application as a central hub for the smart kitchen environment.

2.3 Multimodal intelligence in cooking

The modern paradigm of smart cooking is based on the transformation of the mobile device from a passive means of communication to a proactive culinary knowledge management system, with multimodal artificial intelligence (AI) taking on the role of the central cognitive core [22][6]. This advancement is enabling the redefinition of culinary spaces, shifting the focus from manual data entry to highly efficient and automated environments where AI bridges the "contextual gap" between mere visual identification and real-time practical application of knowledge [6] [23].

Modern culinary assistants are increasingly implementing "decoupled" multimodal pipelines to balance precision and operational efficiency [6], [22]. In this framework, visual models like EfficientNet-B4 (with a proven accuracy of 89.0%) function as "visual tokenizers" that transform photo pixels into semantic tokens (food names), eliminating the barrier of manual data entry [6]. Applying models like YOLO-World enables real-time object detection using an open dictionary, thus overcoming the limitations of fixed cameras in smart refrigerators [22].

After identifying resources, large language models (LLMs) like Gemini 1.5 Pro take on the role of a dynamic knowledge synthesis engine, achieving a factual accuracy of 9.2/10 [6]. Unlike static databases, these models synthesize instructions according to the user's current inventory and skill [24]. The stability of the e-business process is ensured by response engineering (Prompt Engineering), which forces the model to generate a structured JSON format, enabling seamless display of data within mobile interfaces [2].

AI service integration solves decision fatigue through predictive analytics, with personalized plans seeing 15% higher adherence rates [7][23]. Technology is moving towards an "agent era" with multimodal AI agents (like Gemini 2.0) capable of autonomously managing the cycle: from automated grocery shopping to programming smart devices [7]. In this way, AI evolves into a "resource keeper" that reduces household food waste by an estimated 30% to 40% [23].

2.4 Sustainability in a smart culinary ecosystem

The digital transformation of the culinary sector directly addresses the critical issue of sustainability, given that approximately a third of all food produced is thrown away globally [1][25]. A significant part of this waste is generated within households due to suboptimal meal planning, excessive procurement and neglect of food items that remain hidden in stocks until the expiration date [25]. In this context, smart culinary applications assume the strategic role of "resource custodians" through the active promotion of the "cook with what you have" philosophy [1]. Instead of a traditional procurement model for specific recipes, intelligent technology enables creative use of existing raw materials, while advanced systems through barcode scanning and automatic expiration notifications can reduce household food waste [26].

Future culinary services will completely connect food procurement, meal preparation and personal health into one unified system. Technologies such as health sensors, AR glasses and automated ordering will enable hyper-personalized nutrition adapted to

the current needs of the organism. In such an environment, visual recognition of foods and smart generation of recipes become the key foundations of modern cuisine.

3 Solution architecture

More generally, the architecture of the "Sous Chef" system is designed as a modern response to the need to digitize culinary processes within households, shifting the focus from expensive hardware to flexible software solutions. The system is based on a client-server model, where the functionality relies on the synergy of the mobile application, scalable cloud infrastructure and external artificial intelligence services.

3.1 Technologies used

The backend architecture utilizes Supabase, a Backend-as-a-Service (BaaS) platform that eliminates local server maintenance through a PostgreSQL database ensuring the integrity of culinary, history, and inventory data[27][28]. Security is enforced via Row Level Security (RLS), isolating data at the user account level [27].

The frontend is developed using React Native, enabling a uniform cross-platform code-base for iOS and Android that accelerates development and simplifies maintenance [29]. This approach significantly speeds up the development cycle and makes it easier to maintain the application. Additional value is achieved by leveraging Expo, a modern SDK that streamlines integration with native device functionalities - such as the camera and local storage. This capability is critical for applications operating within dynamic, real-world culinary environments [30].

The primary functionality of the system is based on the integration of multimodal AI models that simultaneously process visual and textual data, utilizing the mobile device as a real-time visual sensor [6]. This workflow operates in two phases: Resource Analytics, which automatically identifies food items via computer vision and Knowledge Synthesis, which transforms those ingredients into recipes using generative AI with structured JSON outputs [6], [31].

3.2 System architecture and data flow

The following diagram (Fig. 1) will illustrate a three-tier application architecture that relies on modern cloud services, eliminating the need for a classic monolithic backend server.

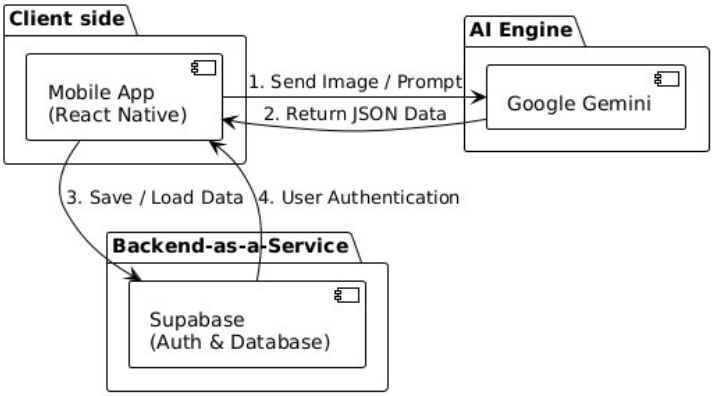


Fig. 1. Application architecture

The system is divided into three key modules:

1. Client side (React Native/Expo): The central orchestrator of the system in charge of collecting visual inputs (photos) and displaying the results. The application does not perform heavy processing locally, but manages the communication between the user, the AI model and the database.
2. AI Engine (Google Gemini): An intelligent layer that recognizes foods and synthesizes recipes through multimodal analysis. A key advantage is the delivery of data in JSON format, which enables immediate mapping of information into the user interface without additional parsing.
3. Backend-as-a-Service (Supabase): A complete infrastructure that provides secure user authentication and persistent data storage in the PostgreSQL database. Enables real-time synchronization of scan history and inventory status with minimal latency.

4 Key functionalities and validation of the system

The "Sous Chef" application is designed as an intelligent system that uses multimodal artificial intelligence to automate food management and personalize the culinary process. Key system functionalities that stand out are:

- Intelligent detection: Using the camera and the Google Gemini model, the system automatically recognizes food items, eliminating the need for manual data entry.
- Contextual recipe synthesis: AI generates recipes in real time based on the current state of the refrigerator, taking into account the permanent stock from the user's "Pantry".

- Personalized filtering: The user defines parameters such as the number of servings, preparation time and weight, which directly affects the structure of the generated JSON output from the AI model.
- History and favorites management: All results are persistently stored in the Supabase database, allowing quick access to favorite dishes and previous scans.

Interaction begins on the Home screen, which serves as the operating core of the system, providing the user with immediate access to the central scanning function while simultaneously providing insight into the history of recent activities. The following image shows the Home screen (see Fig. 2).

The process continues through the Detected Ingredients module (Fig. 3), which serves as a demonstration of computer vision capabilities within a constrained environment. At this level, the multimodal model performs semantic tokenization of image content, converting pixels into digital entities. This screen allows the user to critically validate the AI extraction, which ensures the integrity of the data that will be passed on to further generative processing.

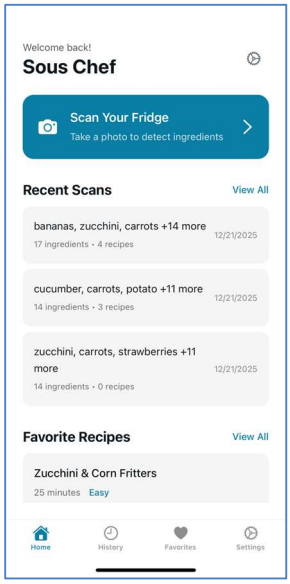


Fig. 2 Home screen

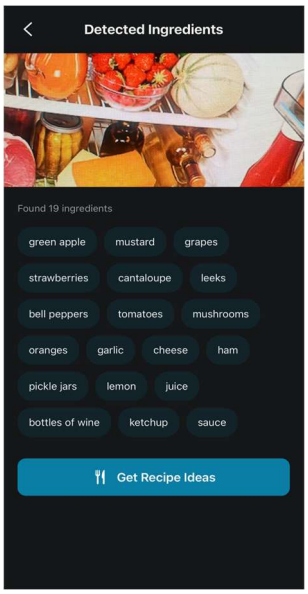


Fig. 3 Detected Ingredients

After confirming the input, the system displays a list of proposed recipes - Recipe Ideas (Fig. 4), which is a direct manifestation of generative AI logic. Instead of a static search, the application offers a selection of dishes that are dynamically synthesized based on identified foods and set user preferences, offering an optimal balance between available resources and culinary creativity.

The cycle is rounded off by the Recipe Details screen (Fig. 5), which delivers the final product of the system. Here, abstract JSON data is transformed into structured and personalized instructions, providing the user with a clear roadmap for completing the

task. This significantly simplifies data processing for the end user, completing the journey from visual detection to the final realization of the meal.

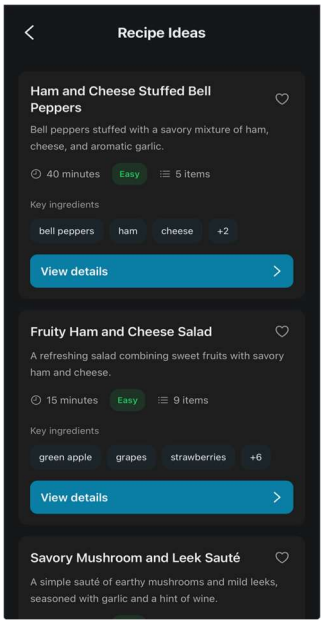


Fig. 4 Recipe Ideas

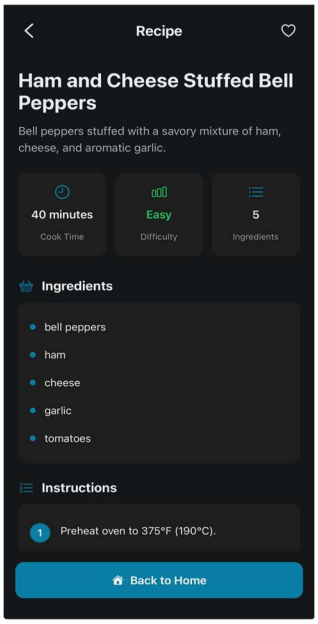


Fig. 5 Recipe Details

4.1 Initial System Validation

In order to perform an initial validation of the developed solution in a real-world environment, a case study was conducted with one end-user (amateur chef). The evaluation focused on usability aspects, interface adaptability and the robustness of the multimodal model when processing 19 detected ingredients (Fig. 3). The user went through the complete life cycle of the application: from registration and refrigerator scanning to dynamic recipe synthesis using the Gemini 1.5 Pro model.

The system successfully transformed pixels into semantic tokens and generated a structured JSON response that enabled an intuitive display of preparation steps without latency in parsing. The user particularly emphasized the speed of access to information and the minimal number of interactions required to generate a recipe, which directly affects the reduction of cognitive load and solving the problem of 'decision fatigue'.

In addition to detection accuracy, the validation included the evaluation of generative logic through the Recipe Ideas (Fig. 4) and Recipe Details (Fig. 5) modules. The user confirmed that the list of suggested dishes (Fig. 4) faithfully reflects the 'optimal balance between available resources and culinary creativity'. The transformation of abstract JSON data into structured instructions on the screen with details (Fig. 5) is particularly significant, which provides the user with a clear roadmap for the realization of

the meal. This confirmed that the system successfully bridges the 'contextual gap' between the mere identification of objects and their practical application in real time. The final test showed that personalized filtering (preparation time and number of portions) directly affects the structure of the output JSON, thus ensuring the relevance of the instructions for the specific needs of the user.

5 Discussion

Implementation of the mobile service "Sous Chef" has confirmed that the gap between physical food and digital data can be successfully bridged with the use of modern phones and artificial intelligence. The key advantage of this solution is the elimination of manual data entry, which in the past was the biggest obstacle for regular use of similar applications.

The integration of the Google Gemini model allowed the system to convert a photo of a refrigerator or pantry into a digital grocery list with high accuracy, with the model proving extremely effective in recognizing fruits, vegetables, and standard packaged goods in adequate lighting. Although the tests confirmed high functionality, certain technical challenges were also identified in the form of a slight drop in accuracy in low light conditions or when the food is physically obscured by other objects, which represents a primary area for future software improvements and further optimization of the system.

The choice of React Native and Expo environment ensured the functionality of the app on Android and iOS platforms while using a single code, thus making the concept of smart cooking widely accessible to smartphone users instead of being limited only to expensive and specific IoT devices. At the same time, security is solved at the level of the base itself through the Supabase platform and RLS (Row Level Security) policies, which guarantee complete isolation and protection of personal data, such as scan history and favorite recipes.

A significant contribution of the proposed system lies in the reduction of "decision fatigue" and the fight against food waste, as the system offers ready-made solutions that save the user time and energy needed for meal planning. By promoting the principle of "cook with what you already have", the app encourages the use of existing supplies before they expire, directly addressing the problem of poor planning that is responsible for almost half of all food waste generated in households.

Usability aspects, observed during the initial validation, indicate that the responsiveness of the interface directly depends on the efficiency of asynchronous data transfer to the Gemini API. The structured JSON format enables an intuitive display of data on the 'Recipe Ideas' and 'Recipe Details' screens, which significantly facilitates interaction in a dynamic environment.

5.1 Future work

The plan for the future development of the "Sous Chef" service is focused on three key directions: technical upgrading, personalization according to health parameters and the introduction of innovative ways of interaction.

Initial validation confirmed that the physical obscuration of food (occlusion) represents a significant technical challenge, which is consistent with earlier research on the spatial limitations of 2D cameras. To address these edge cases and ensure maximum data integrity regardless of lighting conditions or item layouts, first phase of improvements will focus on implementing a barcode scanning function for packaged products, thereby eliminating recognition errors for factory-declared items. Furthermore, integration with the Supabase Storage service is planned to enable secure storage and cross-device synchronization of food images. Finally, the system will be extended to enrich each generated recipe with a detailed display of nutritional values, providing users with a comprehensive view of caloric intake and meal composition.

A more advanced stage of development involves connecting the application with wearable technology, such as smart watches and medical sensors. In this way, the system could dynamically adjust recipe suggestions based on the user's real physiological data – such as current glucose levels, energy consumption or specific health conditions. This approach transforms the application from an ordinary culinary assistant into a tool for personalized health management.

As the final step in the digital transformation of the cooking process, the application of augmented reality (AR) is planned. Using smart glasses or mobile devices, cooking instructions would be projected directly over the physical ingredients in the kitchen. Such real-time visual assistance would allow users to follow the preparation steps without having to take their eyes off the desktop, making the cooking process more efficient and intuitive.

6 Conclusion

The development of the "Sous Chef" application confirmed that the synergy of mobile development and multimodal artificial intelligence facilitates the utilization of smartphone hardware for proactive culinary resource management. By using advanced models like Gemini 1.5 Pro, in-depth understanding of food context and high accuracy recipe generation in real time is achieved. The key results of the work indicate a significant paradigm shift: the digital transformation in cooking is no longer exclusively about food delivery services, but the focus is shifting towards smart and personalized management of resources within the household itself.

The initial validation confirms that integrating multimodal intelligence into a mobile service is not only technically feasible but also practically applicable within the smart home ecosystem. In conclusion, this service represents a solid foundation for the development of future culinary ecosystems that directly save users' time and money, while taking on the role of "resource custodians". By promoting more efficient use of supplies, "Sous Chef" offers a concrete answer to the global challenge of food waste, thereby directly contributing to environmental sustainability and more responsible

consumption. In the era of widespread artificial intelligence adoption, the application serves as an example of practical and ethical application of technology that brings real, concrete value and improves the quality of everyday life in a sustainable way.

Acknowledgments. This research was supported by the University of Belgrade – Faculty of Organizational Sciences and, in part, by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia through institutional funding (grant number: 200151).

Acknowledgment of AI Use: During the preparation of this manuscript, AI-based tools were used exclusively for translation assistance, language editing, and improving readability. All scientific content, methodology, system architecture, and conclusions were entirely developed and verified by the authors. The authors take full responsibility for the integrity and accuracy of the final work.

References

1. A. Zick, X. R. Schmidt, and C. Reynolds, "Operationalising sustainability in professional kitchens: The interplay of chef competencies, environmental values and human resource management strategies," 2025, doi: 10.1016/j.ijgfs.2025.101330.
2. A. D. M. R, V. KUMAR, B. Kumar, A. Raj, and A. K. Suman, "AI-Powered Voice-Controlled Kitchen Assistant for Hands-Free Cooking Support," *Iconic Research And Engineering Journals*, vol. 9, no. 6, pp. 1689–1691, Dec. 2025, doi: 10.64388/IREV9I6-1713013.
3. A. Zibaie, M. Beliaev, M. Alizadeh, and R. Pedarsani, "Dynamic Pricing for Multi-Modal Meal Delivery Using Deep Reinforcement Learning," *Future Transportation 2025, Vol. 5, Page 112*, vol. 5, no. 3, p. 112, Sep. 2025, doi: 10.3390/FUTURETRANSP5030112.
4. A. P. Zibran, R. Hajar, and P. Sejati, "Restaurant Sales Optimization Through Collaborative Filtering and Market Basket Analysis Methods," *Journal of Scientific Research, Education, and Technology (JSRET)*, vol. 4, no. 4, pp. 2803–2825, Dec. 2025, doi: 10.58526/JSRET.V4I4.974.
5. B. Basarir-Ozel, H. B. Turker, and V. A. Nasir, "Identifying the Key Drivers and Barriers of Smart Home Adoption: A Thematic Analysis from the Business Perspective," *Sustainability 2022, Vol. 14, Page 9053*, vol. 14, no. 15, p. 9053, Jul. 2022, doi: 10.3390/SU14159053.
6. R. Khoirul Anam, "Evaluating Gemini LLM in Food Image-Based Recipe and Nutrition Description with EfficientNet-B4 Visual Backbone," *ArXiv*, p. arXiv:2511.08215, Nov. 2025, doi: 10.48550/ARXIV.2511.08215.
7. Raksha Sharma, "AI-Generated Cooking Recipe Market Research Report 2034," Mar. 2026.
8. A. Shokeen, "IMPACT OF DIGITAL FOOD DELIVERYPLATFORMS ON CONSUMER BEHAVIOR: ANINTERNATIONAL PERSPECTIVE," *JETIR2603866 Journal of Emerging Technologies and Innovative Research*, vol. 13, 2026.

9. Khushi Rastogi, “Food Delivery Market: Digital Shift & Growth in 2026.” Accessed: Apr. 23, 2026. [Online]. Available: https://www.kenresearch.com/articles/food-delivery-market-digital-shift-and-opportunities
10. “US Dark Kitchens/Ghost Kitchens/Cloud Kitchens Market Size.” Accessed: Apr. 23, 2026. [Online]. Available: https://www.coherentmarketinsights.com/industry-reports/us-dark-kitchens-ghost-kitchens-cloud-kitchens-market
11. D. T. da Cunha, M. P. Hakim, M. M. Alves, M. S. Vicentini, and M. Z. Wiśniewska, “Dark kitchens: Origin, definition, and perspectives of an emerging food sector,” *Int. J. Gastron. Food Sci.*, vol. 35, p. 100884, Mar. 2024, doi: 10.1016/J.IJGFS.2024.100884.
12. S. Jangra, “THE IMPACT OF CONSUMER BEHAVIOR ON THE GROWTH OF ONLINE FOOD DELIVERY SERVICES,” *Asia Pacific Journal of Research in Business Management*, vol. 16, no. 7, Jul. 2025.
13. S. Raj, A. Jana, A. Ps, S. Lazar, H. Vidya, and A. Prof, “Survey On Smart Cooking Technologies,” *International Journal of Computer Science Trends and Technology*, vol. 11, 2023.
14. K. P. Johare, D. R. Vasant, G. Wagh, D. R. Arvind, and D. Shaligram, “Present and Future Possibilities for Intelligent Kitchen with AI and IoT,” *Iconic Research And Engineering Journals*, vol. 6, no. 3, pp. 122–127, Sep. 2022, Accessed: Jun. 02, 2026. [Online]. Available: https://www.irejournals.com/paper-details/1703813
15. P. Shahi, “Factors Influencing the Adoption of Smart-home Appliances: Users’ Perspectives,” University of Missouri-St. Louis, 2022. Accessed: Jun. 02, 2026. [Online]. Available: https://irl.umsl.edu/dissertation/1273
16. S. Ceccacci, R. Menghi, and M. Germani, “Example of a New Smart Kitchen Model for Energy Efficiency and Usability,” in *SMART 2015, The Fourth International Conference on Smart Systems, Devices and Technologies* , Brussels, Belgium, Jun. 2015, pp. 12–18.
17. Data Bridge Market Research, “Smart Kitchen Appliances Market Growth Surges from USD 21.66 Billion to USD 68.64 Billion by 2032.” Accessed: Apr. 23, 2026. [Online]. Available: https://www.openpr.com/news/4477053/smart-kitchen-appliances-market-growth-surges-from-usd-21-66
18. Euroconsumers group, “Software obsolescence as abusiness model?” Accessed: Apr. 23, 2026. [Online]. Available: https://www.euroconsumers.org/wp-content/uploads/2025/10/Software-obsolescence-as-a-business-model.pdf
19. E. Poppe, E. Wagner, M. Jaeger-Erben, J. Druschke, and M. Köhn, “Is it a bug or a feature? The concept of software obsolescence,” 2021, doi: 10.34961/1754.
20. “Mobile App Development | Chat Bots Development | A-bots.” Accessed: Apr. 23, 2026. [Online]. Available: https://a-bots.com/blog/app-development-for-restaurants
21. “Voice-controlled Kitchen Devices Market Size, CAGR, Trends and Forecast 2034.” Accessed: Apr. 23, 2026. [Online]. Available: https://dimensionmarketresearch.com/report/voice-controlled-kitchen-devices-market/
22. G. Asada and A. Hedbäck, “Computer Vision in Cooking Support SystemsLeveraging Multimodal LLMs for Dynamic, Open-Vocabulary,Real-Time Object Detection and Tracking,” *KTH Royal Institute of Technology*.

23. "AI Kitchen Market Growth Analysis - Size and Forecast 2025-2029 | Technavio." Accessed: Apr. 23, 2026. [Online]. Available: https://www.technavio.com/report/ai-kitchen-market-industry-analysis
24. "Recipe Generation AI Market Research Report 2033." Accessed: Apr. 23, 2026. [Online]. Available: https://marketintelo.com/report/recipe-generation-ai-market
25. S. Bolwig, A. N. Tanner, P. Riemann, B. Redlingshöfer, and Y. Zhang, "Reducing consumer food waste using green and digital technologies," 2021, *UNEP DTU Partnership*. Accessed: Jun. 02, 2026. [Online]. Available: https://orbit.dtu.dk/en/publications/reducing-consumer-food-waste-using-green-and-digital-technologies/
26. J. Liegeard and L. Manning, "Use of intelligent applications to reduce household food waste," 2019, doi: 10.1080/10408398.2018.1556580.
27. "Supabase Documentation." Accessed: Apr. 24, 2026. [Online]. Available: https://supabase.com/docs
28. "PostgreSQL: Documentation." Accessed: Apr. 24, 2026. [Online]. Available: https://www.postgresql.org/docs/
29. "Introduction · React Native Documentation." Accessed: Apr. 24, 2026. [Online]. Available: https://reactnative.dev/docs/getting-started
30. "Expo Documentation." Accessed: Apr. 24, 2026. [Online]. Available: https://docs.expo.dev/
31. "Gemini API | Google AI for Developers." Accessed: Apr. 24, 2026. [Online]. Available: https://ai.google.dev/gemini-api/docs

Original Research Article *https://doi.org/10.5281/zenodo.23475555*

A Web Application for Understanding Conflicting Historical Narratives with User Contributions and GenAI

Petar Tomić[1][0009-0001-9964-5938] and Miloš Jolović[1][0009-0003-6580-2039]

[1] University of Belgrade – Faculty of Organizational Sciences, Jove Ilića 154, 11000 Belgrade, Serbia

Abstract. This paper aims to present an application which generates historical narratives based on facts submitted and reviewed by users, with the help of AI. The purpose of this application is to engage users in contributing to historical knowledge by collecting a multitude of perspectives from them, and to provide the same users with insights into diverse views on the same topics, seen here as a necessity due to history's interpretive nature. The paper reviews existing digital solutions in the domains of crowdsourcing and generative artificial intelligence and their application in the field of history. A new software solution has been designed and implemented as a functional hybrid (web and mobile) application. Limitations and challenges of such a system remain, including encouraging good user behavior in the crowdsourcing model and addressing issues of financial efficiency of AI API calls.

Keywords: digital history, crowdsourcing in history, AI history, historical narratives.

1 Introduction

Today, history is a frequent topic for content on digital platforms. Ecosystems for learning history, which encourage interactions between users, the presence of a multitude of sources, but also the cooperation of lay people and science in the creation of content, have gained more traction than traditional learning methods. However, new methods struggle with the problems of superficiality and bias, due to most modern systems being based on maximizing user engagement through algorithms that will only pander to them and deepen divisions in both objective knowledge of facts and deeper contextual understanding. Recently, LLMs and other forms of AI brought a new host of issues.

AI tools bring numerous uses and challenges to scholars of history. The ability to analyze and synthesize information provided by modern AI tools should be combined with the knowledge of scientists and experts (Boroș et al., 2024). The contributions of non-professional enthusiasts to historical awareness are increasingly pronounced thanks to the democratization of the study of history brought about by today's digital platforms. Such models have many advantages: accessibility, a breadth of experience and connections between users, but also disadvantages: unverified information, group efforts in rewriting history, and inability to understand context (Kaldeli et al., 2021).

The specific idea of this paper is to combine user-generated content in the form of submitting and voting on historical facts, and generative artificial intelligence services in order to reveal different historical narratives, all while incorporating professional expertise of trained historians. Users of different levels of education, political orientations and other demographic categories (such as nationality, gender, age) could submit facts and vote for their accuracy. After some submissions have been made and votes have been cast, they could view the narrative that is most prevalent among a specific group of users on a specific topic.

The theoretical principles are described in Chapter 2, with Chapter 2.1 reviewing the literature in the area of digital history and its crowdsourcing, and Chapter 2.2 focusing on forays of LLMs in this field. Chapter 3 presents the developed application, with Chapter 3.1 detailing the idea and purpose, Chapter 3.2 its design and 3.3 its implementation in Laravel, Angular and Ionic (example of one use case). Chapter 4 covers the testing of the application (Chapter 4.1), and analyzes the results (Chapter 4.2). Chapter 5 concludes the paper and outlines future development directions.

2 Theoretical Background

2.1 Digital History and Crowdsourcing

Democratization. History is one of the most popular topics on the Internet. It gathers a large number of interested laypeople and enthusiasts, as well as historians with formal education. The Internet is full of historical content presented for different levels of expertise and in all languages. Research has shown that there is a significant degree of participation of the general public in online platforms related to history that are easily accessible.

So-called “digital historiography” has its own methods that have proven to be as good or better than traditional ones (Ridge, 2016). Awareness of history is often a decisive factor in the formation of a person's social thinking, their political orientation and life values, observing events through certain “filters” (Zanazanian, 2019). Newer methods of teaching history in modern educational systems emphasize the importance of understanding such “filters” as interpretations, as it fosters critical historical awareness (Paricio, 2021).

Crowdsourcing Platforms. Crowdsourcing means outsourcing content creation to a wider audience. Examples of these history-related applications are often based on the collection of local stories, myths or facts on a system where it is possible to enter the event on a map, describe it in words, or add images or videos (Ziku et al., 2024). Another example is using user knowledge to annotate large historical datasets, contributing to the enrichment of meta-data in digital archives (Kaldeli et al., 2021). In one example from Greece, an event mapping application was used to collect information about the Knights of Rhodes, and users showed a high level of interest in contributing content, despite having no previous experience in similar projects (Ziku et al., 2024).

Some research has shown that younger demographics and students resonate with participatory systems and digital narratives, especially when their credibility is bolstered with precise citations and visible qualifications (Alkhashil et al., 2026). Numerous crowdsourcing and participation-based projects in the field of history attest to the validity of this method of both gathering and learning information.

2.2 Generative AI in Historical Scholarship

Early Application. The application of generative artificial intelligence in history and its related sciences, as well as in most social disciplines, is still in its infancy. While models are only getting better, researchers still emphasize that even advanced models require careful supervision to reduce errors in conclusion that arise from misunderstanding historical context in automatically generated historical narratives (Boroş et al., 2024).

The first field of application where this technology found its place is the analysis of historical documents: their transcription, normalization and digitization. One study showed satisfactory results achieved by ChatGPT in the normalization of texts from the letters of a Sicilian noble dynasty (Spina, 2023).

An area that is important for historical sciences, and in which artificial intelligence gives promising results, is text translation. Existing AI translation tools have democratized the translation process, improved efficiency and scalability, but they are not perfect. They face numerous challenges, such as culture-specific terms, idiomatic expressions and ambiguous contexts, which often lead to incorrect translation and still require a balanced approach that combines AI technologies with the expertise of professionals (Shahmerdanova, 2025).

Knowledge and Hallucination. A question that often arises in the application of any AI tool is the accuracy of the answers they provide. Current AI models are known to hallucinate, or provide a presenting false or misleading information as fact. Because of this, benchmarking methods have been developed to measure the accuracy of AI models.

Research on benchmarking historical knowledge of AI models has been scarce. From the results that do exist, the GPT-4 model achieved the highest score on the knowledge test for advanced high school American history classes, and the second highest on the world history test, which is a significant improvement over the GPT-3, but not perfect (Achiam et al., 2023). To significantly reduce hallucinations, many systems use Retrieval-Augmented-Generation (RAG) frameworks that involve data from trusted sources (Rahman et al., 2026).

It can be concluded that the answers that the user receives from currently available artificial intelligence services should not necessarily be considered correct, especially if it is about more complex questions and lesser known areas of history. Scientists conclude that the obtained results do not show knowledge at an expert level, but exceed the probability of random guessing, and certain historiographical biases of the model are also noted (Hauser et al., 2024). On the other hand, AI-assisted learning presents many

advantages, such as the ability to engage the learner in a more direct way: asking questions and getting clear answers, rather than passively consuming long and complicated text. Research indicates that reading AI summaries performs better than more traditional digital learning methods (reading blog posts or Wikipedia articles), both when the text is labeled as AI-generated and when it isn't (Karell et al., 2025).

The benchmarking itself has been subject to criticism from some scholars, asserting that there's a noticeable gap in LLM performance based on the questions presented in the benchmark. In multiple-choice questions, focused on Western knowledge domains, models excel, but they falter when faced with open-ended questions and when global knowledge is required. Scientists argue that these gaps can be explained by the nature of LLMs relying on pattern recognition, and not actual historical reasoning, and emphasize the need for developing better benchmarks (Hutchinson, 2026). Reliable data for SOTA models doesn't exist at the time , but the accuracy of some LLMs in answering history questions is shown in Figure 1. LLMs struggled with 4 available answers (~ 40% vs 25% at random), as well as with 2 available answers (~ 60% vs 50%)

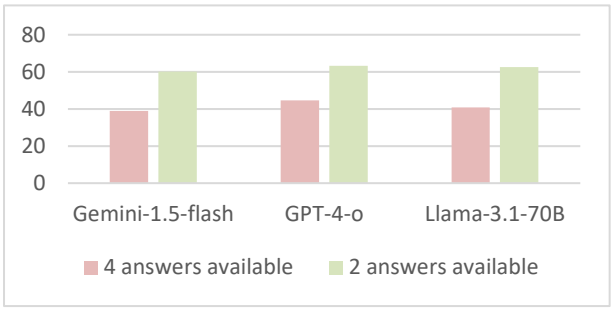


Fig. 1. Accuracy of commercial LLMs in answering history questions (adapted from Hauser et al., 2024).

AI-Augmented Workflows. Some successful international projects applying AI to historical contexts have been developed recently. One of them is the Time Machine project, a European initiative that uses LLMs with Big Data analytics and other statistical techniques, with the aim of reconstructing and contextualizing historical data from different periods in urban environments (Time Machine Project, 2023). Many museums are applying artificial intelligence to digitize their exhibits and enrich the visitor experience (Capgemini, 2025).

There are experimental systems that seek to take advantage of the contextual analysis and reasoning capabilities that modern LLMs provide to extract and present facts in an interesting way (Zeng, 2024; Litaina et al., 2024; Parfenova, 2025). Some historians also constructed workflows incorporating AI as a tool, where it aided in multiple steps such as: primary source collection, thematic clustering, identifying temporal trends, summarization, construction of an interpretive argument and narrative analysis, while still heavily emphasizing human-in-the-loop validation (Henriot, 2025; Yao, 2025).

All the projects presented here show the power of applying generative artificial intelligence, and how science can use it in extraordinary ways to detect new patterns and

ease demanding work. However, by their very nature, AI responses, especially when they require precision in detail, will never be perfect, and it is important to always have disclaimers for users that the answer may not be correct, as well as real experts who would perform quality control.

The Problem of Bias. LLMs often face a problem of bias due to the nature of their training corpora. The so-called "document asymmetry" that exists in digital materials plays a role. Most of the digitized materials relate to the three most recent centuries of history, so the models understand them much better than documents from earlier periods (Santurkar et al., 2023). In the context of practical application, these biases in the models are most easily seen in favoring the narratives of Eurocentrism and presentism. To achieve true equity and fairness in digital historiography, transparency in algorithms, multilingual model training, and validation frameworks developed by domain experts should be advocated (Weidinger et al., 2022).

Aside from latent biases that come built into the model, users can measurably influence the LLM's output by framing the prompt to slant towards one ideological side, even while explicitly telling the model that it should present everything as factual accurate. In one study, researchers found that while the LLM had a latent liberal bias, there was a significant difference in users' understanding of the topic when it was presented them with a "liberal" model and a "conservative" model (Shu et al., 2026).

When we look more closely at the biases present in LLMs, we notice certain epistemological differences between LLMs and real historians. Many historians and other humanities researchers present arguments for a broader inclusion of professionals in this rapidly-advancing field (Klein et al., 2025). Corporate ownership of LLMs may prioritize engagement and profit over scientific values: transparency, reproducibility, accessibility and cultural diversity.

3 Designing a System for Crowdsourcing Multiperspectivity

3.1 Idea and Business Model

Countering the One-Narrative Approach. In the literature analysis outlined in the second chapter, certain shortcomings and needs can be identified. The interpretive nature of history is emphasized, so it is valid to say that alternative understandings of the same historical event, called narratives, are an integral part of history.

The truth cannot be properly understood if only one side of the story is heard, yet most modern education systems offer only that. An authoritative representation of history is presented directly, and often without any comments, to young students, susceptible to influences and with insufficiently developed critical thinking (Paricio, 2021). Students are rarely motivated to research on their own, to read sources that claim the opposite, which is crucial in understanding history. This is especially true when the history taught in school is instrumentalized to propagate nationalistic and other ideological narratives, which is a common problem in developing democratic countries.

As technology advances, the stated problem only seems to be getting worse, mostly thanks to social networks that only look to maximize user engagement through recommendation algorithms. They aim to retain as large an audience as possible for as long as possible, at the cost of fairness and truth. Users are constantly served content that leads them to think as little as possible and instead only confirm their biases. The ubiquity of this social network model testifies to its effectiveness in maximizing profits, however, research shows that it leads to political radicalization and social isolation (Alatawi et al., 2025).

Current AI models and the way they are commonly used threaten to further exacerbate the issue, as LLMs trained on asymmetric data inherit the same structural asymmetries. When asked to reason about contested events, LLMs reproduce the viewpoints of Western states at the expense of others, even when prompts included simple debiasing instructions (Salnikov et al., 2025). Humanities researchers frame this as a structural problem of the LLM architecture, and suggest a solution of deliberately incorporating diverse human knowledge (Klein et al., 2025).

AI-assisted Learning. Looking at the problem and its implications, the idea of this project arises, one which would aim to present history through multiple perspectives. Users (who provide the system with their personal data such as nationality, gender identity, political orientation, or age) are the ones who create the historical text. They do so in small steps, fact by fact. They also provide their judgment on the accuracy of other users' contributions. The LLM API later makes meaningful texts out of those facts, and users are able to see the dominant narrative believed by users from any social group that can be determined based on the data the system has about them, and on facts they've marked as true.

AI-assisted learning has proven effective in summaries of historical events (Karell et al., 2025), and it can be argued engaging the user through their want to know how different people tell the same story might have an even greater effect on learning. As a control mechanism, one must enlist the knowledge of historians, experts on specific topics, to whom the system would give some level of greater voting power. Users should also be allowed to rate each other and thus create a reputation system, where highly rated users have a more powerful voice because the community trusts them. This reputation system should ensure that the application would not become a battleground for coordinated historical revisionism, but also avoid excessive gatekeeping. All users contribute, while professionals offer expert validation to maintain epistemological integrity.

Business Model. The main target demographic would be historians and history enthusiasts, who would find such an application engaging and helpful. Others who might benefit are educators, journalists and civil society, as well as state institutions, as they would have a place where they can see which narratives are dominant in the population, but also which facts are well or poorly understood, helping them shape understand their students, readership and citizens, and shape policy accordingly. A summary breakdown of the potential business's key elements is shown in the Table 1.

The core of the model builds on the idea that value is created through aggregation of diverse perspectives into comparable narratives. Enabling users to both contribute and evaluate others user's inputs, the application establishes a feedback loop in which data quality improves with scale, with the end goal of the platform becoming an evolving knowledge system that reflects ongoing discourse. While early-stage adoption may face challenges due to this data quality scaling, the bulk of the costs remain on the side of LLM API calls, which also scale directly with use. It would be prudent to pursue technical independence from the AI service provider, through a smaller locally-run model.

Table 1. Business model canvas for the proposed idea

<table>
<tr>
<td rowspan="3">Key Partners

Academia

Civil society

Journalism outlets

Cultural heritage organizations</td>
<td>Key Activities

Development / maintenance

Moderation

Community building</td>
<td rowspan="3">Value Proposition

Comparing narratives in one platform

Living dataset of real dominant opinions among groups

Both enthusiasts and professionals

Democratized history education</td>
<td>Customer Relationships

Community-driven

Engagement

Institutional</td>
<td rowspan="3">Customer Segments

Enthusiasts

Professionals

Educators

Civil society

Policy makers</td>
</tr>
<tr>
<td>Resources

Server

LLM API

Codebase

Development team

Endorsements</td>
<td>Channels

Direct

Partnership distribution

Content marketing

API</td>
</tr>
<tr>
<td colspan="3">Cost Structure

LLM API calls

Development / maintenance / hosting

Community management

Outreach</td>
<td colspan="2">Revenue Streams

Institutional licensing

User donations

Grant funding

Premium features</td>
</tr>
</table>

3.2 System Design

The software solution for the problem of crafting and displaying conflicting narratives, relies on wide adoption. During the initial stages, however, content quality must be supported through seed facts generated by the experts and intensive administrator work to prevent bad actors until a healthy and self-sustaining userbase is reached.

Websites or mobile applications would be best suited to achieve this. Accessibility and low-friction interaction on a system with a familiar scrolling feed would result in the system getting more facts submitted and being able to generate more narratives. A hybrid application allows for both mobile and web versions. The mobile version would enable individual fact submissions and quick lookups, while the web versions would be more comfortable for longer reading, detailed contributions, and narrative comparisons.

Figure 2 shows the system design, including the selected technologies. The frontend Layer is built in Angular and Ionic frameworks, which enable a clean and consistent design across both desktop and mobile clients. The backend runs on Laravel, managing business logic and data. Relational databases are well suited for the application's data model. Gemini was chosen because there's a traditionally stable version of it which offers free API calls and relatively high-quality outputs, but when it comes to bias there are other, less biased and Western-centric models (Mak & Luo, 2025).

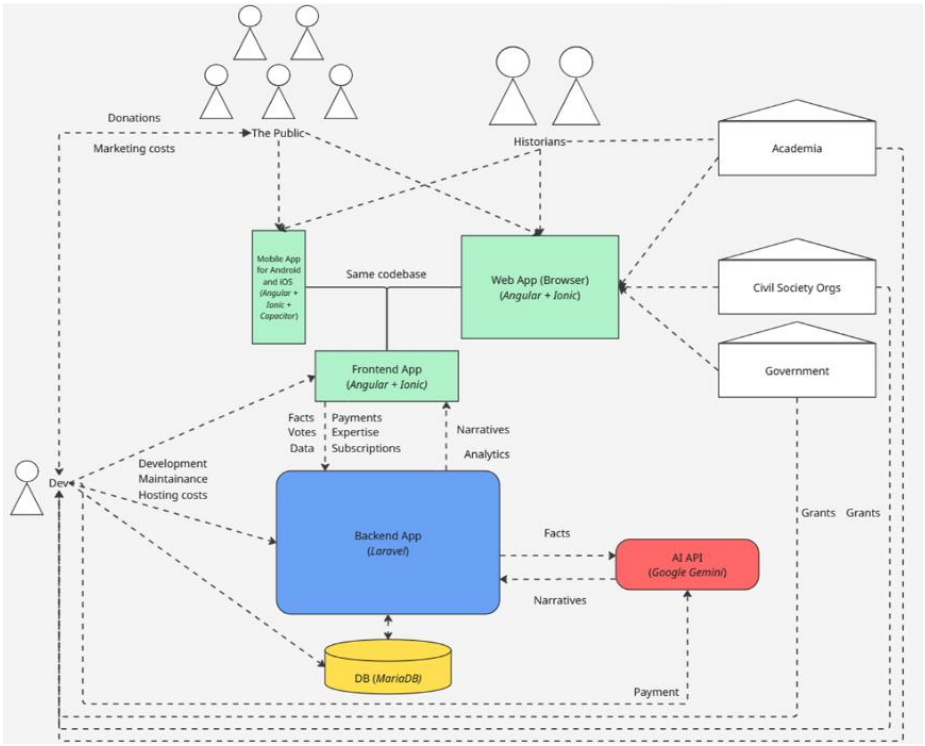


Fig. 2. System design

3.3 Implementation of a Core Use Case

This chapter presents the screens for the most basic use case, in the mobile version. Figure 3 (a) shows the form for adding facts. It works by the user entering the fact itself in the Text field and the source in the Source field. The topic is a search box from which the user searches for existing topics, however, the user has the option to select another topic, if they want, as shown here. For that new topic, they must enter the historical area (a broader term than the topic). The user's entered fact is looked at by the AI model and translated into English, pending user confirmation, shown in Figure 3 (b). Figure 3 (c) shows a scrollable feed of the entered facts.

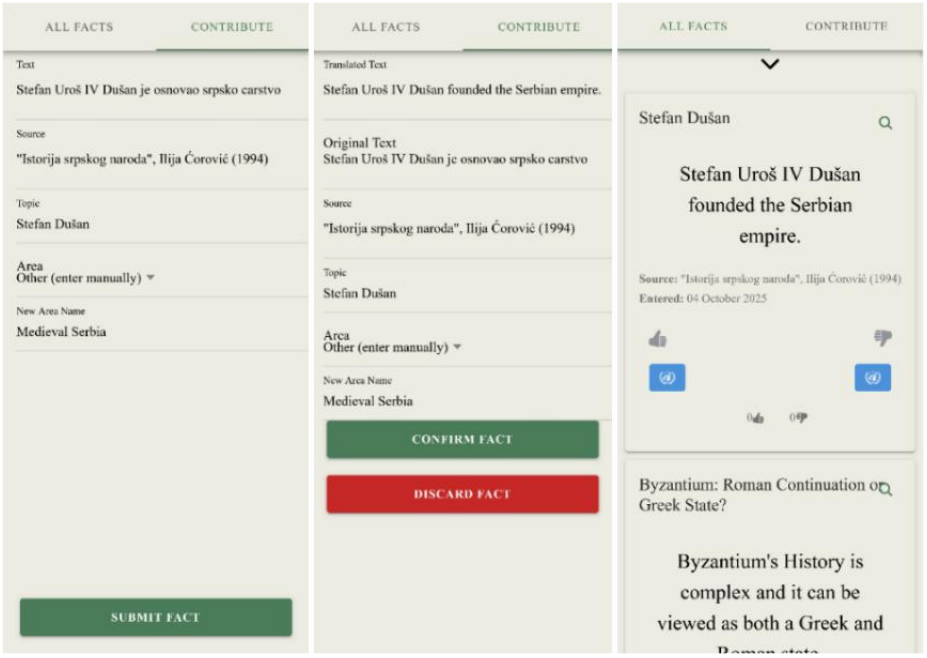


Fig. 3. Submitting (a) translating (b) and viewing (c) facts.

The scrollable feed part of the app displays an accuracy bar for each fact. The user mark a fact as accurate with the “like” button seen in Figure 3 (c), or as false by using the “dislike” button. When enough “likes” and “dislikes” have been made, a bar is formed with the green part representing a proportion of “likes” vs “dislikes”. Votes here are weighted by user expertise in the given area the fact is from, at the moment the vote is cast, so users can more easily tell if a certain fact is highly or lowly rated, taking into account the opinion of the right users. However, this expertise does not influence narrative the narrative generation, as the point of the narrative is to see how prevalent it is in a given population, not to see how deemed true it is by the general public, or even by the experts.

Figure 4 (a) shows the application's main menu, with options available for Home, Profile, Narratives, and Logout. Clicking on the Profile option takes you to the user's profile page, which has two tabs, the user's details, in Figure 4 (b) and the expertise page, in Figure 4 (c). An edit profile window is also available here, as well as an overview of other users and possible ratings of their expertise.

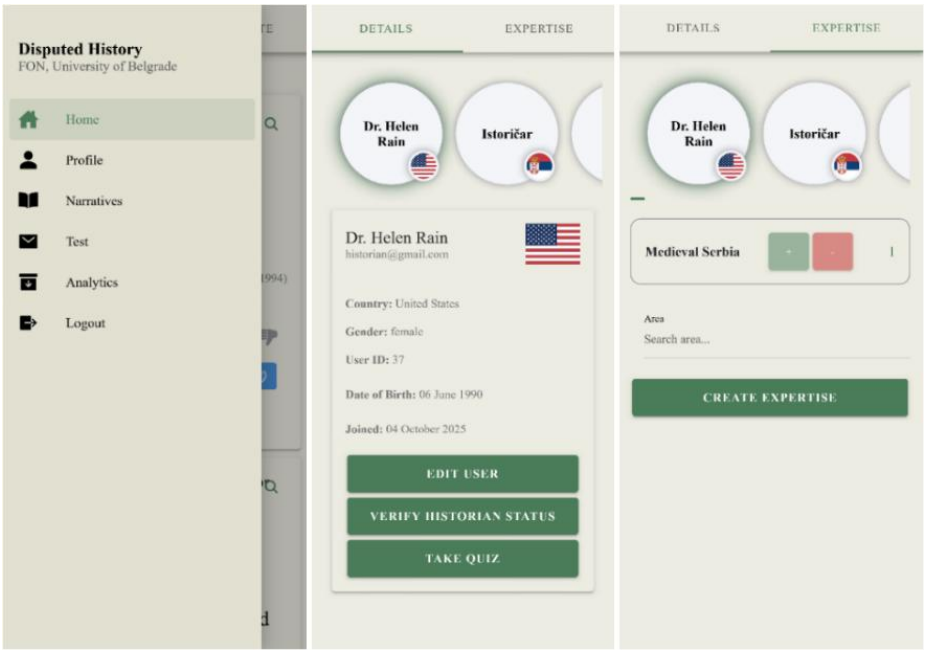


Fig. 4. Side menu (a), user profile (b), and their expertise (c).

The expertise seen in Figure 4 (c) is the basis of the mentioned reputation system, now only implemented as a flat modifier to the user's vote power in the given area the fact is topic of the fact is about. (For example, if a user has been deemed expert in the area of *Medieval Serbia* by 3 members of the community, their vote counts as 4 votes on any topics that relate to the area.) Also, by the current system if the user is a verified historian, their vote is weighted double. It should be noted that this system is a prototype, made in the spirit of implementing the idea of rewarding expertise and balancing it with prevention of abusive conduct such as brigading, and definitely under revision that aims to introduce thresholds and diminishing returns instead of flat modifiers.

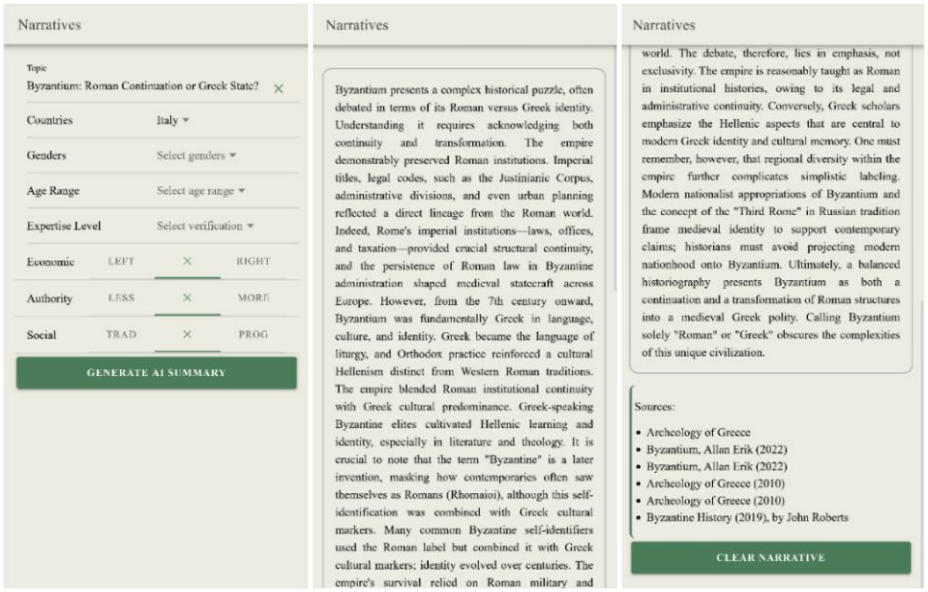


Fig. 5. Narrative filters (a), generated narratives (b) and sources (c)

On the Narratives page, a range of countries and genders can be selected for which perspectives can be viewed. The user can select to view narratives only by country, as well as only by gender, ideology, age or education level. The user can also query for intersections between the presented categories.

When the selection of filters is finished, the user clicks the button and the narrative generation pipeline commences. The system gathers the individual elements from the filter and queries the database for facts which the users belonging to the selected groups marked as true, orders them to return the 100 most voted ones. Figure 5 (a) shows an example of a selection of Italian perspectives about the Byzantine Empire and its culture. Figures 5 (b) and (c) present the generated narrative based on the facts voted by users from Italy, as well as the sources cited in them.

The Gemini prompt underwent a series of revisions, before settling for the following: "*Arrange the information within the following facts into a coherent instructive text. Don't infer any non-obvious information that is not provided within the facts. Do not add interpretations, opinions, or causal connections that are not explicitly stated in the facts. No need to keep the sentence structure the same, make the text stylistically pleasant. Try to arrange the facts chronologically if they are dated*." A list of facts to process is sent. This prompt instructs Gemini to act impartially, intentionally kept minimal to focus the model's capabilities into key directions, instead of relying on its own assumptions. It is perhaps obvious to state that even with an unbiased, encouraging and direct prompt like this, any commercial AI model still refuses to generate certain narratives if they contain obviously malicious historical revisionism, denial of atrocities, and other extreme elements. While this may put a limit the system's possibilities, it can be seen as a built-in layer of moderation to protect users from hate speech and bad faith actors.

4 Evaluation

4.1 Methodology and Results

A small but diverse group of 10 participants, including 8 lay users and 2 trained historians, was observed using the application, both unguided and guided. Unguided sessions lasted 5 minutes, in order to get a feel for how the users would first approach the interface. Guided sessions took 10-15 minutes, and users were tasked to contribute facts, vote, view user profiles, and explore some generated narratives, all of which feedback was used to assess the user interface clarity and the accuracy of generated text, as well as generating new ideas on how to make the application more easily usable and enjoyable with implementing new features. Input of historians was taken into consideration when assessing the validity of this methodology for the field of history as well. Overall, users expressed positive feedback about the app's core use cases and functionalities. However, this section will focus on user complaints and observed areas of difficulty, because the examiner found them more valuable. The most common problems are noted in Table 2, along with the pages on which they appear.

Table 2. Detected issues

Features	Issues	Comments and observations
Narratives page	19	Users reported the most difficulty with understanding this most crucial page of the app. The biggest problems were unintuitive filter design, text being too long or too hard to read, and sources not being more clearly separated and interactive.
Filters	9	
Text	7	
Sources	3	
Facts page	11	While most users were able to navigate the page, some had trouble finding the topics they are interested in and contributing to the facts by a specific topic and area, particularly expressing frustration with topic/area autocomplete.
Feed	6	
Contribution	5	
Profile page	6	Certain users were more concerned with the reputation system and expressed that user search and expertise grading evaluation should be designed better.
Users	3	
Expertise	3	

Other than the issues outlined in Table 2, users reported minor general issues with appearance (page layout, button sizes, navigation sidebar, graph and map sizes on phone screens) and performance of the app (certain loading taking too long, wait time without clear indicators, need for automatic redirection). However, they expressed considerably less interest in having those fixed compared to the main ones they noticed while using the app, so the issues and comments in Table 2 would be a top priority for addressing in any future work.

4.2 Result Analysis

The examinations concluded by users openly discussing suggestions for improvement together with examiner. The main area of interest to focus on would have to be the Narratives page. Trained historians in particular expressed concerns about data transparency and a desire to have every sentence (which may be a fact or an aggregation of multiple ones) be followed up by sources. Users also noted the system should recommend to them which narratives they could realistically select, where a reasonably long text would exist for a particular historical topic. They also expressed that comparison between 2 narratives would be a much more useful tool, as they often wanted to directly access another narrative about the same topic right after. They'd want AI to examine both narratives and clearly outline the most important differences.

Other parts of the app could benefit from improvements mainly to the expertise system, where users wanted to see a way of ensuring no voting manipulation could take place. Some also expressed concerns with allowing such a high advantage to well rated users. Topic and area selection also proved rather troublesome, so potential alleviating solutions could include making it possible for multiple topics to be selected, AI determining the topics automatically (not preferable as it would incur further costs) or making the topic a dynamic attribute subject to voting, just like the accuracy of the fact. Overall, the application was received moderately positively in the testing, but could benefit more from gradual small-scale improvements tackling feature by feature through any future development cycles.

4.3 Quality of Output

Fig. 6. Differing narratives of Greek and Italian users on the issue of Byzantine continuation

Figure 6 shows two outputs for the same selected topic, applying a different nationality filter for each, which resulted in meaningful differences. Here, the topic of contention is the cultural identity of the Byzantine Empire. The narrative generated from facts held true by Greek users emphasizes the dominance of regional Hellenic elites, the Greek liturgical language, the historical combination of Greek cultural markers, and further impact on Eastern Orthodox civilization. On the other hand, the Italian perspective outlines the breadth of economic, scientific and civic traditions and practices, that distinctly hailed from Rome, not Greece, and warns of nationalist attempts to claim Byzantine legacy by the Greeks and other Eastern Orthodox nations. Both of the generated texts can be considered accurate from a Multiperspectivity standpoint, but neither paints a full picture. The value comes from understanding both points of vie: when we look at the areas both interpretations focus on, we can synthesize a well-informed and nuanced answer: that both Greek and Latin civilizations entwine in Eastern Rome, and that this is plainly visible, both in the hovels of the poorest farmers on the fringes, and the purple halls of Constantinople.

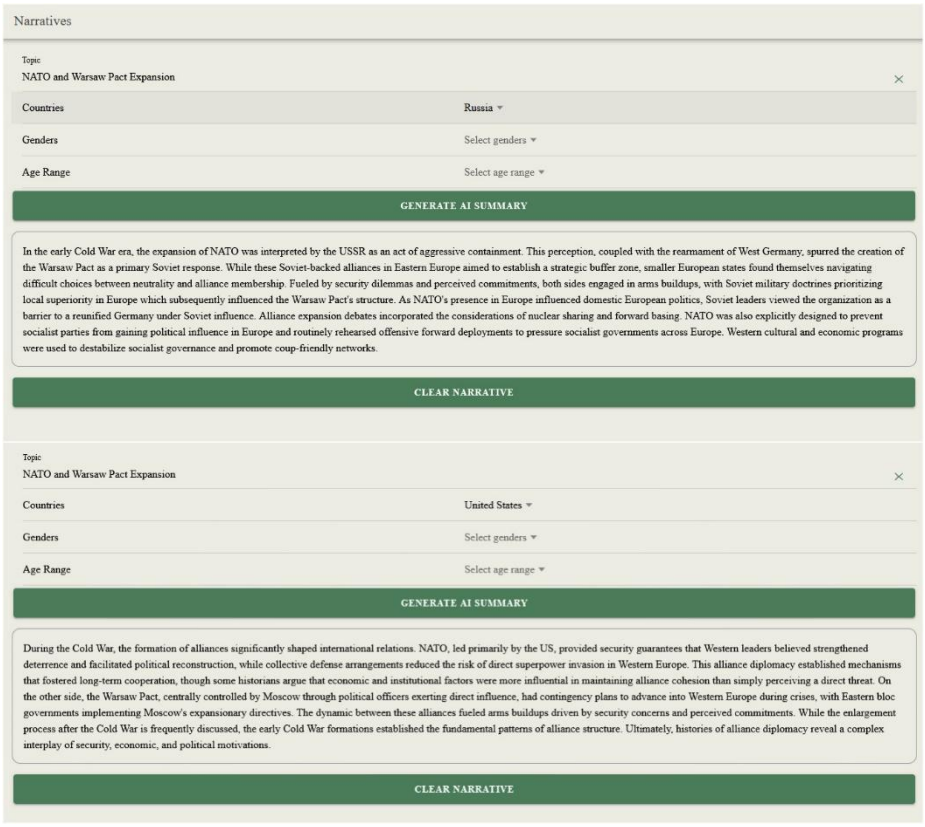


Fig. 7. Differing narratives of Russian and American users on the issue of Cold War Blocs.

Figure 7 is an example from a closer timeframe, showing the difference of opinion on the politics of alliance expansion in the Cold War. The Russian perspective claims that the formation of the Warsaw Pact was a forced response to NATO expansion, and its direct goal of barring German reunification under Soviet influence. Unsurprisingly, the American perspective views this expansion as a necessary security strategy to foster economic cooperation and reconstruction after WW2. This example describes differences that still linger and are harder to reconcile through synthesis like the example from Figure 6. Nevertheless, perhaps through sheer knowledge of what both sides actually believe the user may extract some value out of reading narratives on the more controversial topics.

5 Conclusion

The results of the literature review presented show the potential of the crowdsourcing model and generative AI in responding to the problems and limitations in the ecosystem of digital historiography. Although numerous, the presented challenges can be overcome by achieving a balance between new and old methods, artificial intelligence and expert knowledge (Boroş et al., 2024), and introducing the human element, through crowdsourcing, a well-tested and useful practice successfully applied in social sciences (Kaldeli et al., 2021).

By combining the contributions of lay and professional users, the application provides a way to view history from multiple perspectives and contributes to the inclusiveness and democratization of historiography. It ultimately contributes to overcoming national and cultural bias in understanding history, and raises the public's collective awareness of its diverse interpretations. On the technical side, the implementation of the software system through development frameworks shows how web and mobile applications can support the process of writing and learning history. The integration of the Google Gemini model enables the generation of meaningful narratives based on user-assessed facts. The result is a functional hybrid application that leverages human and artificial intelligence in a way that takes advantage of their complementary strengths while minimizing any identified weaknesses. It is the users who are the creators of the content, they assess the historical accuracy of the facts, aware of the wider context and cultural specificities, while the LLM is the service that processes the large amount of information and derive precise conclusions.

Future development could focus on more direct integration of historical sources and their decomposition into atomic facts, and voting for them. Plans include expanding the mechanisms to prevent bad behavior on the application through moderation and gamification, through the introduction of points and rewards systems, as well as other user connections and elements typical to social networks. Further development is needed in the direction of automated classification of facts by topic and topic by area.

Acknowledgements

This research was supported by the University of Belgrade – Faculty of Organizational Sciences and, in part, by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia through institutional funding (grant number: 200151).

Disclosure of Interests

Authors declare no conflict of interest.

References

1. Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., ... & McGrew, B. (2023). Gpt-4 technical report. *arXiv preprint arXiv:2303.08774.*
2. Alatawi, F., Cheng, L., Tahir, A., Karami, M., Jiang, B., Black, T., & Liu, H. (2021). A survey on echo chambers on social media: Description, detection and mitigation. *arXiv preprint arXiv:2112.05084*.
3. Alkhashil, M. A., Yehia, E. F., & Al-Rashed, S. A. A. (2026). Digital history as participatory communication: platforms, trust, and youth engagement in Saudi Arabia. *Cogent Arts & Humanities*, *13*(1), 2610057.
4. Boroş, E., Ehrmann, M., Romanello, M., Najem-Meyer, S., & Kaplan, F. (2024, March). Post-correction of historical text transcripts with large language models: An exploratory study. In *Proceedings of the 8th joint SIGHUM workshop on computational linguistics for cultural heritage, social sciences, humanities and literature (LaTeCH-CLfL 2024)* (pp. 133-159).
5. Capgemini. (2025). Digitizing history to shape the future through AI with Imperial War Museums. https://www.capgemini.com/news/client-stories/digitizing-history-to-shape-the-future-through-ai-with-imperial-war-museums/ [Accessed: 03.10.2025.]
6. Hauser, J., Kondor, D., Reddish, J., Benam, M., Cioni, E., Villa, F., ... & del Rio-Chanona, R. M. (2024). Large language models' expert-level global history knowledge benchmark (HiST-LLM). *Advances in Neural Information Processing Systems*, 37, 32336-32369.
7. Henriot, C. (2025). The AI-Augmented Research Process: A Historian's Perspective. *arXiv preprint arXiv:2508.01779*.
8. Hutchinson, D. (2026). Benchmarking as Source Criticism: From Recognition to Reasoning in LLM Assessment. *Journal of Open Humanities Data*, *12*(1).
9. Kaldeli, E., Menis-Mastromichalakis, O., Bekiaris, S., Ralli, M., Tzouvaras, V., & Stamou, G. (2021). CrowdHeritage: crowdsourcing for improving the quality of cultural heritage metadata. *Information, 12*(02), 64.
10. Karell, D., Shu, M., Davidson, T., & Okura, K. (2025). Generating the Past: How Artificial Intelligence Summaries of Historical Events Affect Knowledge. *Social Science Computer Review*, 08944393251409744.
11. Klein, L., Martin, M., Brock, A., Antoniak, M., Walsh, M., Johnson, J. M., ... & Mimno, D. (2025). Provocations from the humanities for generative AI research. *arXiv preprint arXiv:2502.19190*.

12. Litaina, T., Soularidis, A., Bouchouras, G., Kotis, K., & Kavakli, E. (2024, August). Towards llm-based semantic analysis of historical legal documents. In *SemDH2024: First International Workshop of Semantic Digital Humanities, co-located with ESWC2024.*
13. Mak, M. K., & Luo, T. (2025). A framework for evaluating cultural bias and historical misconceptions in LLMs outputs. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations*, 100235.
14. Parfenova, A. (2025, April). Unveiling hidden stories: Automated narrative extraction from holocaust diaries with ensemble llms. In *Proceedings of Text2Story—Eighth Workshop on Narrative Extraction From Texts* (Vol. 3964, pp. 081-092).Paricio, J. (2021). Perspective as a threshold concept for learning history. *History Education Research Journal*, *18*(1), 109-125
15. Rahman, S. S., Islam, M. A., Alam, M. M., Zeba, M., Rahman, M. A., Chowa, S. S., ... & Azam, S. (2026). Hallucination to truth: a review of fact-checking and factuality evaluation in large language models. *Artificial Intelligence Review*.
16. Ridge, M. (2016). Making digital history: The impact of digitality on public participation and scholarly practices in historical research. [PhD thesis] *Open University (United Kingdom).*
17. Salnikov, M., Korzh, D., Lazichny, I., Karimov, E., Iudin, A., Oseledets, I., ... & Tutubalina, E. (2025). Geopolitical biases in LLMs: what are the" good" and the" bad" countries according to contemporary language models. *arXiv preprint arXiv:2506.06751*.
18. Santurkar, S., Durmus, E., Ladhak, F., Lee, C., Liang, P., & Hashimoto, T. (2023, July). Whose opinions do language models reflect?. In *International Conference on Machine Learning* (pp. 29971-30004). PMLR
19. Shahmerdanova, R. (2025). Artificial Intelligence in Translation: Challenges and Opportunities. *Acta Globalis Humanitatis et Linguarum*, 2(1), 62-70.
20. Shu, M., Karell, D., Okura, K., & Davidson, T. R. (2026). How latent and prompting biases in AI-generated historical narratives influence opinions. *PNAS nexus*, *5*(3), pgag022.
21. Spina, S. (2023). Artificial Intelligence in archival and historical scholarship workflow: HTS and ChatGPT. *arXiv preprint arXiv:2308.02044*.
22. Time Machine Project. (2023). *About Us*. https://www.timemachine.eu [Accessed: 02.10.2025]
23. Weidinger, L., Uesato, J., Rauh, M., Griffin, C., Huang, P. S., Mellor, J., ... & Gabriel, I. (2022). Taxonomy of risks posed by language models. In *Proceedings of the 2022 ACM conference on fairness, accountability, and transparency* (pp. 214-229).
24. Yao, N. (2025). Integrating AI in historical research: a new method to analyzing national narratives. *Digital Scholarship in the Humanities*, *40*(3), 1033-1040.
25. Zanazanian, P. (2019). Examining historical consciousness through history-as-interpretive-filter templates: Implications for research and education. *Journal of Curriculum Studies, 51*(6), 850-868.
26. Zeng, Y. (2024). HistoLens: An LLM-Powered Framework for Multi-Layered Analysis of Historical Texts--A Case Application of Yantie Lun. *arXiv preprint arXiv:2411.09978*
27. Ziku, M., Kotis, K., Pavlogeorgatos, G., Kavakli, E., Zeeri, C., & Caridakis, G. (2024). Evaluating Crowdsourcing Applications with Map-Based Storytelling Capabilities in Cultural Heritage. *Heritage*, 7(7), 3429-3454.

Original Research Article *https://doi.org/10.5281/zenodo.23475973*

Zero-Knowledge Proofs in E-Commerce: Survey of Applications and a Case Study in Privacy-Preserving Age Verification

Amar Mujezinović[1][0009−0001−9388−2350] and Aleksandar Joksimović[1][0009−0008−5711−7636]

University of Belgrade - Faculty of Organizational Sciences, Belgrade, Serbia
amarmujezinovic1@gmail.com, aleksandar.joksimovic@fon.bg.ac.rs

Abstract. Online commerce increasingly relies on collecting sensitive personal data to satisfy regulatory requirements, a practice that itself creates the privacy risks regulators aim to prevent. Zero-knowledge proofs (ZKPs) offer a way out of this bind: they allow one party to prove knowledge of a fact without revealing anything beyond that fact itself. This paper surveys the main ZKP protocol families, zk-SNARKs, zk-STARKs, and Bulletproofs, and examines where each fits in practical e-commerce scenarios, from KYC and fraud prevention to loyalty programmes and AML reporting. We then develop a detailed case study around privacy-preserving age verification, covering circuit design in Circom, Merkle-tree registry construction, nullifier-based replay prevention, and on-chain verification via smart contracts on both EVM and Solana. We include empirical benchmarks, a security analysis, and a discussion of how the scheme aligns with GDPR, eIDAS 2.0, MiCA, and FATF requirements. This paper contributes an integrative design and benchmark of a ZKP-based age verification architecture for e-commerce deployment, combining per-merchant nullifiers, Merkle registry management, and on-chain verification within a GDPR, eIDAS 2.0, and MiCA-oriented compliance discussion. The key finding is that such a system is technically feasible under the evaluated assumptions, while production deployment still depends on regulatory recognition, governance, and interoperability standards.

Keywords: zero-knowledge proofs · e-commerce · age verification · privacy-preserving protocols · zk-SNARKs · zk-STARKs · GDPR · smart contracts · Merkle trees · nullifiers · KYC · AML.

1 Introduction

Verifying customers is a regular part of online commerce. In many cases a transaction cannot be completed without checking whether the user satisfies a specific requirement. This may include age for restricted products, identity for financial services, residence for jurisdiction-specific rules. In those situations usual way is to collect documents or other personal information and to keep some form of evidence for compliance and audit purposes. Although this approach can satisfy

regulatory requirements, it also introduces additional privacy and security risks. Sensitive data has to be stored, protected, retained, and finally deleted or audited. As a result, even merchants whose primary activity is not identity management may become responsible for data that is valuable and strictly regulated.

Zero-knowledge proofs (ZKPs) provide a different approach. This concept was first introduced by Goldwasser, Micali and Rackoff in 1985 [14], and suggest a proof system in which a prover can convince a verifier that a statement is true without revealing the secret. For a long time this approach was mostly theoretically discussed. This changed with the development of practical proving systems, starting with Pinocchio [22] and continuing with Groth16 [16], PlonK [12], and the STARK family [3, 4]. Today, ZKP systems can take a high-level description of a computation and produce proofs that can be verified efficiently. Bitcoin showed that Merkle trees and cryptographic commitments can be used as part of a public verification mechanism in distributed systems [20]. ZKPs apply related cryptographic ideas to a different problem: verifying claims about users without exposing the data on which those claims are based.

Despite the maturity of ZKP primitives, their application to e-commerce verification workflows has not been systematically addressed in prior work. Existing deployments such as Semaphore and Polygon ID focus on general-purpose group membership and identity layers respectively, without addressing the specific requirements of merchant-facing age verification: per-merchant unlinkability, replay prevention across independent retail contexts, and simultaneous alignment with GDPR, eIDAS 2.0, and MiCA. zkPass targets proof of web2 data attributes but does not provide a deployable architecture for regulated age-gating in retail. Zcash-inspired nullifier constructions [18] address payment privacy rather than eligibility verification. This paper addresses that gap by providing an application-specific integration of these primitives for e-commerce age verification, supported by empirical benchmarks and a regulatory analysis covering the current EU framework.

E-commerce is a relevant context for this technology because many online transactions already require some form of user verification. In many cases, the merchant only needs to know whether a certain condition is satisfied, such as whether the user is old enough, belongs to an allowed jurisdiction, or has completed a KYC process. For this reason, ZKPs can be useful in scenarios where compliance requirements have to be met without exposing more personal data than necessary. This is also relevant in the context of eIDAS 2.0, MiCA, and national age-verification rules, which increase the need for reliable and privacy-preserving verification mechanisms.

This work is related to the broader field of privacy-enhancing technologies, such as anonymisation, pseudonymisation, differential privacy, and homomorphic encryption. These approaches reduce privacy risks in different ways, but ZKPs are especially relevant when the main task is verification without disclosure. In a regulatory context, this is useful because the verifier often needs only the answer to a specific question, for example whether the user satisfies condition X, rather than access to the underlying personal data.

The focus of this paper is on the use of ZKPs in e-commerce, with age verification as the main case study. Section 2 gives an overview of the ZKP foundations and the protocol families most relevant for practical applications, including zk-SNARKs, zk-STARKs, and Bulletproofs. Section 3 discusses selected e-commerce use cases, such as KYC, fraud prevention, loyalty programs, and AML-related reporting. Section 4 presents an architecture for privacy-preserving age verification, including circuit design, registry management, nullifier-based replay prevention, and smart contract integration. Sections 5 and 6 consider the proposed approach from the perspective of performance, security, and regulatory alignment. Section 7 discusses open problems, and Section 8 concludes the paper.

2 Foundations of Zero-Knowledge Proofs

2.1 What Zero-Knowledge Actually Means

A zero-knowledge proof system involves two parties: a prover, who possesses some private information, and a verifier, who wants to be convinced that a certain claim about that information is true. Such a system is usually defined through three core properties. Completeness means that, if the statement is true and the prover follows the protocol honestly, the verifier will accept the proof. Soundness means that, if the statement is false, no dishonest prover should be able to convince the verifier except with negligible probability. Zero-knowledge means that the verifier learns nothing beyond the validity of the statement itself. More formally, there must exist a polynomial-time simulator that can generate transcripts indistinguishable from real protocol executions without knowing the underlying secret.

For practical use, two extensions are especially important. First, many real systems require non-interactive proofs, where the back-and-forth communication between the prover and verifier is replaced by a single proof message. This is usually achieved using a common reference string or by applying the Fiat–Shamir heuristic in the random oracle model. Non-interactivity is particularly important in blockchain and e-commerce systems, where the verifier cannot be expected to engage in a live protocol with every user. Second, many proof systems are also arguments of knowledge. This means that, if a prover can produce an accepting proof, then one can in principle extract a valid witness for the proved statement. This property prevents trivial or contrived strategies in which a prover convinces the verifier without actually possessing the relevant secret.

2.2 Protocol Families

zk-SNARKs zk-SNARKs are among the most widely discussed zero-knowledge proof systems because they provide very short proofs and fast verification. In many deployed constructions, the proof size is constant with respect to the size of the computation being proved. For example, Groth16 proofs are typically only a few hundred bytes, often around 192 bytes on pairing-friendly curves. Verification

requires only a small fixed number of elliptic-curve pairing checks, which makes SNARKs especially attractive for blockchain applications, where on-chain storage and computation are expensive.

The main drawback is the trusted setup. Some SNARK constructions require a structured reference string that is generated before proofs can be produced. If the secret randomness used during this setup, often called "toxic waste", is not destroyed, an attacker could forge valid-looking proofs. Multi-party computation ceremonies reduce this risk by distributing the setup process among many participants. The security intuition is that, if at least one participant behaves honestly and deletes their secret contribution, the final parameters remain safe.

Groth16 [16] remains one of the most efficient constructions in terms of proof size and verification time, but it requires a circuit-specific trusted setup. More recent universal and updatable SNARKs, such as PlonK [12], avoid a separate trusted setup for every circuit. They usually pay for this flexibility with somewhat larger proofs or higher prover costs, but the trade-off is often worthwhile when a system must support many different circuits.

For an e-commerce system, this means that a proof for an age-verification circuit can remain small even if the underlying circuit contains thousands or tens of thousands of constraints. The proof size does not grow linearly with the circuit size; instead, the main cost is shifted to proof generation. This is useful when the verifier is a merchant, a smart contract, or a payment platform that needs quick verification and predictable communication costs.

zk-STARKs zk-STARKs provide a different set of trade-offs. Unlike many SNARK constructions, STARKs do not require a trusted setup. They rely on collision-resistant hash functions and polynomial commitment techniques based on FRI [3, 4]. This makes them transparent: there is no ceremony, no toxic waste, and no need to trust that setup participants behaved honestly. Their security assumptions are also considered post-quantum, because they do not rely on elliptic-curve discrete logarithms in the same way as pairing-based SNARKs.

The main disadvantage of STARKs is proof size. While SNARK proofs are often only a few hundred bytes, STARK proofs are commonly measured in tens or hundreds of kilobytes. Verification is still efficient, but it grows with the size of the computation rather than remaining completely constant. This makes STARKs particularly suitable for large computations, high-throughput systems, and settings where transparency and post-quantum security are more important than minimal proof size.

Bulletproofs Bulletproofs [7] are another important family of zero-knowledge proofs. They do not require a trusted setup, and their proof size grows logarithmically with the number of constraints. This makes them especially attractive for range proofs, where the prover wants to show that a hidden value lies within a certain interval without revealing the value itself.

Their main limitation is verification time, which is linear in the size of the statement. For this reason, Bulletproofs are not usually the best option for large

general-purpose circuits. However, they are well suited for structured claims such as proving that a transaction amount is below a regulatory threshold, or that a private score exceeds a required minimum.

2.3 Building Blocks

Merkle Trees A Merkle tree allows a large dataset to be committed to a single root hash. Each leaf contains the hash of a data item, while each internal node contains the hash of its two children:

$$h_{\text{parent}} = \text{Hash}\left(h_{\text{left}} \parallel h_{\text{right}}\right).$$

To prove that an item d_i belongs to the committed dataset, the prover provides the sibling hashes along the path from the corresponding leaf to the root. The verifier recomputes the root and checks whether it matches the published commitment. Since only one sibling hash is needed at each tree level, the proof size is $O(\log n)$. Merkle trees are useful in zero-knowledge systems because they allow a prover to show that a credential or attribute exists in a committed set without revealing the entire set. However, the choice of hash function is crucial inside ZKP circuits. Traditional hash functions such as SHA-256 are expensive to express as arithmetic constraints. Hash functions designed for prime-field arithmetic, such as Poseidon [15], are significantly more efficient in this context.

Nullifiers A nullifier is a deterministic one-way value derived from a user's private state and a specific context:

$$n = \text{Hash}\left(sk, \text{context}\right).$$

Its purpose is to mark a credential or private note as used without revealing which credential was used. The verifier maintains a registry of previously seen nullifiers and rejects any proof that contains a duplicate nullifier. Since the nullifier is computed using a one-way function, it should not reveal the user's secret key or identity.

This idea was popularised in privacy-preserving cryptocurrencies such as Zerocoin and Zcash [18], where nullifiers prevent double spending of private notes. In e-commerce, the same mechanism can prevent replay attacks. For example, a user may prove that she is over a required age threshold without revealing her exact date of birth. A nullifier tied to a merchant, session, or transaction context can ensure that the same proof cannot simply be copied and reused elsewhere.

3 ZKP Applications in E-Commerce

3.1 Identity Verification and KYC

KYC requirements under European Union anti-money-laundering rules require service providers to verify users before accepting them. The traditional approach

involves collecting documents, checking sanctions lists, and storing identity evidence, which creates significant privacy, compliance, and security risks. A security incident at a KYC provider may expose identity documents at scale.

The zkKYC pattern reduces this exposure by allowing a certified identity provider to verify a document once and publish a commitment, such as a Merkle root, to a registry. The user can then generate a zero-knowledge proof showing that their committed attributes satisfy a merchant's conditions, such as being over 18 or not appearing on a sanctions list. The merchant receives only the verification result and a nullifier, while documents, names, dates of birth, and other identity attributes remain hidden. This model still requires trusted and audited identity providers, but it limits disclosure to relying parties.

3.2 Age Verification

Age verification is one of the most practically relevant applications of zero-knowledge proofs in e-commerce. The UK Online Safety Act, the EU Digital Services Act, and a growing number of US state laws are creating new legal obligations around age-based content restrictions. Existing solutions involve important trade-offs. Self-declaration of age is easily bypassed, credit card proxy verification excludes users who do not have a card, and uploading documents to a third party may create additional GDPR compliance obligations while increasing the impact of potential security incidents. We develop a complete alternative based on zero-knowledge proofs in Section 4.

From an economic point of view, the costs of the current situation are very high. Age verification providers charge fees per verification that can range from 0.20 euros to over 1.00 euro, depending on the verification method and the country the user is in. For platforms with high volumes, this is quite a large expense. A system based on zero-knowledge proofs changes the cost structure: the AVS charges once at registration, and subsequent verifications may have substantially lower marginal costs, apart from proof verification and infrastructure costs. For high-volume platforms, this may make ZKP-based verification economically attractive, provided that onboarding, wallet integration, and operational costs are controlled.

3.3 Payment Fraud Prevention

Online payment fraud remains a significant economic and operational challenge. Fraud detection systems use the transaction history of users, which is personal data, and this creates a conflict with GDPR. The approach based on zero-knowledge proofs works by having the bank create a model of normal behavior for each user. When the user pays, their device only proves that the transaction matches that model, without sending the full history. Bulletproofs are suitable for this type of range-oriented verification because they avoid a trusted setup and provide compact proofs for structured statements. An important thing is that the model itself must also be protected. If an attacker sees what the model looks like, they can create a fake transaction that passes the check. That is why the model

is kept in a protected form and the bank signs it before each check. This makes the approach a bit more complex, but current ZKP tooling can support such constructions, although implementation complexity and model confidentiality must be carefully addressed.

3.4 Loyalty Programmes

Loyalty profiles are among the most detailed and most valuable datasets in retail. Because of that they are also more sensitive under GDPR than other data. A loyalty architecture based on zero-knowledge proofs allows a user to prove membership in a group ("has spent over €500 euros in category X in the past 90 days") without revealing individual purchases. Multi-merchant loyalty programs benefit especially from this: partner merchants can verify that a user qualifies for a shared benefit without exchanging purchase histories with each other and without exposing sensitive customer data.

The technical challenge is that loyalty conditions are often dynamic, segments change as purchase history grows. Efficient zero-knowledge proofs over frequently updated data require either re-proving everything from scratch at each update, which would be very expensive, or gradually updating the proof without needing to recalculate everything from the beginning. For most loyalty applications, it is enough for the user to prove membership when they want to use a benefit, not at every moment while they are in the app. The data needed for the proof can be updated in the background between purchases.

3.5 AML Reporting

AML regulations require that transactions above 10,000 euros be reported, along with information about the sender and receiver of the transaction. Zero-knowledge proofs make the following possible: the bank proves to the regulator that no transaction has exceeded the threshold, without revealing data about individual transactions. FATF guidelines for virtual assets [11] already anticipate this kind of approach, and this is a concrete example of how it can look in practice. STARKs are a suitable candidate for this setting because they are designed for large computations and do not require a trusted setup. The regulator could verify a compact proof instead of inspecting each transaction individually, depending on the concrete implementation and reporting requirements. Compared with bulk data reporting, such an approach could reduce unnecessary data exposure and focus regulatory review on verifiable compliance claims.

3.6 Summary

Table 1 maps protocol families to use cases based on proof size, verification speed, setup requirements, and computational structure. Groth16 is used when small proofs and fast verification are most important; when the requirement is post-quantum security then STARKs are preferred; Bulletproofs work well in scenarios that involve a lot of range proofs and strict speed requirements.

Table 1. Protocol selection guide for e-commerce use cases.

Use Case	Preferred Protocol	Proof Size	Trusted Setup	Latency Req.
KYC / Identity	Groth16 / PlonK	192 – 800 B	Yes (MPC)	< 5 s (reg.)
Age Verification	Groth16	192 B	Yes (MPC)	< 5 s (reg.)
Payment Fraud	Bulletproofs	1 – 2 KB	No	< 500 ms
Loyalty / Segments	Groth16 / PlonK	192 – 800 B	Yes (MPC)	< 5 s
AML Reporting	STARK / PlonK	40 – 100 KB	No / Universal	< 60 s (batch)

4 Case Study: Privacy-Preserving Age Verification

4.1 Why Existing Solutions Fall Short

Age-restricted goods and services cover a wide range, for example alcohol, tobacco, gambling, adult content, and certain medicines. Age limits vary by country (usually 18 in the EU, while in many US states the limit for alcohol is 21). However, the verification problem is structurally the same everywhere and the existing solutions share the same weaknesses.

Self-declaration of age (clicking "I confirm I am over 18") is easily bypassed and is generally not considered a sufficient check. Uploading a document to a third party is the most common "serious" solution and at the same time the most problematic one: the user sends a government-issued ID or passport to some company that stores it, which creates exactly the type of data liability that GDPR was meant to prevent. Several major age verification providers have already had security breaches and hacks. On top of that, the user has to go through the whole verification process again with every new merchant, which is impractical.

A ZKP-based approach can address these limitations by separating eligibility verification from disclosure of the underlying identity data. Specifically, the system should satisfy four conditions: R1: the merchant receives confirmation from a trusted provider that the user meets the age requirement; R2: no personal data reaches the merchant, only the result of the check; R3: the same user generates a different identifier at each merchant, limiting cross-site tracking; and R4: the provider that issues confirmations remains accountable to regulators without routinely disclosing data about individual users.

4.2 Architecture and Protocol

Three parties are involved in the system: the User, the Age Verification Service (AVS), and the Merchant. Two on-chain contracts provide trustless infrastructure: the Merkle Root Registry Contract (MRRC) publishes the current AVS registry root, and the Verifier Contract checks proofs and maintains the nullifier registry.

Registration happens once, off-channel. The user presents identity documents to the AVS through a regulated channel. The AVS verifies the document, extracts the date of birth and document identifier, checks sanctions lists, and computes a leaf commitment: leaf = Poseidon(Hash(DoB), Hash(doc_id), salt_u). Here

salt_u is a random value generated client-side and returned to the user—the AVS never stores it. The leaf is inserted into an incremental Merkle tree of depth $d = 20$, giving capacity for roughly one million registered users. The updated root R is published on-chain via the MRRC. The AVS sends the user their Merkle path and salt_u; after that, the AVS retains only the leaf hash. No document image, no date of birth, no salt.

Verification is triggered by each purchase of age-restricted goods. The user's client generates a ZKP from the stored witness bundle and submits it to the Verifier Contract along with a nullifier and the current root. The contract checks three things: R matches the MRRC's current root, the proof verifies against the registered verification key, and the nullifier has not been seen before. If all three pass, a VerificationPassed event is emitted and the merchant grants access. The full protocol is shown in Fig. 1.

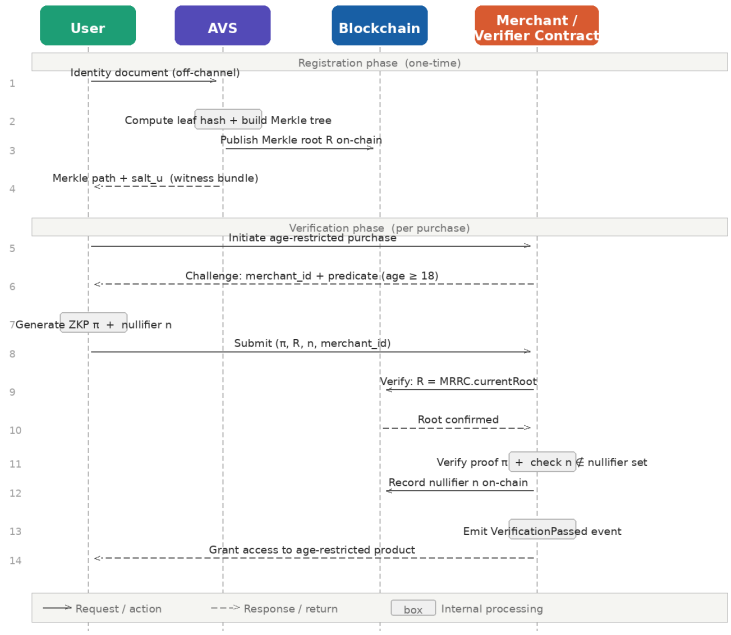


Fig. 1. Sequence diagram of the ZKP-based age verification protocol. Solid arrows denote requests; dashed arrows denote responses. Boxes on lifelines indicate local computation.

4.3 Circuit Design

The circuit is written in Circom 2.0 [2] and compiled to R1CS for proving with Groth16 over BN254 via SnarkJS. Private inputs are: merkle_path (20 sibling

hashes), leaf_preimage (DoB year, month, day), salt_u, and private_key (a BabyJubJub scalar). Public inputs are: R (the Merkle root), n (the nullifier), the predicate threshold year, the current date components, and merchant_id.

The circuit has three constraint groups. The Membership group (˜18,000 constraints) recomputes the leaf hash from leaf_preimage and salt_u via Poseidon, then walks up the tree for d = 20 levels, at each step hashing the node with the appropriate sibling based on a private path-direction bit. The final hash must equal R. This is the core of the membership proof: if this constraint is satisfied, the user's date of birth genuinely appears in the AVS's registry.

The Age Predicate group (˜6,000 constraints) compares the committed date of birth against the threshold. The comparison uses 64-bit binary decomposition to implement: (current_year × 10000 + current_month × 100 + current_day) - (DoB_year × 10000 + DoB_month × 100 + DoB_day) ≥ threshold_years × 10000. The only date-related public value is the predicate threshold year—no absolute birth date appears in the public inputs.

The Nullifier group (˜4,000 constraints) enforces n = Poseidon(private_key, merchant_id). The private_key here is the BabyJubJub key whose public counterpart was embedded in the leaf at registration, binding the nullifier to both the specific credential and the specific merchant. Total constraint count is approximately 28,000, comfortably within the 2^{15} boundary of publicly available Powers-of-Tau ceremonies.

4.4 Smart Contract Integration

The MRRC is a minimal contract with a single updateRoot function that is only available to the AVS multisig. It keeps a timestamped history of roots so that proofs generated against a recently replaced root stay valid for a configurable grace period (we recommend 48 hours).

The Verifier Contract is generated directly from the Groth16 proving key using the SnarkJS exportSolidityVerifier function. It exposes verify(uint[2] a, uint[2][2] b, uint[2] c, uint[] input) → bool, which runs two BN254 pairings through Ethereum precompile functions at addresses 0x06 to 0x08. For Solana, the equivalent BPF program calls alt_bn128_pairing.

The nullifier registry is a mapping from bytes32 → bool. On each call to the verify function, the contract checks require(!nullifiers[n]) before running the pairing computation, then sets nullifiers[n] to true on success. The merchant backend listens for VerificationPassed events and grants access when it receives one. Everything runs as a single unit, either everything succeeds and the state is written, or the transaction is cancelled.

5 Performance and Security Evaluation

5.1 Benchmarks

Proof generation time was evaluated in three different systems, namely, laptop (Intel Core i7-12700H, 32 GB RAM, Ubuntu 22.04), mid-end smartphone (Pixel

7), and server (AMD EPYC 7543, 256 GB RAM). In each of these three systems, we have tried two different ways to generate proof: SnarkJS version 0.7 in WASM environment using Node.js and rapidsnark version 1.0 as a native prover. The result is given in Table 2.

Table 2. Proof generation and verification benchmarks ($d = 20$, approximately 28,000 constraints).

Environment	SnarkJS (s)	rapidsnark (s)	Verification (ms)
Laptop (i7-12700H)	2.1	0.18	0.7
Smartphone (Pixel 7)	4.3	0.62	N/A (on-chain)
Server (EPYC 7543)	0.9	0.07	0.5
On-chain verify (EVM)	–	–	~230,000 gas
On-chain verify (Solana)	–	–	<0.001 SOL
Proof size / Witness bundle	3.2 KB	–	–

The 2.1 s WASM execution time on the laptop is acceptable for a registration flow, where a short delay can be expected, but it may be marginal for checkout-time verification. By contrast, the 180 ms rapidsnark execution time on the laptop and the 620 ms execution time on the smartphone indicate that native proving can support interactive purchase flows. On Ethereum, verification requires approximately 230,000 gas, corresponding to roughly €0.04–€0.12 under late-2024 price assumptions, while the Solana verification cost remains below 0.001 SOL. These results suggest that proof generation and verification costs are feasible, although deployment economics depend on network fees, wallet integration, and operational infrastructure.

5.2 Security

The scheme rests on four security assumptions. Poseidon collision resistance [15] prevents leaf fabrication. BN254 discrete-log hardness prevents proof forgery. Trusted setup integrity prevents toxic waste attacks, which is mitigated by running an MPC ceremony with public transcripts and at least one honest participant. AVS integrity ensures that registered leaves correspond to real verified identities, and this relies on the AVS being a regulated entity subject to audit, not just on cryptography.

The nullifier design handles three attack types cleanly. Replay at the same merchant is blocked by the on-chain nullifier registry. Cross-merchant replay is structurally impossible because the nullifier encodes the merchant ID, so a proof generated for merchant A has a nullifier that is mathematically unrelated to the one for merchant B, even for the same user. Sybil registration, where one person gets multiple leaves, requires either forging documents or corrupting the AVS, and neither of these is a ZKP level problem.

One less obvious risk is timing correlation: an observer watching on-chain nullifier submissions might figure out that a particular blockchain address made

a purchase at a particular time, even without knowing who the user is. Sending the proof through the merchant backend instead of directly from the user wallet breaks this connection, and so does using a privacy preserving transaction relay.

The security model here is layered. The ZKP layer handles proof correctness and zero knowledge. The nullifier layer handles replay prevention. The AVS layer handles real world identity integrity. The on-chain layer handles finality and public verifiability. Each layer has its own trust assumptions and a failure in one does not necessarily affect the others. A corrupt AVS cannot break ZKP soundness and a broken hash function would not immediately compromise the on-chain contract logic. This layering is a deliberate feature of the architecture, rather than an accidental property.

6 Regulatory Alignment

6.1 GDPR

GDPR's data minimisation principle in Article 5(1)(c) [9] requires personal data to be adequate, relevant, and limited to what is necessary for the purpose of processing. The proposed scheme addresses this requirement at the architectural level: the merchant receives only a cryptographic proof and a nullifier, not the user's identity document, date of birth, or underlying attributes. The nullifier is a deterministic pseudorandom value used to prevent replay or double use, rather than to identify the user.

The AVS stores commitments or Merkle leaf hashes instead of plaintext identity attributes after registration. Whether such values fall outside the scope of personal data depends on whether the user can still be identified by means reasonably likely to be used. For this reason, the scheme should use high-entropy salts, for example 128-bit salts, and ensure that stored commitments cannot be linked back to the original attributes without the user's private input. This supports Article 25's privacy-by-design requirement structurally rather than only through organisational policy.

Article 35 GDPR requires a Data Protection Impact Assessment where processing is likely to create high risk, especially in large-scale identity-related processing. An AVS would still need to assess the registration stage, where identity evidence may be processed. However, compared with a traditional document-upload provider, the residual operational dataset is narrower: subsequent checks rely on commitments, Merkle membership proofs, and nullifiers rather than repeated disclosure of identity documents to merchants. This can simplify the DPIA and reduce the merchant's compliance burden, although it does not remove all GDPR obligations from the system.

There is also a question of legal basis. Traditional age verification may rely on Article 6(1)(c), where verification is required by law, or Article 6(1)(f), where legitimate interest is appropriate and properly balanced. Under the proposed scheme, the AVS processes identity information during registration, while later merchant-side verification can be performed without the merchant receiving the

underlying identity data. This reduces the amount of data the merchant must justify, secure, retain, or disclose in the event of an incident.

6.2 eIDAS 2.0

The revised eIDAS framework requires European Digital Identity Wallets to be made available, with wallet availability expected by the end of 2026. The EUDI Wallet Architecture Reference Framework [8] supports selective disclosure of attributes through formats such as SD-JWT and mdoc. ZKP-based selective disclosure is not identical to these mechanisms, but it follows the same architectural objective: proving a required attribute without revealing unnecessary additional data.

In eIDAS terminology, the wallet or wallet-controlled component can serve as the user's presentation mechanism, the identity or attribute provider can act as the AVS or credential issuer, and the merchant or Verifier Contract can act as the relying-party verifier. This mapping is conceptually compatible with the eIDAS 2.0 direction, although concrete deployment would depend on wallet standards, certification rules, and national implementations.

One important open question is whether a ZKP-based age proof could qualify as, or be derived from, a Qualified Electronic Attestation of Attributes (QEAA). QEAAs carry a high assurance level and benefit from legal recognition across Member States. If an age attribute is issued by a recognised or qualified provider and presented through an EUDI Wallet, a ZKP-based presentation could in principle provide a privacy-preserving proof of eligibility. However, this remains a regulatory and certification question rather than a purely technical one.

6.3 MiCA and FATF

MiCA [10] establishes a harmonised EU framework for crypto-asset markets and crypto-asset service providers. A zkKYC architecture can support regulated compliance workflows if the AVS or credential issuer is an appropriately authorised and supervised identity provider. The ZKP layer reduces disclosure during verification, but it does not remove the need for a legally recognised identity-verification process at onboarding.

This is consistent with FATF's risk-based approach to virtual assets and virtual asset service providers [11]. Digital identity systems may support customer due diligence when they provide appropriate assurance and governance. ZKP-based attribute proofs can therefore be understood as a privacy-preserving mechanism for proving compliance-relevant facts, such as successful KYC completion, age eligibility, or jurisdictional status, without disclosing the full identity record to every relying party. The Basel Committee's crypto-asset prudential framework [1] reinforces the broader regulatory expectation of assurance, auditability, and risk control.

6.4 What Is Still Unresolved

Two issues remain unresolved. First, the legal status of a ZKP as evidence of verification has not been fully tested. Some age-verification regimes may require merchants to retain evidence that a check was performed. A cryptographic proof can provide strong technical evidence, especially if linked to a timestamp, verifier context, and issuer commitment, but whether this is legally sufficient depends on national rules and regulatory interpretation. Regulatory sandbox engagement before production deployment would therefore be prudent.

Second, liability allocation remains unclear. If a user fraudulently obtains a credential from an AVS and later uses it with a merchant, it is not obvious whether liability should fall on the AVS, the merchant, the credential issuer, or another party. This is similar to liability allocation in payment-card schemes, where technical verification must be supported by contractual rules. ZKP-based systems will therefore need clear contractual arrangements, audit procedures, and sector-specific standards, not only cryptography.

7 Discussion

The proposed system is technically feasible under the evaluated assumptions, but several deployment challenges remain. The implemented circuit, benchmark results, and regulatory analysis indicate that ZKP-based age verification can be practically implemented; however, production deployment depends on mobile performance, interoperability standards, legal recognition, and governance arrangements. There are three issues that need to be addressed.

The latency experienced in mobile devices is the most noticeable technical challenge. The 4.3 s proof-generation time on the Pixel 7 may be acceptable during account creation, but it is likely too high for checkout-time verification. There are three solutions to this issue. First, the data required for proof generation can be precomputed in the background by the user before making the payment. Secondly, rapidsnark can be integrated as a native library that reduces down proof generation time to 620ms. Lastly, there are recursive proof creation techniques such as Nova and Halo2 [5] that enable the gradual updating of proofs without redoing all the work.

The second challenge is one of standardization. At present each AVS uses a different circuit and API. An e-commerce company trying to integrate with three different providers will have to create three separate integrations. The W3C Verifiable Credentials framework provides a plausible interoperability layer. The EUDI wallet ARF offers a legal means to do so within the EU context. Addressing this challenge could improve interoperability and reduce integration costs across providers.

The third challenge is that of selective disclosure of data to regulators, assuming that there exists a dedicated communication channel to regulators directly from users. One possible direction is to combine ZKP-based verification with threshold cryptography for controlled regulatory disclosure. User identity can only

be determined if most of the regulatory nodes work together in collusion, providing cryptographically imposed constraints on the regulatory process rather than relying solely on procedural safeguards. This ties in with efforts in accountable anonymity research.

Going beyond age verification, the very same architecture – Merkle registry, ZKP circuit, nullifier, and on-chain verifier – can be used for verifying the existence of a condition of the binary type for a registered attribute. Income status for lending eligibility. Certification status for regulatory compliance of a service. Medical fitness for telemedicine. Residence for geographic content licensing. The circuit will be different in each case, but the registry management, nullifier setup, and on-chain verifier can be reused for all of them. The ability to reuse such infrastructure justifies its creation even when starting out with a narrow use case. The cost of implementing an additional condition is about the cost of creating and auditing another Circom circuit.

There is also a chance of merging ZKP attribute proofs with decentralized identity standards in general. The W3C Decentralized Identifier and Verifiable Credentials specifications describe a user-controlled identity that is becoming increasingly popular outside the realm of blockchain technologies as well. By implementing ZKP selective disclosure as a feature within this VC ecosystem rather than a technology specifically for blockchains, the potential markets will be greatly expanded. This would also make the approach more accessible to merchants that do not otherwise rely on blockchain infrastructure.

8 Conclusion

This paper examined whether zero-knowledge proofs can be used for age verification in e-commerce while reducing the data exposure associated with current verification solutions. Based on the proposed design and benchmark results, the answer is technically positive under the assumptions considered in this work. The Groth16/Circom circuit, combined with Poseidon hashing, a depth-20 Merkle registry, per-merchant nullifiers, and on-chain verification, addresses the four design criteria defined above: data minimization, unlinkability, replay resistance, and AVS accountability.

The benchmark results indicate that the approach is practical for the evaluated deployment scenario. Native proof generation with rapidsnark takes less than one second on the tested devices, while on-chain verification remains within a feasible cost range, with verification below €0.15 on Ethereum and substantially lower costs on Solana. From a regulatory perspective, the protocol is technically aligned with GDPR principles of data minimization and privacy protection, and it follows the selective-disclosure direction promoted by eIDAS 2.0. In MiCA- and AML-related contexts, such mechanisms could support compliance workflows if combined with legally recognized identity providers, appropriate governance, and regulator-approved procedures.

The remaining challenges are therefore not purely technical. They also include regulatory recognition, liability allocation for AVS providers, interoperability with

verifiable credential standards, and operational governance. Regulators would need to clarify under which conditions ZKP-based attestations can constitute sufficient evidence of validation, while industry actors would need to define accountability frameworks for providers issuing such credentials. Standardization efforts, including ZK-compatible verifiable credential profiles, will also be important for broader adoption. These challenges are substantial, but ongoing work on digital identity wallets, verifiable credentials, and ZKP infrastructure provides a realistic basis for further development. Organizations that invest early in ZKP-based identity architectures and standards participation may therefore be better positioned as regulatory and technical frameworks mature.

Acknowledgments. Gemini was used only to help with proofreading, grammar correction, and minor style improvements of text written by the authors. It was not used for choosing the literature, designing the research, performing the technical analysis, interpreting the benchmark results, or writing the conclusions.

Disclosure of Interests. The authors declare that they have no conflicts of interest regarding this study's content.

References

1. Basel Committee on Banking Supervision: Prudential treatment of cryptoasset exposures. Bank for International Settlements (2023)
2. Belles-Munoz, M., Isabel, M., Munoz-Tapia, J.L., Rubio, A., Baylina, J.: Circom: A circuit description language for building zero-knowledge applications. IEEE Transactions on Dependable and Secure Computing **20**(6), 4733–4751 (2023). https://doi.org/10.1109/TDSC.2022.3232813
3. Ben-Sasson, E., Bentov, I., Horesh, Y., Riabzev, M.: Scalable zero knowledge with no trusted setup. In: Advances in Cryptology – CRYPTO 2018. Springer (2018)
4. Ben-Sasson, E., Chiesa, A., Tromer, E., Virza, M.: Succinct non-interactive zero knowledge for a von neumann architecture. In: Proceedings of the 23rd USENIX Security Symposium. pp. 781–796. USENIX Association (2014)
5. Boneh, D., Drake, J., Fisch, B., Gabizon, A.: Halo infinite: Proof-carrying data from additive polynomial commitments. In: Advances in Cryptology – CRYPTO 2021. Springer (2021)
6. Boneh, D., Shoup, V.: A Graduate Course in Applied Cryptography. Self-published online manuscript (2020), `https://crypto.stanford.edu/~dabo/cryptobook/`
7. Bunz, B., Bootle, J., Boneh, D., Poelstra, A., Wuille, P., Maxwell, G.: Bulletproofs: Short proofs for confidential transactions and more. In: Proceedings of the 2018 IEEE Symposium on Security and Privacy. pp. 315–334. IEEE (2018). https://doi.org/10.1109/SP.2018.00020
8. European Commission: Architecture and reference framework for the european digital identity wallet. Tech. rep., European Commission (2024), version 1.4.0
9. European Parliament and Council: Regulation (EU) 2016/679: General data protection regulation (GDPR). Official Journal of the European Union (2016)
10. European Parliament and Council: Regulation (EU) 2023/1114 on markets in crypto-assets (MiCA). Official Journal of the European Union (2023)

11. Financial Action Task Force: Guidance for a risk-based approach to virtual assets and virtual asset service providers. Tech. rep., FATF (2019)
12. Gabizon, A., Williamson, Z.J., Ciobotaru, O.: Plonk: Permutations over lagrange-bases for oecumenical noninteractive arguments of knowledge. IACR ePrint 2019/953 (2019)
13. Gennaro, R., Gentry, C., Parno, B., Raykova, M.: Quadratic span programs and succinct nizks without pcps. In: Advances in Cryptology – EUROCRYPT 2013. Springer (2013)
14. Goldwasser, S., Micali, S., Rackoff, C.: The knowledge complexity of interactive proof systems. SIAM Journal on Computing **18**(1), 186–208 (1989)
15. Grassi, L., Khovratovich, D., Rechberger, C., Roy, A., Schofnegger, M.: Poseidon: A new hash function for zero-knowledge proof systems. In: Proceedings of the 30th USENIX Security Symposium. pp. 519–535. USENIX Association (2021)
16. Groth, J.: On the size of pairing-based non-interactive arguments. In: Advances in Cryptology – EUROCRYPT 2016. Springer (2016)
17. Kate, A., Zaverucha, G.M., Goldberg, I.: Constant-size commitments to polynomials and their applications. In: Advances in Cryptology – ASIACRYPT 2010. Springer (2010)
18. Miers, I., Garman, C., Green, M., Rubin, A.D.: Zerocoin: Anonymous distributed e-cash from bitcoin. In: Proceedings of the 2013 IEEE Symposium on Security and Privacy. pp. 397–411. IEEE (2013). https://doi.org/10.1109/SP.2013.34
19. Moser, M., Bohme, R., Breuker, D.: An inquiry into money laundering tools in the bitcoin ecosystem. In: Proceedings of the 2013 APWG eCrime Researchers Summit. pp. 1–14. IEEE (2013). https://doi.org/10.1109/eCRS.2013.6805780
20. Nakamoto, S.: Bitcoin: A peer-to-peer electronic cash system (2008), `https://bitcoin.org/bitcoin.pdf`
21. Narayanan, A., Bonneau, J., Felten, E., Miller, A., Goldfeder, S.: Bitcoin and Cryptocurrency Technologies. Princeton University Press (2016)
22. Parno, B., Howell, J., Gentry, C., Raykova, M.: Pinocchio: Nearly practical verifiable computation. In: IEEE Symposium on Security and Privacy. pp. 238–252 (2013)
23. Ruffing, T., Moreno-Sanchez, P., Kate, A.: P2P mixing and unlinkable bitcoin transactions. In: Proceedings of the 24th Annual Network and Distributed System Security Symposium. Internet Society (2017). https://doi.org/10.14722/ndss.2017.23415
24. StarkWare: What are STARKs? Online resource (2021), `https://starkware.co/stark/`, accessed 3 June 2026
25. Yakovenko, A.: Solana: A new architecture for a high performance blockchain v0.8.13. Tech. rep., Solana Labs (2018)

Original Research Article *https://doi.org/10.5281/zenodo.23476012*

A CCTS-Based Conceptual Foundation for Explainable Schema Matching Decisions

Ivan Jovanović*[1][0009-0003-1097-7720], Milica Škembarević Sretenović[1][0000-0003-0649-3005], Ivan Luković[1][0000-0003-1319-488X], Nenad Ivezić[2][0000-0003-1863-4370], Anđela Karašićević[2][0009-0001-8731-8243]

[1]University of Belgrade - Faculty of Organizational Sciences
[2]Advanced Integration Research LLC

Abstract. Schema matching is a fundamental problem in data integration. Previous studies have proposed numerous techniques for discovering correspondences between schema components, including linguistic, structural, graph-based, semantic, and hybrid approaches. However, in practice, schema matching remains challenging because schema components are often interpreted without sufficient understanding of the broader semantic context in which they are defined and used. Matching decisions are frequently made manually and ad hoc by integration experts, while matching between components lacks explicit meaning and cannot be automatically reused. To address this gap, we introduce a conceptual foundation for schema matching as a context-aware decision problem, aiming to enable verifiable, repeatable, and traceable matching decisions. The proposed conceptual foundation relies on the Core Components Technical Specification (CCTS) as a methodological basis for interpreting and structuring the semantics of data elements. Based on the previous research done in this area, this approach extends the functionalities of the CCTS based framework by adding key context dimensions and matching shared surface in order to enhance determinism, traceability, explainability, and reusability of the defined matchings. Also, it identifies and describes conceptual matching decisions, which can occur in the data matching process.

Keywords: data integration, schema matching, schema mapping, CCTS, business context

1 Introduction

Schema matching is seen as the task of identifying correspondences between the elements or components of two schemas. The problem of schema matching arises in many domains. These include XML message mapping, data warehouse loading, schema integration, e-business, mediators, and heterogeneous databases [1, 2, 3]. In these scenarios, independently developed schemas often use different terminology, structures, levels of granularity, and schema modeling approaches. This occurs even when they describe related or partially overlapping business concepts. However, in practice, schema matching remains difficult due to the inability to understand the semantic meaning in a

broader sense of schema components, which is a challenge related to issues of semantic ambiguity and lack of shared conceptualization in distributed systems [4]. Despite a long-standing research effort in the field, schema heterogeneity remains a major bottleneck in deploying large-scale integration systems, as noted by [5]. In recent research, according to [6], there are various approaches that address the heterogeneity of data exchange standards and different settings in which business systems operate, but they can be time-consuming and expensive. Also, extension or integration of ontologies that already exist can be quite difficult and costly [7].

The importance of schema matching has been recognized for several decades, as said in [3]. Authors [1] argue that matching should be studied as an independent, generic schema manipulation operation. Their view is that schema matching has broad applicability across data models and domains. These observations show that schema matching is a recurring problem in the development and evolution of interoperable information systems. For this reason, before transforming data between schemas, it is necessary to understand the schemas. It is also necessary to identify semantically equivalent elements.

Our main motivation for this research is to identify and propose concepts that would establish a procedure for optimal schema component matching management that would be context-aware and based on a repeatable, traceable, explainable, and verifiable protocol. The central hypothesis of this paper is that schema matching should be treated as a context-aware decision problem, asserting that determining the validity of correspondences requires considering the business context, not just names, structure, or syntactic similarities. To address this, we use the Core Components Technical Specification (CCTS) as the methodological basis for interpreting and structuring data element semantics, continuing the previous work from [8]. CCTS enables real-world concepts to be abstracted into reusable core components and their context-specific forms, thereby supporting semantic interpretation and facilitating mappings between components.

The aim of the research presented in this paper is to introduce and relate the concepts required for further algorithm development, propose terminology to support the development of a deterministic, explainable algorithm for context-aware schema matching, and ensure the output is a traceable, explainable, validatable, and reusable decision.

In this paper, there are three main contributions:

- Framing schema matching as a context-aware decision problem, not just as the identification of syntactic or structural correspondences.
- Introducing CCTS and CCL-based (*Core Component Library*) component definition for schema components comparison based on a shared semantic foundation.
- Definition of the concepts and steps necessary for future algorithmic formalization

The remainder of this paper is organized as follows. Section 2 reviews related work in schema matching and provides a classification of existing approaches to this prob-

lem. Section 3 introduces the key concepts of the Core Components Technical Specification (CCTS), which form the conceptual foundation of the proposed approach. Finally, Section 4 introduces the fundamental concepts necessary to establish a theoretical foundation for a future deterministic algorithm.

2 Related Work

Prior work has established schema matching as a fundamental operation in data integration.

Authors [1] define matching as an operation that identifies a relationship between elements in two heterogeneous schemas. They also state that matching is a very important step in the data integration process. They emphasize that schema matching is a complex part of mapping because of differences between schemas; even when two elements appear similar, they may be semantically different. At the time of writing their paper, schema matching was performed manually by domain experts using tools with graphical interfaces. At best, some of those tools could find only the simplest correspondences, while even minor differences in names or constraints would prevent such elements from being automatically paired. They conclude that matching is an important problem that should be addressed independently of mapping and transformation, and they even suggest an independent tool devoted exclusively to this problem, one that can operate across different domains.

Authors [2] also agree that schema matching is an important part of integrating diverse information entities. Likewise, they note that, at the time of writing, schema matching was mostly performed manually with the assistance of domain experts, and that this process consumed significant time and other valuable resources, while being highly error-prone and difficult to validate. They argue that it is necessary to automate this process. They surveyed a large number of works on finding such an approach and concluded that no single, standalone approach had completely solved the problem. They therefore propose a hybrid approach that combines several existing simpler approaches, such as name matching or data type matching.

Authors [3] explain that schema matching is a mandatory initial step in schema integration and view it as the identification of relationships between two schemas. The problem of matching has been studied since the 1980s and arises whenever integration is required between two independently developed schemas, each with its own structure and terminology. In their work, the authors describe the schema matching problem in the domain of e-commerce, where there is a high frequency of exchanges of different messages, rules, and data formats, and where exchanged messages use different meanings, names, data types, and constraints. Based on this example, they identify two types of matching: element-level and structure-level. Element-level matching compares two elements based on their names or data types. Structure-level matching compares multiple elements and their positions within the schema. In this paper, we establish the foundation for an approach that would build on structure-level matching, as it enables the

comparison of components together with their structural and contextual dependencies, which is essential for context-aware pairing decisions.

Authors [9] classify matching into syntactic and semantic matching. They represent a schema as a graph, and in syntactic matching, the goal is to pair node labels and assess similarity using syntactic techniques and comparison measures. In semantic matching, elements are paired according to meaning, that is, by considering the "concepts of nodes." Almost all previous works reviewed and discussed in their paper relied on syntactic matching. The authors explain that, in syntactic matching, meaning refers only to, for example, the label of a node, independently of its position in the graph, whereas this is not the case in semantic matching; for that reason, they argue that semantic matching is superior. Based on this conclusion, our future approach is grounded in semantic matching, which serves as a necessary prerequisite for context-aware matching.

Also, authors [10] add that, in syntactic matching, semantics is not analyzed directly, and that this type of matching frequently suffers from name conflicts, especially when an element with the same meaning in two different schemas has different names. In such cases, pairing cannot be performed using that approach alone. For this reason, they argue that matching should be established through semantic relations between the concepts of two schemas, which in fact constitutes semantic matching. Their paper presents a system they developed for semantic mapping, based on comparing elements in graph hierarchies. The authors in [11] also note that matching hierarchical structures is the right approach to the matching problem, especially when the goal is to achieve integration between heterogeneous sources and targets.

In paper [12], the authors warn that perhaps the greatest challenge in finding a universal solution or algorithm for the matching problem lies in the breadth of the domain itself, since it may apply to very different areas of the economy and society, such as e-commerce, databases, data warehousing, software engineering, CAD systems, bioinformatics, and others. They also argue that this still needs to be attempted, even at the risk that a universal matcher may perform worse than a narrowly specialized domain-specific matcher.

Authors [13] identify schema matching as the key first step in the data mapping process, as it enables the identification of semantic relationships between elements of different schemas, which are later used to generate concrete mappings. In their study, they concluded that the available research at the time focused on simple one-to-one cases, in which one source component is paired with one target component. However, they point out that, in practice, complex pairings are very common, involving combinations of multiple attributes and subsequently requiring transformations. In real schemas, complex matches account for up to half of all mappings, further emphasizing their importance. Their discovery is significantly more difficult, however, because unlike the limited number of possible one-to-one correspondences, the number of candidate complex matches is not bounded in advance, which substantially complicates the matching

process. For this reason, the development of methods for semi-automatic generation of complex matches is essential for efficient data integration and exchange.

In addition to that, the authors in [14] define the integration of data from multiple heterogeneous sources as relying on three key steps. These are schema and data management, schema matching, and the definition of concrete mappings. Before any mapping can be carried out, it is first necessary to understand the schema and then identify similar elements within those schemas, which are often different, incomplete, and otherwise difficult to interpret. The authors argue that this process cannot be fully automated because semantic information is often insufficient to fully understand the meaning of the data, and therefore, additional verification or user assistance is required [15]. They also note that the number of possible pairings can be very large. Finally, mapping uses the identified pairings as the basis for deciding on transformations, where additional reasoning is often necessary. One of the first steps toward solving those problems can be found in [16], where business domain ontology modeling was proposed, as it captures the intended meaning of data exchange by formalizing domain concepts and their relationships. The concept of Business Information Entities (BIEs) has been addressed in prior work on standard lifecycle management and schema matching [17].

In recent years, approaches based on machine learning and large language models have been increasingly explored as potential solutions to reduce manual work in schema matching and to automate the matching process. The literature review identified one such system, LSD, which uses previously defined pairings and mappings to train models that make use of various sources of information, including element names, data, structure, and constraints [18]. Another such work describes the LLMatch tool, which uses a Large Language Model to solve complex schema matching problems. The paper reports an experiment showing that this approach can improve integrator productivity and may even improve matching accuracy [19]. The two previously mentioned works focus primarily on predicting relationships between components and schemas and on improving pairing accuracy, whereas our paper addresses the schema matching problem in a different way and will not rely on large language models or machine learning. Authors in [20] conclude that it is important for any technique developed to solve the mapping problem to satisfy the condition of determinism, that is, when inputs, rules, and constraints are predefined and known, the outcome of the procedure must be unique. Therefore, this paper does not adopt an LLM- or ML-based approach, since they are inherently probabilistic and may produce non-deterministic outputs (discussed in [21] where results show almost deterministic behavior, but never fully deterministic).

3 Conceptual Foundations For Context-Aware Schema Matching

This paper establishes the theoretical foundation for an approach to understanding the decision-making process for pairing components between the source and target schemas, assuming that the meaning of components is not determined solely by their names,

structures, or positions in the document. It assumes that, during component pairing, the context in which components are defined and used must be explicitly considered. In this paper, we define a set of concepts and relations among them that enable consistent and verifiable management of the future process of pairing and subsequent mapping, and that provide the basis for an algorithm that will address the schema matching problem. Since the research presented in this paper builds directly on the research done in [8], where CCTS is adopted as the methodological framework for context-based integration, we also rely on CCTS as the conceptual basis for representing reusable components and their context-specific interpretations. It is important to note that many integration scenarios are not natively CCTS-based, and an additional preparatory step is needed in order to bridge the gap between the schemas that are originally based on and the CCTS, but that part is out of scope of this research.

3.1 Semantic and Contextual Grounding of Schema Components

The inputs that will be used by the future algorithm consist of the source and target schemas, which are represented as graphs of components alongside information on their relationships. Such a representation makes it possible to interpret a component not in isolation, but as part of a broader structure, in which the meaning of one element also depends on its place within the context. Graph representation is also important because matching in certain integration use cases involves not only simple fields but also aggregate structures, relations between components, and the hierarchy of business information. While previous work has explored business context as a factor in integration [22], this paper extends the concept by introducing Key Context Dimensions and a formalized pairing framework.

CCTS (Core Components Technical Specification) is an ISO standard for data exchange that specifies the rules for creating standardized, reusable building blocks based on semantics [15]. The purpose of this standard is to enhance interoperability and capture context-specific information about the schema elements. Business context (BC) describes broader real-life circumstances that shape the form in which a certain component will appear. By capturing factors from different areas (e.g., socio-economic, geographical, etc.), it becomes easier to determine the factors within the context and how they influence the form of the entities within them. Consequently, by knowing the broader context, it becomes easier to determine the exact semantic meaning of a component within it.

A context taxonomy represents an ordered set of context dimensions through which a specific business use case, in which a component has meaning, is described. These dimensions may include categories such as business process, geopolitical location, tax regime, and others relevant to a concrete integration domain. The role of the taxonomy is to provide a consistent, uniform set of categories and values that describe the dimensions of a business use case, thereby affecting the understanding of a component and

giving it specific meaning. Each dimension must therefore be clearly defined, together with the set of admissible values, so that the context is explicit and consistent.

Another concept defined within this standard is a core component (CC). CC captures the semantic meaning of the perceived concept from reality and holds all possible forms of the concept's shape across different business contexts. The way CC manifests when it appears in a certain BC is by taking the predefined parts of the CC that are related to the forms of the concept in that specific use case, which is related to a certain BC. A business information entity is a logical concept that specifies the form of the CC for the given BC. BIEs are contextualized forms of a CC, i.e., a CC with a given context, that therefore use just those attributes relevant to the context in which they are.

By knowing that the core logic is encapsulated in the CC and that each CC will have the same form when it appears in a certain BC, it becomes clear that this mechanism can be utilized for enhancing reusability during the matching process. Namely, the idea is that it is enough to map BIEs (related to the same CC) from one BC to another only once, and then, whenever a new schema from the same BC arrives, if there has been a previously established matching from that BC to another, it is possible to use the same matching logic without the intervention of either the human or the algorithm. However, the prerequisite for this is that both source and target schemas are based on the CCTS.

Core Component Library (CCL) is the storage where all CCs and their corresponding BIEs are held alongside their definitions, metadata, and cardinalities. Two scenarios could happen when a concept is contextualized (CC is assigned a BC):

1. A BIE related to a certain CC and a certain BC does not exist in the CCL
2. A BIE related to certain CC and BC exists in the database

The first case means that a new CCL element appeared and that it needs to be added to the CCL. A new BIE is constructed, assigned to the corresponding CC, and, if necessary, the CC is extended with additional context-specific information derived from the BC that triggered the change. By gradually adding BIEs and expanding CCs the library becomes the source of matching reusability. In the second case, an already existing BIE can be identified and reused when the same CC occurs in the same BC, thus minimizing the reconstruction of BIEs and allowing the same BIEs to be reused.

CCL is regarded as a shared, abstract layer that allows elements from different schemas to connect to a common conceptual foundation. In this way, the comparison between the source and target schema is not based solely on syntax, element names, or technical paths; instead, it relies on a common set of core components and ,their rules, constrained by a specific use case. The role of the CCL is therefore not only structural or descriptive, but primarily semantic, because it provides meaning to the components. The CCL thus provides the basis for both schemas and serves as a prerequisite for the semantic pairing of two components.

Assigning context to components involves defining a set of allowed contextual combinations within a predefined taxonomy, in which components from the CCL are assigned the corresponding admissible subset of taxonomy values. In this way, it is explicitly determined for each component in which use cases it is applicable and semantically valid. This process connects the structural schema model, represented as a graph of components and their relations, to the semantic layer, expressed through the context

taxonomy. The assigned context results from the intersection of predefined business rules and permitted categories in the taxonomy, thereby making sure that each component is restricted to the set of use cases for which its use is semantically and commercially justified.

A profile represents a view of the schema that depends on a concrete use case and contains only those components that have a valid effective context (effective meaning that only the parts of the CC that are relevant to the observed use case will appear, disregarding the parts of CC that are irrelevant to the given use case) in that use case. In that sense, a profile represents a filtered and contextualized version of the existing base schema; in other words, a profile contains only those components that are relevant to the given use case, together with their effective context. The role of the profile is to reduce schema complexity and to shift the consideration of matching from the complete source and target structures to their relevant parts for a specific use case.

3.2 Conceptual Basis for Context-Aware Matching Decisions

The semantic basis of both the source and target schema is defined by the CCTS methodology. It is important to note that this framework assumes that both schemas are based on the CCL and that their components comply with the CCTS modeling rules defined earlier.

Building on this, the concept of the shared matching surface is introduced. This concept denotes the shared semantic scope of the source and target profiles, that is, the part in which meaningful overlap exists between the two schemas. The premise of the shared matching surface is that pairing is not considered across the entire schemas, but only where comparison is justified at all, that is, over the components included in the profiles. This means that decisions about component pairing are made only for the components relevant to a specific use case on both the source and target sides, which facilitates the decision-making process.

The matching of components is based on their connection with the core components from the Core Components Library (CCL). Each component in the profile represents a contextualized instance of a given core component (CC), that is, its business information entity (BIE). In this way, although components from the source and target schema may have different names, structures, or positions in the model, their comparison is based on a common foundation in the form of the same or a corresponding Core Component. Pairing can therefore be understood as an attempt to establish a relation between components that share the same abstract basis but are adapted to different use cases, that is, have different contexts.

It has been observed that different components do not depend on the same aspects of the business environment (e.g., a component that is related to a tax regime is influ-

enced by socio-economic factors and not by natural factors). This means that, in deciding on component pairing, it is not sufficient merely to know that a component has an effective context; it is also necessary to understand which dimensions of that context actually affect its meaning. For that reason, the notion of Key Context Dimensions (KCD) is introduced, and in this paper, we use it to denote those context dimensions that are semantically important or significant for a concrete profile component.

A component for which an effective context has been calculated contains information about all dimensions of the business context; however, not all dimensions are semantically important for all components. KCDs represent a semantically focused subset of all context dimensions associated with a given component. In the current scope of this paper, KCDs are treated as design-time components. They are identified by the model designer or domain expert, based on the component semantics, the context taxonomy, and domain rules indicating which context dimensions may affect the interpretation or applicability of the component. For example, for one component, belonging to a particular business process may be significant, while regulatory or geopolitical conditions may be completely irrelevant. For another component, the geopolitical dimension may be the important one, and changing its value within that dimension, in accordance with the taxonomy's rules, may completely change the component's meaning. The introduction of KCDs, therefore, enables semantic pairing of source and target components, as only the significant context dimensions of the two components are compared during pairing. If a context dimension is not important for a particular component, that is, if that component in no way depends on a change in the value of that dimension, then that dimension should not be taken into account when comparing two components from the source and target schema. KCDs, therefore, represent an important step toward correct reasoning about components and verifiable decision-making when comparing two components or when reusing the same component. Since KCDs are predefined and used as fixed input to the future matching procedure, the determinism of the approach refers to the decision-making process under a given set of schemas, context assignments, and KCDs.

Based on the shared semantic scope and the consideration of context through KCDs, the final outcome should be a decision on the pairing of components between the source and target schema. In this paper, that decision is introduced and explained at the most general level, without elaborating on the procedure for calculating it.

In this paper, schema matching, schema mapping, and transformation are treated as distinct but related concepts. Schema matching refers to the process of identifying semantic correspondences between elements belonging to two different schemas [23]. Schema mapping refers to a high-level declarative specification describing how elements of a source schema relate to elements of a target schema [23]. Transformation refers to the additional logic required when a valid pair cannot be mapped directly because differences in the relevant context dimensions affect how the data should be interpreted or represented. Therefore, the expected output of the proposed approach is a matching decision, possibly accompanied by a transformation requirement, but not a complete executable mapping or transformation specification.

In addition to the decision itself, the future approach is expected to document that decision in a Matching Rule Artifact (MRA). In this paper, the MRA is understood as a conceptual decision record that captures additional information about the decision, including which components are matched, the contextual conditions under which the decision is made, the context dimensions relevant to the specific matching case, the reasoning behind the decision, and the means by which it may be validated. In this way, the MRA links the matching decision to its underlying reasoning, context, and validation requirements, thereby supporting traceability, explainability, and reuse.

Based on the degree of alignment in the relevant context dimensions, three expected types of mapping outcomes are distinguished.

The first type of outcome is seamless matching, in which the source and target components achieve such a level of correspondence in a concrete use case that their pairing and subsequent mapping do not require additional context-dependent transformation logic. In this case, the components preserve the same meaning and form, so the mapping can be considered directly applicable, even if the component names differ. The key characteristic of this type of mapping is its semantic stability. It is important to note that, during pairing, differences in the broader business context, that is, differences in non-key dimensions, do not affect the way in which the components are paired.

The second type of outcome is transformation-required matching, which refers to a situation in which the source and target components are not fully aligned in the relevant context dimensions. In this case, pairing and subsequent mapping are not directly applicable and require additional transformation logic explicitly conditioned by context. In other words, the pairing of components is not semantically stable but depends on the values of certain relevant context dimensions, which necessitates defining rules that determine how the matching changes across different use cases. The key characteristic of this type of decision is that it preserves the shared foundation between the components, but requires a context-conditioned transformation.

The third type of outcome is no matching, which occurs when there is no corresponding counterpart on the other side within the shared scope of consideration. Specifically, this outcome arises when a component is present in one profile, that is, has a valid effective context, while on the other side, there is no component available for pairing in the same scenario, that is, it has no effective context. In such cases, the pairing procedure is not carried out because there is no corresponding component that would be semantically and contextually identical or sufficiently similar. The key characteristic of this outcome is that the absence of a component in one of the profiles directly excludes the possibility of pairing, thereby avoiding the consideration of invalid or arbitrary relations between elements.

The role of the expected outcome types is to establish a clear terminological framework and to show that the goal of the defined framework is not only the identification

of similarities, but the formation of an explicit and verifiable conclusion about the nature of the relationship between components, thereby creating a suitable basis for the reuse of components and of the defined pairings between components, and subsequently also for mapping.

As discussed above, the expected output of the decision-making algorithm should also be a Matching Rule Artifact (MRA). MRA is viewed as an explicit artifact that records not only what has been paired but also under which conditions, based on which assumptions, and with which validation requirements. Its function is to facilitate validation, reusability, and understanding of pairing by replacing the integrator's implicit assumptions with an explicit, verifiable, and manageable trace of the decision and providing a good basis for mapping between components.

Figure 1 provides a simplified MRA template together with an illustrative invoice-related MRA instance.

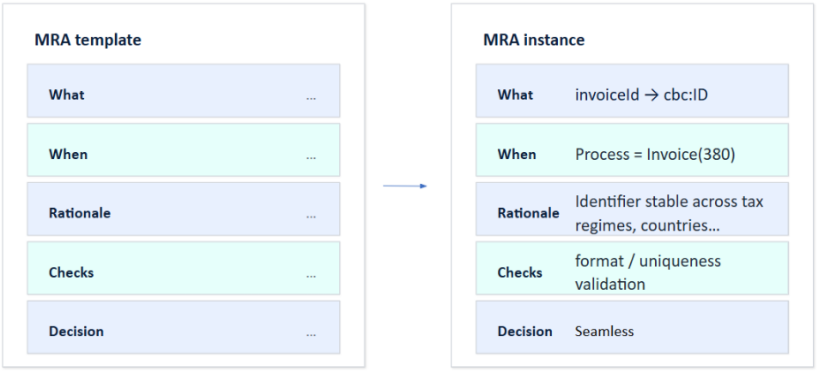


Figure 1. Simplified MRA template and an example MRA instance

The goal of establishing the basis for a future approach and algorithm is not only to produce a correct decision about component pairing, but also to make that decision deterministic, explainable, verifiable, and reusable. This means that the result should not be arbitrary or tied to a one-time interpretation, but should be grounded in a clearly described context and a shared semantic basis, and thereby satisfy the requirements defined above. In addition, each decision made must also be traceable. In other words, for every decision, it must be possible to reconstruct why a given component was considered for pairing within the shared semantic scope, which dimensions of its context were relevant (KCD), and why the final outcome is precisely the decision made (MRA) and the decision. This creates the basis for validation, monitoring, and subsequent reuse of decisions in modified or new integration scenarios.

3.3 Illustrative Example and High-Level Procedure for Context-Aware Matching Decisions

To illustrate how the proposed concepts relate to one another, we consider a simplified invoice-related matching scenario. The source and target sides are described through two business contexts. The assigned business contexts for the source and target sides are summarized in Table 1.

Table 1. Summarized the assigned business contexts for the source and target sides

Context dimension	Source assigned business context	Target assigned business context
Country	DE (Germany)	NL (Netherlands)
Business Process	Invoice 380 (Standard Commercial Invoice)	Invoice 380 (Standard Commercial Invoice)
Tax Regime	S20 (20% rate)	AE0 (0% rate)
Contract Terms	PT30 (Pay in 30 days)	PT30 (Pay in 30 days)

These values represent specific combinations of context dimensions and their assigned values for the source and target use cases. In this sense, they define the assigned business context within which the corresponding components are interpreted and considered for pairing. Only a subset of these dimensions may be relevant for a particular component, which is later captured through the notion of Key Context Dimensions (KCDs).

To illustrate the three expected matching outcomes, the example considers three invoice-related BIEs, each grounded in a corresponding Core Component: InvoiceID, TaxAmount, and PaymentDueDate.

Following the CCTS-based modeling logic, the initial source and target schema components are grounded in reusable Core Components (CC) stored in the CCL, while their context-specific forms are represented as corresponding BIEs. Based on the assigned business context, the effective context is determined for each component, resulting in source and target profiles that contain only those contextualized components that are relevant for the observed use case. The shared matching surface is then derived from the components included in these profiles, while candidate pairings are considered by comparing the underlying CCs of the corresponding BIEs. This enables components to be compared through a shared conceptual basis, while preserving the contextual restrictions.

In this example, InvoiceID and TaxAmount enter the shared matching surface because corresponding BIEs exist on both sides, while PaymentDueDate does not because no corresponding target-side component is available.

The first illustrative case concerns the InvoiceID BIE. This component is present in both the source and target profiles and therefore belongs to the shared matching surface. For this component, the key context dimension (KCD) is the business process. Since the business process is Invoice 380 in both assigned business contexts, the component is considered contextually stable for the observed use case. The corresponding outcome is therefore seamless matching, as no additional context-dependent transformation is required for subsequent mapping.

The second illustrative case concerns the TaxAmount BIE. The TaxAmount is included in both the source and target profiles and may therefore be considered a candidate pair within the shared matching surface. However, for this component, the KCDs are the Business Process and Tax Regime dimensions. While the business process remains Invoice 380 on both sides, the tax regime changes from S20 in the source profile to AE0 in the target profile. For this reason, the outcome is not seamless matching. Instead, the appropriate outcome is transformation-required matching, because the components can be paired, but subsequent mapping would require additional transformation logic conditioned by the tax context. Thus, the proposed approach identifies the need for additional transformation logic and specifies the context dimensions that should guide such a transformation, without defining the concrete transformation rules themselves.

The third illustrative case concerns the PaymentDueDate BIE. In the observed use case, this component is present in the source profile, while no corresponding target component is available. Since it does not appear on both sides, PaymentDueDate cannot enter the shared matching surface, and no candidate pair can be formed. The appropriate outcome is therefore no matching, and the component is excluded from further pairing consideration.

For each matching outcome, the future approach is expected to record the decision in a Matching Rule Artifact (MRA). Table 2 illustrates the proposed MRA structure through three simplified cases, showing how different matching outcomes can be documented together with their context, KCDs, rationale, and validation focus.

Table 2. Proposed MRA structure through three simplified cases

MRA element	InvoiceID	TaxAmount	PaymentDueDate
Component pair	InvoiceID_BIE → InvoiceID_BIE	TaxAmount_BIE → TaxAmount_BIE	No corresponding target component

Shared basis	InvoiceIdentifier CC	TaxAmount CC	PaymentDueDate CC
Source KCD values	BP = Invoice 380	BP = Invoice 380, Tax Regime = S20	BP = Invoice (380), PT=30
Target KCD values	BP = Invoice 380	BP = Invoice 380, Tax Regime = AE0	Component not present on target side
KCDs	Business Process	Business Process, Tax Regime	Payment Terms
Matching outcome	Seamless matching	Transformation-required matching	No matching
Rationale	The key context dimension is aligned	A relevant context dimension differs	No valid counterpart exists in the shared matching surface
Validation focus	Identifier format and uniqueness	Tax calculation, rule consistency	Validation of exclusion from pairing scope

This example illustrates that the proposed approach does not treat component pairing as a simple comparison of names or structural similarity. Instead, candidate components are first grounded in their CCTS-based semantic basis, then considered within scenario-specific profiles and the shared matching surface, and finally evaluated through the context dimensions relevant to the component.

The same conceptual flow can be expressed more generally as a high-level procedure, where the assigned contexts are first used to derive profiles, the profiles define the shared matching surface, KCDs guide the comparison, and the resulting decision is recorded in an MRA.

Algorithm 1. Conceptual Procedure for Context-Aware Matching Decisions

1: ***Input***
2: S $CCTS-based\ source\ schema$
3: T $CCTS-based\ target\ schema\ for\ p\ in\ patterns\ do$
4: CCL $Core\ Component\ Library$
5: CT $Context\ taxonomy$
6: Us $Source\ use-case\ context$
7: Ut $Target\ use-case\ context$
8: KCD $Function\ assigning\ relevant\ context\ dimensions\ to\ components\ or\ component\ pairs$
9: ***Output***
10: MRA_Set $Set\ of\ Matching\ Rule\ Artifacts$
11: ***Notation***:
12: $EC(c, U)$ $effective\ context\ of\ component\ c\ under\ use\ case\ context\ U$
13: $CC(c)$ $Core\ Component\ associated\ with\ component\ c$
14: $project(E, K)$ $projection\ of\ context\ E\ on\ dimensions\ K$

```
Contextualize source and target components
  for each component x in S:
        ECs(x) := EC(x, Us)
  for each component y in T:
        ECt(y) := EC(y, Ut)
Create source and target profiles
   Ps := { x in S | ECs(x) is not empty }
   Pt := { y in T | ECt(y) is not empty }
Identify the shared matching surface
    SMS := empty set
      for each x in Ps:
        for each y in Pt:
          if CC(x) = CC(y):
            add (x, y) to SMS
Decide the matching outcome for candidate pairs
     MRA_Set := empty set
     for each candidate pair (x, y) in SMS:
        K := KCD(x, y)
        Cs := project(ECs(x), K)
        Ct := project(ECt(y), K)
        if Cs is empty or Ct is empty:
          decision := NO_MATCHING
        else if allKCDsAligned(EC_s_K, EC_t_K, K):
          decision := SEAMLESS_MATCHING
        else:
          decision := TRANSFORMATION_REQUIRED_MATCHING
        MRA := createMRA(
              sourceComponent = x,
              targetComponent = y,
              sourceCoreComponent = CC(x),
              targetCoreComponent = CC(y),
              relevantContextDimensions = K,
              sourceContextValues = Cs,
              targetContextValues = Ct,
              matchingDecision = decision,
              validationRequirements = defined according to decision type,
              traceInformation = {S, T, CCL, CT, Us, Ut, KCD, ECs(x), ECt(y)}
            )
         add MRA to MRA_Set
Handle components outside the shared matching surface
     for each component x in Ps with no candidate pair in SMS:
        add createMRA(
            sourceComponent = x,
            targetComponent = none,
            sourceCoreComponent = CC(x),
            targetCoreComponent = none,
            matchingDecision = NO_MATCHING
          ) to MRA_Set
     for each component y in Pt with no candidate pair in SMS:
        add createMRA(
            sourceComponent = none,
            targetComponent = y,
            sourceCoreComponent = none,
            targetCoreComponent = CC(y),
            matchingDecision = NO_MATCHING
          ) to MRA_Set
Return MRA_Set
```

4 Conclusion

The data mapping problem remains a multifaceted challenge, especially when the mapping needs to be done based on semantics rather than the structure itself. By introducing the new concepts (KCD and shared matching surface), the proposed approach offers a repeatable, verifiable, traceable, and explainable process for data matching based on context-aware reasoning. Knowing the KCDs, the algorithm based on this approach will be able to determine the relevant context categories that influence the change in the mapping and consequently in the form of a BIE within it. Once documented, the matching can then be reused when the same situation arises: the same KCDs are present, and the same shared surface is in the intersection of the KCDs. The main benefit of this approach is the determinism that was not guaranteed by other approaches, especially those based on ML or LLMs. Another contribution lies in the identification of three types of decisions that can be made in CCTS-based mapping and the corresponding actions that need to be taken afterwards. Overall, the proposed foundation contributes to the development of more reliable, interpretable, and reusable data matching, supporting consideration of the context and using it to explain the matching decisions. Also, one of the effects of an algorithm shaped in such a way is that it is traceable and every decision that was made by the algorithm can be reconstructed step by step and better understood.

Despite its contributions, this research is just the beginning of the development of a fully specified algorithm for automated and deterministic data matching. The future areas of research include, but are not limited to, the formalization of the decision-making process with the corresponding algorithm, identification of KCDs, and a comparison algorithm for determining the similarities and differences between BIEs related to the same CC, as well as the possible changes that need to happen after seamless matching to accommodate implementation details in a certain system.

Acknowledgments

This research was supported by the University of Belgrade – Faculty of Organizational Sciences and, in part, by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia through institutional funding (grant number: 200151).

References

1. Madhavan, J., Bernstein, P.A., Rahm, E.: Generic schema matching with Cupid. In: Proceedings of the 27th International Conference on Very Large Data Bases (VLDB 2001), pp. 49–58. Morgan Kaufmann, San Francisco (2001).
2. Do, H.H., Rahm, E.: COMA—a system for flexible combination of schema matching approaches. In: VLDB 2002: Proceedings of the 28th International Conference on Very Large

Data Bases, pp. 610–621. Morgan Kaufmann, San Francisco (2002). doi:10.1016/B978-155860869-6/50060-3
3. Rahm, E., Bernstein, P.A.: A survey of approaches to automatic schema matching. The VLDB Journal 10(4), 334–350 (2001). doi:10.1007/s007780100057
4. Pouchard, L., Ivezic, N., Schlenoff, C.: Ontology engineering for distributed collaboration in manufacturing. In: AIS2000 Conference, pp. 1012–1017 (2000)
5. Nguyen, H.-Q., Taniar, D., Rahayu, J.W., Nguyen, K.: Double-layered schema integration of heterogeneous XML sources. Journal of Systems and Software 84(1), 63–76 (2011). doi:10.1016/j.jss.2010.07.055
6. Jelisic, E., Kulvatunyou, B., Sidi Mahaman Malick, S., Oh, H., Ki, J.: Use case decomposition for multi-standard management to enable flexible supply chain. In: ASME 2025 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, vol. 2A, Anaheim, CA, USA (2025). doi:10.1115/DETC2025-168568
7. Obrst, L., Grüninger, M., Baclawski, K., Bennett, M., Brickley, D., Berg-Cross, G., Hitzler, P., Janowicz, K., Kapp, C., Kutz, O., Lange, C., Levenchuk, A., Quattri, F., Rector, A., Schneider, T., Spero, S., Thessen, A., Vegetti, M., Vizedom, A., Westerinen, A., West, M., Yim, P.: Semantic Web and Big Data meets Applied Ontology: The Ontology Summit 2014. Applied Ontology 9(2), 155–170 (2014). doi:10.3233/AO-140135
8. Jelisić, E.T.: Integracija poslovnih sistema primenom poslovnog konteksta na model životnog ciklusa standarda za razmenu poruka. Ph.D. thesis, University of Belgrade, Faculty of Organizational Sciences, Belgrade (2024)
9. Giunchiglia, F., Shvaiko, P.: Semantic matching. The Knowledge Engineering Review 18(3), 265–280 (2003). doi:10.1017/S0269888904000074
10. Giunchiglia, F., Shvaiko, P., Yatskevich, M.: S-Match: an algorithm and an implementation of semantic matching. In: Bussler, C.J., Davies, J., Fensel, D., Studer, R. (eds.) The Semantic Web: Research and Applications, ESWS 2004. LNCS, vol. 3053, pp. 61–75. Springer, Berlin, Heidelberg (2004). doi:10.1007/978-3-540-25956-5_5
11. Avesani, P., Giunchiglia, F., Yatskevich, M.: A large scale taxonomy mapping evaluation. In: Gil, Y., Motta, E., Benjamins, V.R., Musen, M.A. (eds.) The Semantic Web – ISWC 2005. LNCS, vol. 3729, pp. 67–81. Springer, Berlin, Heidelberg (2005). doi:10.1007/11574620_8
12. Melnik, S., Garcia-Molina, H., Rahm, E.: Similarity flooding: a versatile graph matching algorithm and its application to schema matching. In: Proceedings of the 18th International Conference on Data Engineering (ICDE 2002), pp. 117–128. IEEE (2002). doi:10.1109/ICDE.2002.994702
13. Dhamankar, R., Lee, Y., Doan, A., Halevy, A., Domingos, P.: iMAP: discovering complex semantic matches between database schemas. In: Proceedings of the 2004 ACM SIGMOD International Conference on Management of Data, pp. 383–394. ACM, New York (2004). doi:10.1145/1007568.1007612
14. Miller, R.J., Hernández, M.A., Haas, L.M., Yan, L.-L., Ho, C.T.H., Fagin, R., Popa, L.: The Clio project: managing heterogeneity. ACM SIGMOD Record 30(1), 78–83 (2001). doi:10.1145/373626.373713
15. UN/CEFACT: Core Components Technical Specification (CCTS), Version 3.0. UNECE (2009). https://unece.org/DAM/cefact/codesfortrade/CCTS/CCTS-Version3.pdf. Accessed 23 Apr 2026
16. Vujasinovic, M., Ivezic, N., Kulvatunyou, B., Barkmeyer, E., Missikoff, M., Taglino, F., Marjanovic, Z., Miletic, I.: Semantic mediation for standard-based B2B interoperability. IEEE Internet Computing 14(1), 52–63 (2010). doi:10.1109/MIC.2010.17

17. Oh, H., Kulvatunyou, B., Jones, A., Finin, T.: Employing word-embedding for schema matching in standard lifecycle management. Journal of Industrial Information Integration 38, 100547 (2024). doi:10.1016/j.jii.2023.100547
18. Doan, A., Domingos, P., Halevy, A.Y.: Reconciling schemas of disparate data sources: a machine-learning approach. In: Proceedings of the 2001 ACM SIGMOD International Conference on Management of Data, pp. 509–520. ACM, New York (2001). doi:10.1145/375663.375731
19. Wang, S., Li, Y., Xiao, H., Dai, B.T., Lee, R.K.-W., Dong, Y., Deng, L.: LLMATCH: a unified schema matching framework with large language models. In: Li, J., Chbeir, R., Li, L., Zong, C., Zhang, Y., Zhang, M. (eds.) Web and Big Data. APWeb-WAIM 2025. LNCS, vol. 16116, pp. 343–356. Springer, Singapore (2026). doi:10.1007/978-981-95-5722-6_37
20. Bonifati, A., Mecca, G., Papotti, P., Velegrakis, Y.: Discovery and correctness of schema mapping transformations. In: Bellahsene, Z., Bonifati, A., Rahm, E. (eds.) Schema Matching and Mapping, pp. 111–147. Springer, Berlin, Heidelberg (2011). doi:10.1007/978-3-642-16518-4_5
21. Cui, A.Y., Yu, P.: Do language models have Bayesian brains? Distinguishing stochastic and deterministic decision patterns within large language models. arXiv preprint arXiv:2506.10268 (2025). doi:10.48550/arXiv.2506.10268
22. Jelisic, E., Ivezic, N., Kulvatunyou, B., Milosevic, P., Babarogic, S., Marjanovic, Z.: A novel business context-based approach for improved standards-based systems integration—a feasibility study. Journal of Industrial Information Integration 30, 100385 (2022). doi:10.1016/j.jii.2022.100385
23. Dimitrieski, V., Kordić, S., Ristić, S., Kern, H., & Luković, I. (2025). Mapping-based approach to integration of technical spaces. *Computer Science and Information Systems, 22*(1), 279–310. doi: 10.2298/CSIS240701009D

Original Research Article *https://doi.org/10.5281/zenodo.23476090*

Digital Transformation of a Private Higher Education Institution: Evidence from a Serbian University and Implications for the Sector

Miloš Vrbanac[1,2][0009-0005-5514-5703]

[1] University of Belgrade, Faculty of Organizational Sciences, Belgrade, Serbia
[2] University Union Nikola Tesla, Belgrade, Serbia
mv20253261@stud.fon.bg.ac.rs · milosvrbanac@unionnikolatesla.edu.rs

Abstract. Private higher education institutions in transition economies face a specific challenge when it comes to digital transformation: the pressure to modernize is real, but the resources to do so comprehensively are limited. This paper looks at how University Union Nikola Tesla (UNT) in Belgrade, Serbia, approached this challenge. Using a case study methodology that draws on internal institutional data, competitive benchmarking across six institutions, and an analysis of the full student journey from application through enrollment to recurring administrative cycles, the paper identifies the main obstacles to digital maturity and proposes a phased transformation roadmap structured around that journey. The institution's problems are not caused by missing technology, but by a digital service environment that does not communicate with the user proactively at any point in the journey, from the prospective applicant to the enrolled student. The paper concludes with recommendations that, given the structural similarities between UNT and comparable private universities in the region, may be applicable beyond this single case.

Keywords: Digital Transformation · Higher Education · Business Model Canvas · Process Redesign · Private University · Transition Economy

1 Introduction

Digitalization in higher education is no longer an optional upgrade. Students increasingly compare the quality of university administrative services to commercial digital services they use daily, and they expect a similar level of convenience [1]. For private universities, this creates a particularly difficult position: they compete with public institutions that often have larger budgets, more staff, and decades of established digital infrastructure, while operating with significantly fewer resources.

The problem this paper investigates is twofold. First, why do private HEIs in transition economies often appear digitalized on the surface but still leave students dependent on manual queries and in-person visits? Second, what kind of transformation approach is realistic for an institution with a small marketing budget, limited IT capacity, and direct competition from publicly funded universities offering free tuition? While the

broader topic of digital transformation in higher education is well covered in the literature, the specific situation of mid-sized private institutions in transition economies has received comparatively less attention.

Most of the current research on digital transformation in higher education focuses either on large public universities with substantial resources, or on the introduction of specific tools such as learning management systems and online platforms [2, 6]. Studies that examine the strategic transformation of mid-sized private institutions in Central and Eastern Europe are less common, and those that do exist tend to remain at the descriptive level rather than offering a structured framework for action [7]. The contribution of this paper, accordingly, is to offer a concrete, applicable model for institutions that face budget and structural constraints similar to those of UNT. The paper combines a current-state diagnosis using SWOT and Business Model Canvas, a benchmarking comparison with three domestic and three international institutions, and a phased roadmap that connects administrative process redesign with a digital marketing strategy aimed at improving enrollment outcomes.

A note on the role of digital marketing in this study is appropriate at the outset. The marketing automation component developed in later sections is not treated as a separate marketing topic, but as part of the student-service transformation layer. The application-to-enrollment workflow is the first sustained interaction a prospective student has with the institution's digital services, and the quality of that interaction is shaped by the same design principles that define the administrative side of the transformation: process transparency, proactive communication, and reduced friction. For this reason, enrollment conversion improvement is used as a measurable indicator of digital transformation effectiveness, alongside more conventional indicators such as registration completion rates and the reduction of administrative queries.

The remainder of the paper is organized as follows. Section 2 reviews the relevant theoretical background. Section 3 describes the case study methodology and data sources. Section 4 presents the current-state analysis, including the stakeholder context, identified pain points across the student journey, and a SWOT framed in terms of digital transformation readiness. Section 5 develops the proposed transformation model, presenting the AS-IS and TO-BE Business Model Canvas, benchmarking findings, and a three-phase roadmap structured around the student journey. Section 6 discusses the implications and limitations of the findings, and Section 7 concludes.

2 Theoretical Background

The distinction between digitization, digitalization, and digital transformation is well established in the literature. Digitization refers to converting information from analog to digital form. Digitalization means changing how work is done using digital tools. Digital transformation goes further: it involves redesigning processes, value propositions, and organizational relationships in a way that digital technology makes possible [2, 10]. In practice, many institutions stop at the second stage. They introduce digital tools but do not change the underlying logic of their processes, which often means that inefficiencies are preserved in a digital format rather than eliminated [3].

Several digital transformation frameworks have been proposed for the higher education context. Kopp, Gröblinger, and Adams identify recurring barriers to digital transformation in universities, including organizational inertia and the gap between technological provision and actual use [6]. Sandkuhl and Lehmann frame digital transformation in higher education as a structured progression across organizational, process, and technological dimensions rather than a one-off technology adoption [3]. The organizational dimension is particularly important in institutions operating under digital disruption, where the effectiveness of internal coordination and knowledge transfer shapes the success of transformation efforts [12]. These frameworks share a common emphasis: transformation succeeds or fails at the level of process and organizational design, not at the level of tool availability.

The Business Model Canvas (BMC), introduced by Osterwalder and Pigneur [4], is widely used as a tool for mapping and redesigning organizational models. Although originally developed for commercial companies, it has been applied in higher education research to analyze how institutions create and deliver value to their stakeholders [5]. Its advantage in this context is that it forces institutions to make explicit what they actually offer and to whom, which often reveals gaps between the intended and the actual student experience.

A growing body of work frames institutional digital quality through the student journey perspective, treating the prospective and enrolled student as a continuous user passing through successive digital touchpoints: the application, the enrollment confirmation, the first administrative interaction, and the recurring exam and scheduling cycles [9]. This perspective argues that digital maturity must be assessed across the full journey rather than at any single point, since the user experiences the institution as one continuous relationship rather than as a set of separate functional services. The growing role of mobile channels in this journey is well documented: studies of students in the regional context show a strong propensity to adopt mobile learning and mobile-first services, which has direct implications for how institutions design their digital touchpoints [11]. Studies on digital maturity in higher education identify several factors that distinguish well-functioning digital environments from poorly functioning ones: the ability of users to track the status of their requests in real time, the ability to complete administrative tasks independently, and proactive communication from the institution before the user needs to ask [6]. Private HEIs in Central and Eastern Europe often struggle to reach these standards, primarily due to budget limitations and the incremental way in which digital systems were introduced over time [7].

3 Methodology

The study uses a single case study approach [8], with UNT as the unit of analysis. This methodology is appropriate when the aim is to understand a complex organizational situation in context, rather than to test a hypothesis across a large sample. The findings are intended to be useful for thinking about comparable institutions, not to be statistically representative of the sector.

Data was collected from three sources. The first was internal institutional information, including enrollment figures for the 2024/25 academic year, the annual marketing budget and its breakdown by channel, and documentation of current administrative systems. The second was direct observation of how the key student-facing processes work in practice, covering the application stage, exam registration, academic scheduling, and student communication with the administration. The third was a benchmarking analysis covering six institutions, selected according to explicit criteria described below.

Benchmarking Criteria. The benchmarked institutions were selected to satisfy three criteria. First, comparability of context: three domestic institutions (FON Belgrade, Singidunum University, and the Faculty of Architecture Belgrade) operate in the same national regulatory and market environment as UNT. Second, demonstrated digital maturity: three international institutions (the Technical University of Munich, KU Leuven, and Georgia State University) were chosen as recognized reference points for digital student services, each illustrating a specific mature practice. Third, relevance of practice: each institution was included because it exemplifies at least one of the three digital maturity dimensions used in the analysis, namely process transparency, self-service capability, or proactive notification. Institutions were assessed across these same three dimensions to keep the comparison consistent.

The analytical tools used were SWOT analysis framed around digital transformation readiness, an AS-IS Business Model Canvas representing the current institutional model, a TO-BE Business Model Canvas representing the proposed redesign, and a conversion funnel model for the marketing strategy.

4 Current State Analysis

4.1 Institutional Context and Stakeholder Map

UNT is a private integrated university founded in 2011 in Belgrade, Serbia, comprising seven faculties with approximately 1,800 enrolled students. It holds a distinctive market position as the only private institution in the Western Balkans offering accredited programs in Architecture and Civil Engineering, alongside programs in Ecology, International Politics, Entrepreneurial Business, Informatics, and Economics. Despite this positioning, it faces challenges representative of its institutional category: an annual marketing budget of approximately 15,000 EUR, 160 unfilled enrollment places in 2024/25 out of 720 accredited, and administrative processes where digital tools exist but are supplemented by manual workarounds.

The stakeholder environment consists of three groups with different needs. Students, especially those who are also employed, form the primary user group for all digital services; their administrative needs often arise outside standard working hours, which makes the availability and clarity of digital processes particularly important. Academic and administrative staff are both users and producers of information within the institution's systems, and the quality of those systems directly affects how much time they spend handling routine queries. External stakeholders include IT vendors, the Ministry

of Education, and competing institutions, whose service levels shape what students expect.

In terms of competitive position, the institution has two genuine advantages: its unique program offering in Architecture and Civil Engineering, and its affiliation with an engineering company that guarantees internship placements for students. Its main structural disadvantage is a marketing budget roughly five to eight times smaller than its primary private competitor.

4.2 Identified Pain Points Across the Student Journey

Looking at the student journey end to end, a consistent pattern emerges: the digital tools are there, but they were built around what is convenient for administration, not what is clear to the student or applicant. The clearest example within the enrolled-student phase is exam registration. The final confirmation step of the registration process looks almost identical to the steps that come before it. As a result, many students complete most of the process but do not realize they have not actually finished it. Since the system sends no automatic confirmation, students have no way to check whether their registration went through without contacting the office directly. The practical consequence is that the institution maintains parallel paper records as a backup, which adds work without solving the underlying problem.

Scheduling is another weak point. With seven faculties operating independently, timetable conflicts are common and are typically resolved through informal channels rather than through the system. Student communication with administrative services relies almost entirely on email and in-person visits. There is no centralized notification system and no way for students to track the status of requests they have submitted. Tuition payments are handled through traditional banking channels rather than through an integrated online payment option, which adds an additional friction point compared with what students see at competing institutions.

The same pattern extends to the pre-enrollment phase. Prospective students who submit an online application typically receive no structured follow-up communication, no clarity about what document is missing, and no confirmation of where their application stands. The institution loses approximately 30% of applicants between application and enrollment, and 160 of 720 accredited places remained unfilled in 2024/25. These figures are usually treated as marketing performance indicators, but viewed alongside the administrative pain points they reflect the same underlying issue: a digital service environment that does not communicate with the user proactively. From the user's perspective, the friction encountered during the application process and the friction encountered later during exam registration are symptoms of one continuous problem, not two separate ones.

The SWOT analysis (Table 1) situates these findings within the broader institutional context, framing each item in terms of digital transformation readiness.

Table 1. SWOT analysis framed by digital transformation readiness.

Strengths (capabilities for transformation)	Weaknesses (barriers to transformation)
Existing digital infrastructure as a base to build on; unique program portfolio providing a distinctive value proposition; guaranteed internship placement as a differentiator; central location and flexible payment terms.	Fragmented systems with parallel manual workflows; UX designed around administrative convenience; small marketing budget (~15,000 EUR/year); 160 unfilled enrollment places reflecting both a marketing and an administrative communication gap.
Opportunities (transformation levers)	**Threats (risks to manage)**
Unfilled places as direct revenue potential addressable through better digital communication; alumni network as a low-cost channel; demand for flexible study formats among employed adults; phased transformation feasible within current budget.	Free tuition at public universities; FON dominates IT-related programs; declining pool of secondary-school graduates; data security and compliance risks (PCI DSS, GDPR-equivalent) during digitalization.

5 Proposed Transformation Model

5.1 AS-IS and TO-BE Business Model Canvas

Table 2 presents the institution's current (AS-IS) business model alongside the proposed (TO-BE) redesign. The AS-IS column serves as the diagnostic baseline; the TO-BE column shows the redesign. The most significant changes are concentrated in the Value Proposition, Channels, Customer Relationships, and Key Activities, the elements that directly determine the student experience and where the gap between current and desired state is largest.

Table 2. AS-IS and TO-BE Business Model Canvas comparison.

BMC Element	AS-IS (current)	TO-BE (proposed)
Customer Segments	Domestic high-school graduates; smaller share of working adults; limited international reach.	As current, with stronger focus on the working-adult segment as a growth target.
Value Propositions	Accredited programs incl. unique Architecture and Civil Engineering; internship guarantee; flexible payment; central location.	Programs retained; added: transparent processes, self-service, proactive communication.
Channels	Website; Instagram and Facebook; email; in-person office visits.	Mobile portal with Single Sign-On; push notifications; structured social-media content.

BMC Element	AS-IS (current)	TO-BE (proposed)
Customer Relationships	Reactive across applicant and enrolled phases; no proactive notifications or status tracking.	Proactive; status, deadlines, and alerts communicated before the student asks; five-step applicant follow-up.
Revenue Streams	Tuition fees; some support from affiliated companies.	As current, with higher conversion from improved process and communication.
Key Resources	Academic staff; facilities; legacy IT; social media (~1,590 followers).	Plus redesigned portal, centralized notification system, alumni network.
Key Activities	Teaching and research; admissions; manual processing alongside digital entry.	Process standardization across journey touchpoints; marketing automation; ongoing UX work.
Key Partnerships	Affiliated engineering company; Ministry of Education; accreditation bodies.	Plus payment processor and marketing-automation provider.
Cost Structure	Salaries; facilities; marketing (~15,000 EUR/year); IT maintenance.	Baseline plus ~75,000 EUR phased investment over 18 months.

5.2 Benchmarking Findings

Table 3 summarizes the benchmarking results across the three maturity dimensions. The main takeaway is that the better-performing institutions are not necessarily those with the most features; what distinguishes the international reference institutions is that their systems tell users what is happening and what to do next at every step, while domestic institutions have digitized many processes but still leave the student to guess, ask, or come in person to confirm something.

Table 3. Benchmarking results and applicability.

Institution	Key practice	Applicability
TU Munich (DE)	Application wizard with visible progress steps and real-time registration status.	Redesign the exam registration confirmation as a distinct step; add status tracking (Received–Processing–Confirmed).
KU Leuven (BE)	Automatic conflict detection in course selection; personalized schedule generation.	Introduce a basic conflict check across faculties; reduce informal scheduling coordination.

Institution	Key practice	Applicability
Georgia State (US)	Early-warning system based on multiple risk signals; proactive outreach.	A simpler version: automated alert if a first-year student misses the first exam attempt.
FON (RS)	Centralized portal; automated exam registration; online tuition payment.	Centralize student services into one portal; stop maintaining parallel paper records.
Singidunum (RS)	Clear interface design; instant payment; Single Sign-On.	Apply UX-first design; integrate an online payment option.
Faculty of Architecture (RS)	Information hub portal; limited transactional capability.	Information centralization is useful but insufficient; students need to act, not just read.

The gap between international and domestic examples is not about budget scale. It is about how much the system supports the user, whether it guides them through a process or leaves them to figure it out. These benchmarking findings translated directly into the roadmap priorities: the TU Munich practice of visible progress steps informs the confirmation transparency and status tracking elements; the Singidunum and FON examples shape the online payment integration; and the Georgia State approach forms the basis for the simplified early alert mechanism.

5.3 Transformation Roadmap Across the Student Journey

The transformation roadmap is structured around the student journey rather than along functional lines. The same user moves through the application phase, the enrollment phase, and the recurring administrative cycle, and the digital experience at each step has to follow the same logic. Treating pre-enrollment communication as a marketing function and post-enrollment processes as an administrative function is precisely what produces the inconsistencies described in Section 4. The roadmap therefore intervenes at the points where the journey is most visibly broken, regardless of which functional category the intervention has traditionally belonged to.

The first phase (0–6 months, estimated cost around 25,000 EUR) addresses the most acute friction points. At the point of application, a five-step automated follow-up sequence is introduced, triggered by online application submission: confirmation with a documentation checklist on day zero, a reminder on day three, an SMS invitation to a consultation on day seven, information about the installment payment option on day fourteen, and a final reminder before the enrollment deadline on day thirty. This addresses the approximately 30% of applicants who currently do not complete enrollment. Within the same phase, the exam registration confirmation step is redesigned with explicit visual confirmation and automatic email notification, the scheduling system is centralized across the seven faculties, and basic push notifications replace email as the primary channel for time-sensitive communication. The applicant follow-up and the

exam registration redesign share the same design pattern, which is what makes them components of one transformation rather than two separate projects.

The second phase (6–18 months, around 50,000 EUR) extends the same logic. A mobile-compatible student portal with Single Sign-On consolidates the touchpoints into a single interface, a basic early-alert mechanism notifies students at risk of falling behind, and an integrated online payment module compliant with PCI DSS requirements completes the financial side of the journey. The marketing channel mix is rebalanced based on conversion data from phase one: Instagram Ads at 40% of the 15,000 EUR annual budget targeting the 18–22 group through video content, Google Ads at 20%, content production at 15%, ambassador and influencer activity at 10%, Facebook Ads aimed at parents at 10%, and email and SMS automation at 5%. The two underperforming programs, Ecology (around 40% filled) and Civil Engineering (around 67%), receive dedicated campaign attention. The third phase (18–36 months) adds analytics linking enrollment, marketing, and administrative performance in a single view, automated FAQ support, and an alumni platform that doubles as a low-cost promotional channel.

The minimum viable product for immediate implementation combines the redesigned exam registration with the five-step applicant follow-up sequence, since these are the two interventions where the design principle is identical and the impact is measurable on both sides of the enrollment threshold. The expected outcomes are a reduction in failed exam registrations to under 2% of submissions, a significant drop in registration-related queries, and a measurable improvement in the application-to-enrollment conversion rate for the 2026/27 cycle. These metrics tie the success of the MVP to both administrative efficiency and enrollment outcomes, consistent with the framing of enrollment conversion as an indicator of digital transformation effectiveness introduced in Section 1.

6 Discussion

The case shows a pattern that is not unusual for private universities of this size in the region. Digital tools were introduced over time, usually in response to immediate needs or competitive pressure, without a systematic look at whether they actually made things easier for the student at any particular stage of the journey. The result is that the institution is technically digitalized in many areas, but users still regularly need to contact someone or come in person to resolve issues that the system should handle on its own.

The benchmarking comparison points to something specific: the difference between well-functioning and poorly functioning digital environments is mostly not about how many features are available. It is about whether the system keeps the user informed and reduces uncertainty at each touchpoint. A student who knows their registration went through does not call the office. An applicant who receives a clear next-step prompt after submitting an application is more likely to complete enrollment than one who hears nothing for two weeks. Fixing this does not require a new system; it requires a decision to look at the existing journey from the user's side and make targeted adjustments.

This is also why treating the 30% application dropout rate as a marketing problem alone misses the structure of what is going on. The applicant who drops off and the enrolled student who calls the office about a registration are reacting to the same kind of failure: a digital environment that does not tell the user what is happening or what to do next. Improving one stage of the journey without the others leaves a significant part of the problem unresolved.

This study has limitations worth acknowledging. It is based on a single institution, so the findings cannot be assumed to apply everywhere without further investigation. The roadmap is a proposal, not an account of implemented changes; its actual effects remain to be measured against the indicators defined above. A useful next step would be to track outcomes after implementation and to compare the experience with other private universities in the Western Balkans.

7 Conclusion

Digital transformation in a resource-constrained private university does not have to start with a large investment or a complete system replacement. What such an institution needs in the short term is not more technology but better use of what already exists, beginning with the points along the student journey where the failure rate is both measurable and avoidable: the applicant follow-up sequence on one end and the exam registration confirmation on the other.

The Business Model Canvas proved useful here not as a planning tool for a new product, but as a way to make visible what the institution currently offers and where the gap between the intended and the actual user experience lies. The AS-IS to TO-BE comparison shows that the most consequential changes are concentrated in a few canvas elements, which makes the transformation more focused and more affordable than an undirected redesign would be.

Structuring the roadmap around the student journey rather than along functional lines is what holds the model together. The same communication weakness that frustrates an enrolled student also costs the institution applicants, and treating these as separate problems is exactly what produces the inconsistencies the case study identified. A coordinated approach across the journey is more efficient than addressing any single stage in isolation.

The roadmap and the specific MVP proposed here are designed with the constraints of a mid-sized private university in mind. The structural characteristics that shaped this approach, namely a limited budget, several faculties, a significant share of working-adult students, and competition from larger public institutions, are not unique to UNT, which is why the conclusions may be relevant for similar institutions in the region.

Acknowledgments. During the preparation of this manuscript, the author used generative AI tools (large language model-based assistants) for improving the language, readability, and structure of the text. The author reviewed and edited the content as needed and takes full responsibility for the accuracy and integrity of the final version.

Disclosure of Interests. The author has no competing interests to declare that are relevant to the content of this article.

References

1. Vial, G.: Understanding digital transformation: A review and a research agenda. J. Strateg. Inf. Syst. 28(2), 118–144 (2019). https://doi.org/10.1016/j.jsis.2019.01.003
2. Gong, C., Ribiere, V.: Developing a unified definition of digital transformation. Technovation 102, 102217 (2021). https://doi.org/10.1016/j.technovation.2020.102217
3. Sandkuhl, K., Lehmann, H.: Digital transformation in higher education – the role of enterprise architectures and portals. In: Rossmann, A., Zimmermann, A. (eds.) Digital Enterprise Computing (DEC 2017), pp. 49–60. Gesellschaft für Informatik, Bonn (2017)
4. Osterwalder, A., Pigneur, Y.: Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers. John Wiley & Sons, Hoboken (2010)
5. Stensaker, B., Frølich, N., Huisman, J., Waagene, E., Scordato, L., Pimentel Bótas, P.C.: Factors affecting strategic change in higher education. J. Strateg. Manag. 7(2), 193–207 (2014). https://doi.org/10.1108/JSMA-12-2012-0066
6. Kopp, M., Gröblinger, O., Adams, S.: Five common assumptions that prevent digital transformation at higher education institutions. In: INTED2019 Proceedings, pp. 1448–1457. IATED, Valencia (2019). https://doi.org/10.21125/inted.2019.0445
7. Hénard, F., Mitterle, A.: Governance and Quality Guidelines in Higher Education: A Review of Governance Arrangements and Quality Assurance Guidelines. Organisation for Economic Co-operation and Development (OECD), Paris (2010)
8. Yin, R.K.: Case Study Research and Applications: Design and Methods, 6th edn. SAGE Publications, Thousand Oaks (2018)
9. Malešević, A.: Perspective chapter: digitization of traditional higher education touchpoints. In: Kang, K., Namisango, F. (eds.) E-Service Digital Innovation. IntechOpen, London (2022). https://doi.org/10.5772/intechopen.105926
10. Henriette, E., Feki, M., Boughzala, I.: The shape of digital transformation: a systematic literature review. In: Pucihar, A. (ed.) MCIS 2015 Proceedings, pp. 431–443. AIS Electronic Library, Samos (2015)
11. Milošević, I., Rakić, A., Arsić, S., Stojanović, A., Nikolić, I., Đorđević, P.: Model for considering the propensity of students to accept m-learning. Manag. J. Sustain. Bus. Manag. Solut. Emerg. Econ. 27(1), 13–24 (2022). https://doi.org/10.7595/management.fon.2020.0021
12. Tornjanski, V., Petrović, D., Nešić, S.: Effectiveness of knowledge transfer between project team members in digitally disrupted organizations. Manag. J. Sustain. Bus. Manag. Solut. Emerg. Econ. 25(2), 1–14 (2020). https://doi.org/10.7595/management.fon.2018.0034

Original Research Article *https://doi.org/10.5281/zenodo.23476122*

Application of 3D Indoor Navigation Systems in Academic Environments

Danica Živković[1][0009-0008-5345-4168], Dušan Kostić[2][0009-0000-3432-6074] and Marijana Despotović-Zrakić[3][0000-0002-6458-1575]

[1] University of Belgrade, Faculty of Organizational Sciences, Serbia
[2] University of Belgrade, Faculty of Organizational Sciences, Serbia
[3] University of Belgrade, Faculty of Organizational Sciences, Serbia

Abstract. This paper presents the theoretical aspects of development and application of a 3D navigation system within the Unity Engine, designed to optimize indoor orientation through complex academic infrastructures, specifically the Faculty of Organizational Sciences. This research explores the impact of 3D navigation systems and the benefit they provide to users, as well as the synergy between game engine technologies and functional spatial visualization. The study details a methodology for architectural modeling and user-oriented navigation logic that assists students and guests of the faculty. The proposed system aims to reduce factors of spatial disorientation and anxiety within large-scale buildings through a structured implementation plan. The study concludes by discussing the theoretical aspects of using rendering engines for functional navigation, providing a base for future development and system deployment.

Keywords: 3D modeling, Navigation system, Academic infrastructure, Unity engine.

1 Introduction

As humans interact with their surroundings, they are required to navigate through unfamiliar infrastructures and complex environments. Through this experience, individuals are expected to gather spatial experience and form internalized cognitive maps. These mental blueprints allow individuals to determine optimal routes and create navigational patterns that become semi-automated over time. Verma and Mettler [1] mention how this capability is not only merely a spatial navigation skill for physical movement but a fundamental cognitive process applied in high-stakes tasks, such as surgery, where spatial awareness is directly connected to precision and success. However, the initial mention of "wayfinding" in a new environment remains an exhausting experience characterized by cognitive load and potential disorientation.

The primary challenge addressed by this research is the lack of immediate, accessible and precise information regarding complex building infrastructures. This absence of information leads to a phenomenon known as "spatial anxiety", where guests or

newcomers spend a significant amount of time physically searching for rooms or asking for verbal direction assistance from others. It's highlighted how navigation to the place of work is often typical, while navigating in a novel environment can be stressful and mentally draining [2]. The delays that occur daily through academic or professional settings result in productivity loss, diminished user experience and anxiety [3].

The goal of this research paper is the optimization of navigational processes within the Faculty of Organizational Sciences (later referred to as FON) in Belgrade, Serbia. The faculty building represents a particularly convoluted infrastructure, having undergone recent expansions that divided the infrastructure complex, now referred to as the "old" and "new" building. For first-year students and international guests, the building's layout doesn't often seem intuitive. The complexity of multiple levels, connected or prolonged corridors, and non-linear room numbering system frequently leads to disorientation, therefore, finding specific halls or amphitheaters becomes a recurring event during the first few weeks of the academic year.

Traditional solutions, such as static 2D floor plans, often fail to bridge the gap between a flat map and three-dimensional reality in the building. This research implies that the solution to tightening the gap lies in the intersection of game engine technology and functional utility. By using the high-definition rendering and movement implementation capabilities of the Unity engine, it is possible to create a digital copy of the faculty that mirrors the user's physical experience.

The goal of this paper is to establish comprehensive theoretical framework and technical design for a 3D navigation system. Beyond a simple visual representation, the study explores the technical possibilities and spatial logic, pathfinding algorithms, and the expansion of these systems into bigger initiatives. By introducing a conceptual blueprint for a functional application, this research aims to provide a scalable solution to the mentioned problem that enables users to easily navigate through the buildings of Faculty of Organizational Sciences and reduce the time spent on wayfinding and appearing anxiety, as well as provide exact information.

The paper is organized as follows: the second section offers an insight into similar works regarding 3D navigation systems, modeling principles and their importance in various fields. The third section further explains the system design and technologies planned to use during the development phase of the application. The fourth section delves into the implementation methods of the system. The fifth section describes the future work regarding this topic. The last section of this paper concludes everything previously said and provides an insight of important theoretical aspects for the development of a 3D navigation system.

2 Literature Review

As the majority of the world's population owns a smartphone and has an internet connection, it is difficult to imagine daily life without the services these devices provide. As of early 2026, there are approximately 5.7 billion unique smartphone users[4]. With the proliferation of mobile technology, "digital positioning systems" grew along, representing frameworks that provide real-time position estimations [5].

While Global Positioning Systems (GPS) have revolutionized outdoor navigation, their precision worsens within indoor environments due to signal attenuation caused by construction materials. Apart from Global Positioning Systems, Indoor Positioning Systems (IPS) utilize localized technologies such as Wi-Fi fingerprinting, Bluetooth Low Energy (BLE) beacons and Ultra-Wideband (UWB) to estimate a user's location inside a building [5].

However, traditional IPS solutions often face a barrier during the deployment phase. These systems require beacons to be installed, maintained and calibrated, while the accuracy of the positioning is sensitive to environmental changes. Therefore, knowing an individual's precise coordinates on a 2D map does not always translate to a successful navigation in a building.

This paper identifies a distinction between positioning, the mathematical process of locating a device, and wayfinding as the cognitive task of a human reaching a desired destination. By focusing on a Visual Navigation System, developed in Unity Engine through the ProBuilder Add-on and C# scripts, this research avoids hardware dependent tracking and prioritizes spatial context and visual recognition. This approach allows the application to be utilized by a student while at home to plan a route or by a guest standing in the faculty halls, without the requirement of connecting to internal building sensors.

The importance of 3D-based navigation is driven by the limitations of traditional 2D floor plans. Academic research into spatial cognition suggests that humans process 3D environments more intuitively because they mirror natural perspective. The to-be developed 3D navigation application aims for reduction of cognitive load, pre-visit orientation and universal accessibility

2.1 3D Environment in Education

Planning and maintaining urban environments often require complex efforts that encompass various urban subsystems, including transportation and mobility, urban energy, water supply, as well as buildings and infrastructure [6]. While the original 3D virtual model of real-world objects serve primarily to support advanced manufacture and analytics, another study proposes the idea of Covenant University location navigation system built on the Android platform. A 3-D map of Covenant University was packaged into a location navigation system and was implemented using Unreal Engine, enabling the free-roam navigator with no downtime on access and little to no data loss [7].

Another great example of 3D environments in education is a digital twin of the West Cambridge site of the University of Cambridge in the United Kingdom. The mentioned project goes through the process of creating a digital twin, as well as research efforts aimed at the created system [8]. This system highlights the bridge built to reduce the gap between human relationships with buildings or cities.

Dealing with the same problems of large and complex buildings, the campus of Al-Zaytoonah University of Jordan developed Virtual tours as a campus guide. The focus of such project is finding the location of an office, a lecture hall and a service center as

it is not an easy task for most visitors [9]. Their study showed user satisfaction with the tours and the efficiency of using their website without special virtual reality devices.

2.2 Unity in Non-Gaming Context

While known for its video game development purposes, Unity Engine can be used for non-gaming purposes due to their technical integration with various IDEs and programming languages. Haas [10] mentioned how architects can prototype ideas, artists can create interactive art installations and how researchers can use engines for data visualization. Furthermore, Unity's non-gaming revenue grew over 35% as of January 18th 2026 [11].

Another interesting example of using Unity Engine for environment building is the project of third-year architecture students in Belgium who redesigned the public square in the city of Leuven, made it interactive in Unity game engine and created augmented reality (AR) application using smartphones [12]. Results of this study show how the qualities of this architectural education approach simulate user experience.

3 Conceptual Framework and Environmental Analysis

The importance of System Design lies within its impact on the rest of the application segments that are waiting to be developed as it affects the connected processes and functionalities and clears the development path. The process of designing the system refers to the different techniques and approaches to develop systems more effectively and efficiently [13].

It is important to design a system based on behaviors individuals express during wayfinding. Humans acquire spatial knowledge about their environment in three stages: the landmark knowledge is acquired first, then route knowledge and finally survey knowledge. The landmark knowledge is represents memorizing the appearance of distinct objects along the way. Route knowledge represents learning the sequence of direction choices at intersections. Lastly, survey knowledge represents forming a cognitive map of the environment based on everything the particular individual has seen [14].

The proposed 3D navigation project is architected specifically to support these three stages. Unlike traditional 2D maps, which force the user to mentally rotate the flat image, the 3D perspective provided by this application offers a high-fidelity visual representation that makes the user more "present". By emphasizing visual landmarks through realistic rendering, the system is expected to reduce cognitive load to transition through these three stages. Steck and Mallot [15] mention how observers use local and global landmarks to define direction and position. This 3D navigation project focuses merely on the local landmarks since the system provides only the interior of the building.

Furthermore, the design of this system addresses the "environmental stress" often associated with navigating through complex academic infrastructures. By providing an interactive, exploratory interface that can be accessed outside of the physical buildings,

the system should allow students and guests to build their landmark before reaching the faculty. This strategic design choice shifts the application from being a simple reference tool to becoming an active participant in a learning progress.

The 3D navigating application proposes the following features:

- Complete replication of both of the FON buildings
- Ability to move through the building halls
- Searching for a specific room, hall or amphitheater
- Following a path that leads to a room, hall or amphitheater
- Resetting positions in case of bugs

These features will add additional support to users during the process of building exploration, enhancing user experience and allowing individuals to quickly access targeted data.

3.1 Environment Creation

The process of environment creation begins with the translation of 2D floor plans into three-dimensional space. To ensure the 3D model serves as a faithful copy, the development follows the process of mapping the walls, floors, corridors and doorways in the Unity scene from the 2D blueprint. By importing these 2D maps as reference planes within the Unity editor, the created environment provides a layout that reduces space for human error. Using the ProBuilder add-on, the creation of objects is simplified, allowing seamless transition from a flat representation to a volumetric structure. Apart from looking good, the created assets and objects must work with the software being user to create the application [16].

The technical workflow for level creation is divided into faculty floors. Each floor of the Faculty of Organizational Sciences is treated as a distinct level, modeled with 0.5-unit snapping grid. This snapping logic is important when connecting gaps between objects, as it ensures that the stairwells and elevator shafts align with the vertical axis. Such gaps between objects often result in light leaks or script bugs.

To enhance the utility of the system for diverse user groups, the environment is designed with two distinct entry points, which are the main entrance of the “old” building and the main entrance of the “new” building. In the system design, these serve as the primary spawn points for both the user model and the navigation camera. The navigation camera follows behind the user, enabling the third person perspective and providing the user with more insight of the building interior. By allowing the users to choose their starting location, the application acknowledges that students and guests often arrive with incomplete information about a room or hall. This method of creating the environment ensures the model is not just a static gallery, but a functional tool for real orientation.

3.2 Navigation Logic

The selection of the Unity engine and the ProBuilder add-on is driven by the need for a high-performance and interactive development environment that bridges the gap between architectural modeling and interactive simulation. Unity provides the real-time physics and scripting capabilities that are necessary to transform a 3D model into a navigation tool. This technology enables the ability to attach behaviors to digital objects through C# scripting. In this system, the user model, which the real user controls, is defined by the movement script. This script handles the translation of user inputs into physical actions, such as walking, running and rotating, ensuring that the movement feels responsive and grounded within the faculty. The system provides a sense of presence through different states, such as idle, walking and sprinting.

To further support the acquisition of survey knowledge, the application utilizes a third-person camera perspective. This camera logic is programmed to follow the user model from a specific offset, providing a wider field of view that the first-person perspective would allow. This design choice is intentional, allowing the user to judge distances and see around corners. Using a third-person perspective facilitates a better understanding of the spatial relationship between the user's current position and the surrounding landmarks such as room entries or corridor junctions. The camera's rotation and dampening prevent sudden movements and ensure the orientation process remains comfortable.

The broader navigation logic also includes the underlying Navigation Mesh system, which represents the walkable surface of the buildings. This system is important because it defines the paths the user can take and preventing the user to walk through walls and fall through the floor. The integration of ProBuilder allows the developer to mark specific objects during the modeling phase. Using navigation mesh, the process of future pathfinding becomes simple.

4 Implementation

The development of the 3D navigation application is rooted in the specific spatial challenges posed by the Faculty of Organizational Sciences in Belgrade. The faculty building represents a complex architectural case study characterized by the integration of two distinct construction phases which are the original "old" building and a modern "new" building. This structural duality has created a convoluted internal layout where corridors, floor levels, and stairs do not always follow a linear or intuitive progression. The implementation phase focuses on the translation of theoretical design into a functional Unity application. This process is divided into two main parts which are the structural reconstruction of the environment and the development of the navigation system.

What is specific to this problem is the spatial confusion experienced by the primary user groups which are first-year students and guests. The building's room numbering system and the intersection of corridors often lead to a high degree of spatial disorientation. Traditional 2D floor plans aren't enough to bridge the gap between flat orientation and the three-dimensional reality of the building. Consequently, users

frequently suffer from spatial anxiety, spending a significant amount of academic time physically searching for specific amphitheaters or halls.
This application is being developed as a solution to these challenges. Unlike traditional positioning systems that require a user to be physically present and connected to local hardware, this Unity-based tool allows for pre-visit orientation. By creating a digital model of the faculty, the application provides an environment for users to build landmark and route knowledge before they even enter the physical infrastructure, therefore increasing institutional accessibility and reducing time lost to wayfinding.

The foundation of the application is the conversion of 2D architectural sketches into a 3D environment. The 2D sketches were imported into Unity as reference textures on a horizontal plane. The modeling process utilized a modular approach, where walls, floors and ceilings were created using ProBuilder add-on.

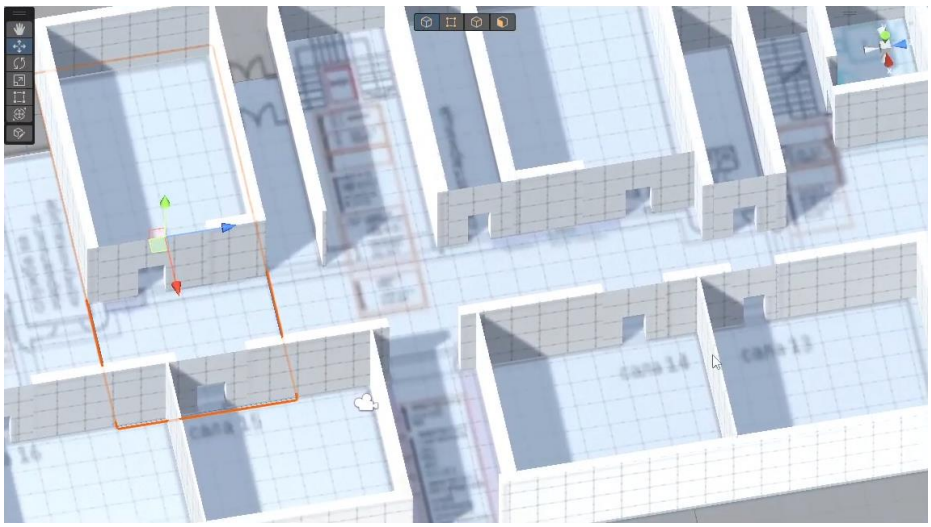


Fig. 1. The creation of a 3D model based on a 2D blueprint

A critical technical constraint implemented during this phase was the strict adherence to a grid-based snapping system. By setting the grid to 0.5-unit increments, the model ensures that all vertices and edges align perfectly. This precision is not just aesthetic, but a functional requirement to ensure the entire model becomes "connected". Without this snapping precision, small gaps between meshes could cause the Unity engine to create light gaps and errors in movement.

The movement is handled by a custom C# controller attached to the user representative object. This script utilizes a state-based system to manage player actions. Users can toggle between walking and running states, with speed variables adjusted to represent real human pace relative to the building scale.

The third-person camera allows the user to rotate their view freely while maintaining consistent distance from player object.

Fig. 2. Third-person camera following the user model

To mitigate potential disorientation, there is a "reset position" function implemented in the UI, allowing the user to quickly get back to one of the entrance spawn points. As mentioned, this feature enables users to switch positions without feeling frustrated in case of any unrecognized errors.

The UI is designed to be minimalistic, but functional, ensuring that the 3D view remains the primary sense of information. The application will have one simple UI panel providing the user with options of resetting the position, switching to the adjacent building and searching for a specific room. Once the destination is selected, the system calculates the shortest path and visualizes the path towards the selected room. This path is visualized in the 3D space, guiding the user through corridors and between buildings.

Navigating between the old and new buildings requires specific transitions. Rather than using loading screens that break immersion, the implementation utilizes position transitions. By pressing the building switch button, the user calls for a function which changes the current position of the user model object.

Fig. 3. UI panel with four options

To clarify the scope of this work, the current implementation of the system focuses on the foundational mechanics of the environment, specifically the first versions of the "old" building 3D model, movement scripts, third-person camera control and UI panels. While the broader architecture is presented as a theoretical framework for the future, these core components have been successfully deployed and tested to validate the physics and path visualization. Furthermore, it should be noted that the figures presented through this paper are included for representational purposes to illustrate the conceptual workflow and will be further refined in later iterations.

5 Discussion and Future Work

While the application is currently in the primary development phase, the architectural and logical framework in this paper provides a significant foundation for enhancing institutional wayfinding. The transition from traditional, hardware-dependent indoor positioning to a visual 3D navigation model addresses the barrier that sometimes prevents universities form deploying navigation tools. As Peng mentions [17], social-technical systems such as virtual environments can be deployed to support institutional learning and innovation in campus planning. The Faculty of Organizational Sciences can offer tools that are equally valuable for pre-visit orientation at home as it is for real-time guidance.

The potential effectiveness of this system can be projected through the lens of cognitive load reduction. Based on the "Landmark-Route-Survey" hierarchy discussed in the Section 2, the outcome of this application is a faster acquisition of survey knowledge among the user body. In the absence of empirical results, theoretical analysis suggests that the third-person perspective and the fidelity of the ProBuilder objects will reduce snapping grid errors. Qualitative feedback from early-stage

conceptual reviews suggests that users prioritize visual recognition, confirming that the high-precision modeling approach is aligned with user needs.

Future work will focus on the empirical evaluation of the system through structured user testing. Once the implementation is finalized, a series of metrics will be conducted to measure the efficiency and satisfaction of the interface, as well as test the reduction of users experiencing spatial anxiety. Furthermore, the nature of the Unity environment allows for the potential integration of Augmented Reality (AR) features. By implementing the live camera feed, the system could evolve from a screen-based 3D guide into an immersive AR assistant.

Additionally, the expansion of the system could include dynamic data integration. This would involve linking the application logic to the faculty's schedule database. In such a scenario, a student would not only find the location of a hall, but also receive information about class changes or room availability. Finally, the goal of this research is to move beyond a standalone application and toward a comprehensive 3D navigation system of the Faculty of Organizational Sciences.

6 Conclusion

This paper has presented a comprehensive theoretical and technical framework for the development of a 3D navigation system within the Unity environment, specifically focused on the Faculty of Organizational Sciences in Belgrade. By shifting the focus from hardware-dependent positioning to a visual wayfinding approach, the proposed system addresses the complexities of modern academic infrastructures. The integration of high-precision ProBuilder modeling with a C# movement logic ensures that the application is not just a set of images, but a functional tool designed to support human spatial cognition. Through the systematic translation of 2D blueprints into a 3D digital model, this research establishes a scalable methodology for institutional orientation that prioritizes accessibility and user experience.

Ultimately, the importance of this work lies in its potential to migrate the cognitive load and spatial anxiety experiences by students and guests. The established system design principles, ranging from strict grid-snapping to position reset, provide a blueprint for future implementation and empirical testing. As universities continue to grow in structural complexity, the adoption of game engine technologies for functional navigation represents a necessary evolution in campus management. This research serves as a foundational step toward a fully realized digital ecosystem, offering a sustainable and flexible solution for navigating the complex environments for the future.

Acknowledgments. This research was supported by the University of Belgrade - Faculty of Organizational Sciences and, in part, by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia through institutional funding (grant number: 200151).

Disclosure of Interests: The authors declare no conflict of interests.

Declaration of AI Use: During the preparation of this manuscript, AI-assisted tools were user for language editing, readability improvement and structural formatting. The scientific content, methodology, implementation and results were developed, reviewed and verified by the authors. The authors take full responsibility for the integrity and final content of the manuscript.

References

[1] A. Verma and B. Mettler, "Investigating Human Learning and Decision-Making in Navigation of Unknown Environments," *IFAC-PapersOnLine*, vol. 49, no. 32, pp. 113–118, Jan. 2016, doi: 10.1016/J.IFACOL.2016.12.199.

[2] Boone, A. P., Bullock, T., MacLean, M. H., Santander, T., Raymer, J., Stuber, A., Jimmons, L., Okafor, G. N., Grafton, S. T., Miller, M. B., Giesbrecht, B., & Hegarty, M. (2024). Resilience of navigation strategy and efficiency to the impact of acute stress. Spatial Cognition & Computation, 24(3), 195–226. https://doi.org/10.1080/13875868.2024.2321920

[3] Informatics, U., Xu, H., Omitaomu, F., Sabri, S., Zlatanova, S., Li, X., & Song, Y. (2024). Leveraging generative AI for urban digital twins: a scoping review on the autonomous generation of urban data, scenarios, designs, and 3D city models for smart city advancement. Urban Informatics 2024 3:1, 3(1), 29-. https://doi.org/10.1007/S44212-024-00060-W

[4] "How Many People Have Smartphones Worldwide (2026)." Accessed: Apr. 25, 2026. [Online]. Available: https://www.bankmycell.com/blog/how-many-phones-are-in-the-world

[5] G. M. Mendoza-Silva, J. Torres-Sospedra, and J. Huerta, "A Meta-Review of Indoor Positioning Systems," *Sensors 2019, Vol. 19, Page 4507*, vol. 19, no. 20, p. 4507, Oct. 2019, doi: 10.3390/S19204507.

[6] M. Fedorowicz, J. Schilling, E. Bramhall, B. Bieretz, Y. Su, and K. Steven Brown, "Leveraging the Built Environment for Health Equity Promising Interventions for Small and Medium-Size Cities," 2020.

[7] M. Adebiyi *et al.*, "A 3-D MODEL OF AN INSTITUTIONAL LOCATION NAVIGATION SYSTEM (NAVILOC) (A CASE STUDY OF COVENANT UNIVERSITY)," 2019.

[8] Q. Lu *et al.*, "Developing a Digital Twin at Building and City Levels: Case Study of West Cambridge Campus," *Journal of Management in Engineering*, vol. 36, no. 3, p. 05020004, Mar. 2020, doi: 10.1061/(ASCE)ME.1943-5479.0000763.

[9] M. Salah, A. Abdalla, M. Abdallah, A. A. Mazhar, B. Alokush, and I. Jebril, "Using Virtual Tours as a University Campus Guide: Al-Zaytoonah University Case Study," *Information Sciences Letters An International Journal*, vol. 12, no. 9, p. 2961, 2023, doi: 10.18576/isl/120906.

[10] J. Haas, "A History of the Unity Game Engine An Interactive Qualifying Project Submitted to the Faculty of WORCESTER POLYTECHNIC INSTITUTE in partial fulfillment of the requirements for graduation".

[11] "What is Customer Demographics and Target Market of Unity Software Company? – PortersFiveForce.com." Accessed: Apr. 25, 2026. [Online]. Available: https://portersfiveforce.com/blogs/target-market/unity

[12] E. Nofal, "Studying User Experience in Architectural Education Using Unity Game Engine," *Man and Place*, pp. 33–43, 2024, doi: 10.1007/978-3-031-49903-6_3.

[13] Alan. Dennis, B. Haley. Wixom, and R. M. . Roth, "Systems analysis and design," 2009, Accessed: Apr. 25, 2026. [Online]. Available: https://books.google.com/books/about/Systems_Analysis_and_Design.html?id=gFkYqK84rVgC

[14] K. Kim and O. Bock, "Acquisition of landmark, route, and survey knowledge in a wayfinding task: in stages or in parallel?," *Psychological Research 2020 85:5*, vol. 85, no. 5, pp. 2098–2106, Jul. 2020, doi: 10.1007/S00426-020-01384-3.

[15] S. D. Steck and H. A. Mallot, "The role of global and local landmarks in virtual environment navigation," *Presence: Teleoperators and Virtual Environments*, vol. 9, no. 1, pp. 69–83, 2000, doi: 10.1162/105474600566628.

[16] L. Ahearn, *3D game environments: create professional 3D game worlds*. CRC Press, 2017. doi: 10.1201/9781315687445.

[17] C. P.-I. J. of A. Computing and undefined 2011, "uCampus: Can an open source 3D interactive virtual campus modelling platform support institutional learning and innovation?," *journals.sagepub.comC PengInternational Journal of Architectural Computing, 2011•journals.sagepub.com*, vol. 9, no. 3, pp. 303–324, Sep. 2011, doi: 10.1260/1478-0771.9.3.303.

Original Research Article *https://doi.org/10.5281/zenodo.23476172*

Toward a Secure-by-Design Higher Education Ecosystem: Decentralized Identity, Privacy, and Trust

Tamara Naumović[1][0000-0001-9849-7665], Danica Živković[1][0009-0008-5345-4168], Marija Janković[2][0000-0002-3600-3014], Dušan Barać[1][0000-0003-0517-2525]

[1] University of Belgrade – Faculty of Organizational Sciences, Belgrade 11000, Serbia
[2] Centre for Research and Technology Hellas, Information Technologies Institute, Thessaloniki, Greece
tamara.naumovic@fon.bg.ac.rs

Abstract. The rapid digital transformation of higher education institutions necessitates secure, interoperable, and student-centered systems for managing micro-credentials. Traditional centralized infrastructures are characterized by institutional data silos, inefficient manual verification procedures, and significant privacy risks. To address these structural gaps, this paper proposes a five-layer secure-by-design micro-credential platform designed as a foundational upgrade to the European Blockchain Services Infrastructure (EBSI). The proposed architecture leverages principles of Self-Sovereign Identity and Decentralized Identifiers to reposition the credential earner as the primary controller of their academic records. Key innovations include a secure oracle layer that mitigates the validation inconsistency gap through multi-source validation, and a privacy-preserving verification layer utilizing zero-knowledge proofs. This integration enables selective disclosure, allowing earners to prove specific academic claims without revealing sensitive personal data. This framework provides a scalable blueprint for a digital ecosystem that balances institutional trust with individual autonomy.

Keywords: secure-by-design, zero-knowledge proof, decentralized identifiers, EBSI, micro-credentials.

1 Introduction

The growing importance of micro-credentials (MCs) in higher education reflects broader changes in learning, work, and credentialing [1]–[3]. While MCs allow learners to document specific skill sets more granularly than traditional degrees, their effective management remains obstructed by rigid infrastructures [4]. Current centralized credentialing systems are characterized by institutional data silos, low interoperability, and inefficient manual verification procedures, which on average take up to 12 days [2], [5]. Furthermore, reliance on centralized servers creates "single points of failure" (SPOF) that expose sensitive student records to data breaches and credential fraud [3], [6].

Emerging digital trust infrastructures such as blockchain technology, Decentralized Identifiers (DIDs), and Verifiable Credentials (VCs), offer a paradigm shift toward Self-Sovereign Identity (SSI), where learners maintain primary control over their academic records [2], [6], [7]. However, existing blockchain implementations often remain technically partial, failing to address critical security gaps in relation to data. Unreliable source data enters the ledger, or does not align with complex GDPR privacy requirements [2], [4]. This underscores the need for ecosystems where security and privacy are not reactive measures but foundational design principles.

Against this background, this paper proposes a secure-by-design framework for micro-credential management. The study focuses on integrating advanced privacy-preserving layers, such as Zero-Knowledge Proofs (ZKP), into interoperable infrastructures to support student autonomy and institutional trust. Particular attention is given to the European Blockchain Services Infrastructure (EBSI) [8], which provides a cross-border "governmental trust anchor" for credential exchange.

The contribution of this paper is twofold. First, it identifies systemic limitations in current HEI digital ecosystems, from verification bottlenecks to privacy risks. Second, it proposes a five-layer secure-by-design micro-credential platform that extends EBSI's core services with a secure Oracle layer and ZKP-based selective disclosure mechanisms

The remainder of the paper is structured as follows. Section 2 presents the theoretical background, covering secure-by-design and privacy-by-design principles, self-sovereign identity, decentralized identifiers, verifiable credentials and zero-knowledge proofs. Section 3 reviews current solutions and identifies their main limitations in higher education credentialing. Section 4 introduces the proposed secure-by-design micro-credential platform. Section 5 discusses how the proposed architecture addresses the key gaps identified in the literature and related work. Finally, Section 6 concludes the paper and outlines implications for future research.

2 Theoretical Background

The digital transformation of higher education institutions (HEI) has intensified the need for secure, interoperable, and privacy-aware infrastructures for managing learner identities, academic records, and micro-credentials [1], [9], [10]. Traditional centralized systems are often characterized by data silos, limited interoperability, manual verification procedures, and exposure to credential fraud, identity misuse, and single points of failure [2], [5], [11], [12]. In such settings, security can no longer be treated as an additional technical layer, but rather as a foundational design principle embedded into the architecture of the system itself [2], [13].

The secure-by-design approach assumes that confidentiality, integrity, authentication, authorization, and auditability should be built into the system from the earliest stages of design [14]–[16]. In the context of higher education, this implies the development of infrastructures that reduce dependence on vulnerable centralized repositories, support tamper-evident operations, and ensure that only authorized actors can issue, access, validate, or revoke academic credentials [4]. Closely related to this is privacy-

by-design, which emphasizes data minimization, purpose limitation, and privacy-preserving data processing as core system requirements rather than ex post compliance measures [4], [10], [16]. This logic is particularly relevant in educational ecosystems, where institutions process highly sensitive personal, academic, and behavioral data [11], [13], [17]. A major challenge in blockchain-based educational systems lies in the tension between immutability and data protection requirements. Having in mind that HEIs manage sensitive personal and academic data that fall under legal frameworks, such as the GDPR Article 25, [18], it is necessary to ensure that only limited data needed for a specific purpose is collected and stored on-chain [19], [20].

This also includes architectural choices that limit unnecessary exposure of student data, support role-based interactions between stakeholders [1], and provide verifiable yet privacy-respecting mechanisms of trust. This creates the basis for credential infrastructures in which security, transparency, and privacy are not competing objectives, but mutually reinforcing design goals [16].

2.1 Self-Sovereign Identity as a Learner-Centric Identity Paradigm

Self-Sovereign Identity (SSI) emerged as a response to the limitations of centralized and federated identity systems, in which user identity is largely controlled by external authorities, platforms, or institutional intermediaries [21]. In such systems, individuals often lack visibility into how their data is stored, shared, or reused, and have limited control over the disclosure of personal information [13]. SSI introduces a different paradigm in which the individual becomes the primary controller of digital identity data and related credentials [22], [23].

From the perspective of higher education, SSI is particularly important because it repositions the earner from a passive subject of institutional identity management to an active participant. Instead of requesting repeated manual confirmation from the issuing institution, the earner can store and present credentials in a controlled and portable way [13]. This is consistent with the “trust triangle” involving issuer, holder (earner), and verifier [8], [17], [24].

Such a model is especially relevant in contemporary higher education ecosystems characterized by mobility, lifelong learning, stackable credentials, and cross-institutional recognition [1], [25], [26]. In these contexts, SSI provides a conceptual foundation for learner-centered control, identity portability, and selective disclosure, thereby supporting both institutional trust and student autonomy [3], [9], [11].

2.2 Decentralized Identifiers and Verifiable Credentials as the Technical Foundation of Trust

The SSI paradigm is operationalized primarily through Decentralized Identifiers (DIDs) and Verifiable Credentials (VCs) [11], [13], [27]. DIDs are globally unique identifiers that enable an entity to prove control over its identity through cryptographic mechanisms without depending on a centralized registry or certificate authority [13], [27]. A DID is associated with a DID document containing public keys and service endpoints required for trusted digital interaction [28]. In contrast to traditional

identifiers, DIDs are designed to support user control, privacy preservation, and interoperability across different technical environments [11].

Verifiable Credentials represent the digital equivalent of trusted documents such as diplomas, certificates, transcripts, or records of competency acquisition [8], [11], [21]. Their key advantage lies not merely in digitization, but in cryptographic verifiability. A VC allows claims to be issued in a machine-readable and tamper-evident format, making it possible for a verifier to validate authenticity and integrity without relying on manual institutional confirmation [19], [26].

In the context of higher education, the DID–VC combination forms the technical backbone of a decentralized trust infrastructure [11]. It supports credential portability, and cross-institutional interoperability, while also reducing the administrative burden associated with conventional certificate verification [1]. Recent research in decentralized education ecosystems emphasizes that this combination can strengthen earner ownership of credentials, improve trust, and support long-term verifiability of academic achievements [2], [6].

2.3 Zero-Knowledge Proofs and Selective Disclosure for Privacy-Preserving Verification

While DIDs and VCs enable decentralized and machine-verifiable trust, they do not by themselves fully resolve the tension between verifiability and privacy. This is where Zero-Knowledge Proofs (ZKPs) become particularly relevant. ZKPs are cryptographic techniques that allow one party to prove the truth of a statement without revealing the underlying data used to prove it [29]. Applied to educational contexts, they enable an earner to demonstrate possession of a valid academic credential without disclosing the full document or unrelated personal data [16], [20], [29].

This capability directly supports selective disclosure, a principle according to which only the minimum necessary information is shared in a verification process. For example, an employer may need confirmation that a learner has completed a specific program, but not access to the learner's personal identifiable information (PII), such as date of birth, national identity, or any additional records [6]. Similarly, another educational provider may require proof of eligibility for enrollment without access to the earner's complete academic history. In such cases, ZKPs enhance both privacy and trust by ensuring that verification can occur without excessive disclosure, transforming privacy from reactive measure to proactive design principle [4].

From a secure-by-design perspective, ZKPs add an advanced privacy-preserving layer to credential ecosystems [10], [16]. Contemporary academic verification frameworks increasingly integrate ZKPs, zk-SNARKs, or zkEVM-based architectures to support secure validation, improve regulatory compliance, and reduce the exposure of sensitive student information [3], [16], [29].

3 Related Work

Current academic data management systems still face serious structural challenges, where academic records are mostly managed through centralized databases under the control of institutions, making them prone to inefficiencies and data leakage [14]. As previously stated, shortcomings related to data silos, low interoperability, and reliance on manual verification directly contribute to the global problem of diploma forgery and the flourishing of the so-called "degree mill" industry [6]. This issue generates billions of dollars annually through the issuance of false certificates and transcripts, undermining trust in educational institutions and the labor market [29], [30]. Manual verification of diploma authenticity today takes about 12 days on average, which is unacceptable in the modern digital society and creates an administrative bottleneck for student mobility and employment [5].

In response to these challenges, several notable blockchain solutions have been developed. Blockcerts, an initiative of the MIT Media Lab, represents one of the oldest and most widespread standards for issuing digital certificates anchored on the Bitcoin or Ethereum blockchain [6], [29]. Although Blockcerts has proven technical feasibility, it suffers from high transaction costs and a lack of native privacy protection mechanisms (like ZKP), which can lead to the exposure of sensitive meta-data during the verification process [17], [29]. On the other hand, the EduCTX platform pioneered the concept of digital credits aligned with the ECTS system, using ARK and Ethereum infrastructure for the global transfer and accumulation of credits [6], [31]. However, early versions of this system did not fully utilize the advantages of smart contracts to automate complex business logic, nor did they implement standardized decentralized identifiers (DIDs) that would allow students full autonomy. The Bloxberg specialized academic network, based on Proof of Authority (PoA) consensus, offers an infrastructure tailored to research and educational institutions, ensuring the longevity and verifiability of credentials even if the issuing institution ceases to exist [19]. However, the lack of detailed protocols for user experience and mass adoption, as well as challenges in integration with the university's legacy IT systems, can also be observed at Bloxberg.

Of particular importance in the European context is the European Blockchain Services Infrastructure (EBSI), a joint initiative of the European Commission and 27 EU member states, which lays the foundations for the cross-border exchange of academic data through the Web3 trust model [8], [21]. EBSI integrates the European Self-Sovereign Identity Framework (ESSIF), enabling the development of digital wallets that give students full control over their own identity without relying on central registries. Pilot projects within EBSI, such as collaborations between universities in Belgium, Italy and Romania, have shown that a model based on the SSI and Verifiable Credentials standards is ready for production at the national level, achieving a high success rate of real-time verification [4], [26]. EBSI transforms the trust paradigm from the Web 2.0 model of "trust the platform" to the Web 3.0 imperative of "don't trust, verify," where the credibility of a degree is cryptographically verified, without the need to contact the institution directly.

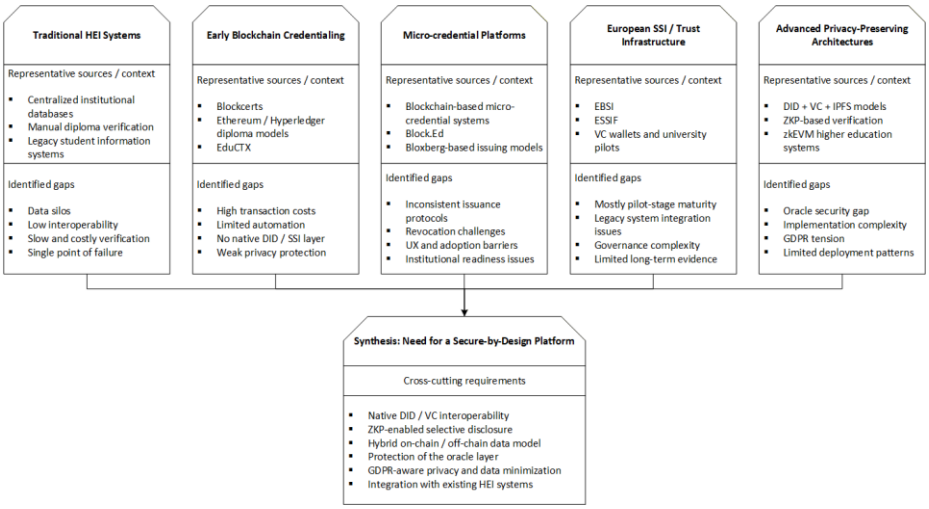


Fig. 1. Comparison of current credentialing systems and platforms

Despite progress, analysis of current solutions reveals critical gaps (Fig. 1) that prevent the full realization of a secure-by-design ecosystem. The biggest risk is the so-called "Oracle security gap" - a point of vulnerability at the border between legacy systems (like Moodle) and the blockchain, where incorrect data can be permanently and irreversibly written to the ledger. Also, most existing platforms still face the conflict between blockchain immutability and the GDPR right to be forgotten, as they often inadequately separate personal information (PII) from cryptographic evidence. The lack of native integration of Zero-Knowledge Proof (ZKP) technology in most current solutions forces students to share entire transcripts even when only partial verification is required, thus violating the principle of data minimization. These identified shortcomings - from the insecurity of the Oracle layer to incomplete privacy protection - form the basis for the proposal of a new, more robust architecture that will be defined through the principles of secure-by-design design in further work.

4 Secure-by-Design Micro-credential Platform

Based on the current technological gaps identified in higher education ecosystems and the established potential of the European Blockchain Services Infrastructure (EBSI), this paper proposes a Secure-by-Design Micro-credential Platform. The core philosophy of this proposal is not to build a new, standalone blockchain from scratch, but rather to design a foundational upgrade and extension to the existing EBSI framework, leveraging its multi-layer, modular architecture to enhance privacy and security. The platform is conceived as a layered digital trust infrastructure that supports the full lifecycle of a micro-credential identified in previous research [1]: evidence capture, validation, issuance, storage, presentation, verification, update, and revocation. Rather than treating security and privacy as additional modules, the platform embeds them into the

architecture from the beginning. In this way, the platform connects institutional learning systems with a broader digital trust ecosystem, while preserving the roles of universities, earners, employers, and governance actors within the credential lifecycle.

Figure 2 presents the proposed platform as a five-layer architecture. The figure should be interpreted not only as a technical system design but also as a model of stakeholder participation, showing how different actors contribute to the use, maintenance, and legitimacy of the platform.

At the foundation of the architecture is the Secure Oracle Layer, which contains the systems in which evidence of learning is generated and maintained. These include learning management systems such as Moodle, Canvas or Teams, student information systems, assessment platforms, and other repositories of competency-related data. These systems provide the educational evidence from which proofs are derived, such as course completion, assessment outcomes, participation in extracurricular activities, or validated non-formal learning experiences. However, because institutional systems may differ in quality, completeness, and internal governance, this layer is not treated as a simple data input environment. Instead, it operates as a Secure Oracle, whose purpose is to validate source data before they enter the credentialing workflow. This secure oracle performs multi-source validation, for example by cross-checking academic records, assessment outputs, and digitally signed institutional confirmations. In this way, it addresses the oracle security gap by ensuring that only verified and internally consistent data are transformed into credential proofs.

The second layer is the Identity and Registry Layer, which directly integrates with established trust services for decentralized identity and credential governance. At this level, the platform uses decentralized identifiers to represent the key actors of the ecosystem, including students as credential holders and institutions as trusted issuers. The role of this layer is to ensure that each actor can be represented through a globally unique and verifiable identifier, while trusted registries provide the basis for recognizing legitimate issuers and approved credential structures. This layer supports a self-sovereign identity logic in which earners retain control over the credentials they hold, while institutions preserve their authority as recognized issuers. In practical terms, it allows micro-credentials to be embedded into a broader trust framework rather than being limited to a single university or vendor-controlled environment.

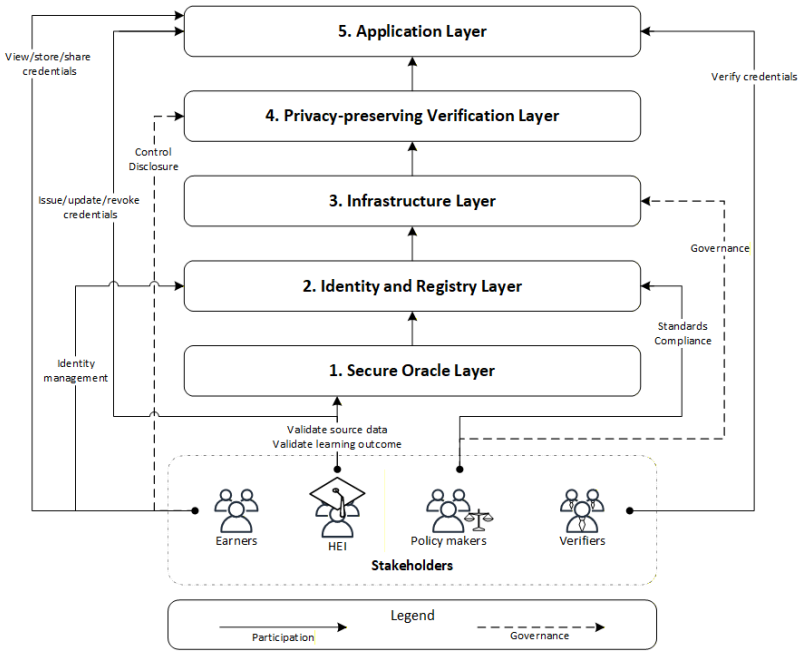


Fig. 2. EBSI-enabled five-layer architecture and stakeholder participation

The third layer is the Infrastructure Layer, which relies on the EBSI network as the trust backbone of the platform. This layer provides the distributed node environment in which credential-related references, issuer attestations, status information, and integrity proofs can be anchored. Its function is to support authenticity, traceability, and resilience through a trusted public-sector infrastructure. In this sense, the infrastructure layer acts as a governmental trust anchor, offering a level of legal and institutional reliability that isolated private platforms cannot easily replicate. At the same time, it supports interoperability across institutional and national contexts, which is particularly important in higher education systems increasingly shaped by lifelong learning, micro-credentials, and learner mobility.

Above this foundation, the architecture introduces the Privacy-Preserving Verification Layer, which represents the main functional enhancement proposed in this paper. While standard digital credential flows often rely on digital signatures that still require the presentation of entire documents or excessive metadata, higher education verification scenarios frequently demand a more selective and privacy-aware approach. Employers, partner institutions, or accreditation bodies often need to verify only one specific claim, such as whether a learner completed a certain course, possesses a valid micro-credential, or satisfies a threshold requirement, without accessing the learner's complete transcript or unrelated personal data. To support this need, the proposed architecture incorporates Zero-Knowledge Proof mechanisms that enable selective disclosure. Through this layer, a learner can prove the validity of a claim without revealing the full credential content. This makes the platform substantially more aligned with privacy-by-design principles, since verification is transformed from full document

disclosure into claim-based proof. Conceptually, this layer acts as a bridge between trusted credential infrastructure and real-world privacy requirements in higher education.

At the top of the model is the Application Layer, which contains the user-facing services through which the platform is used in practice. For students, this layer is primarily represented by digital wallets through which they receive, store, manage, and present their micro-credentials. For universities, the application layer includes issuer interfaces and administrative tools that support credential creation, issuance, update, and revocation. For employers and other verifiers, it includes verification portals or APIs through which submitted presentations can be checked efficiently and reliably.

Taken together, the five layers of the proposed platform form a coherent secure-by-design architecture for micro-credential management in higher education. The model secures the input side of the system through a validated oracle-based connection with institutional learning systems, and secures the output side through privacy-preserving verification. Between these two points, it relies on decentralized identity, trusted registries, and interoperable infrastructure to ensure that micro-credentials are not only technically verifiable, but also institutionally meaningful and transferable across contexts. In that sense, the proposed architecture should be understood not merely as a technical solution for issuing digital credentials, but as a structured trust model for the next generation of higher education credential ecosystems.

4.1 Implementation and Evaluation Roadmap

In the previous research [1], [9] we identified key stakeholders, processes and aspects in relation to micro-credentials as a concept:

- Stakeholders: earners, course providers, issuers, verifiers and policy makers
- Processes: earning, issuing, maintaining and sharing
- Needs and values for earners and verifiers: trusted infrastructure, data security, transparency and reliability, efficiency

that will impact the development of micro-credential platform.

To move from conceptual architecture toward practical adoption, the proposed five-layer platform should be implemented through a phased roadmap that begins with a minimum viable product (MVP), continues through controlled institutional evaluation, and gradually expands toward broader organizational and inter-institutional integration.

Minimum viable product (MVP) development – The development is currently in this stage, where the priority is to identify and model the core entities and processes required for micro-credential management.

1. The entities include the micro-credential itself, the earner, the issuing institution, the issuer role, the verifier, the credential schema, learning evidence, proof objects, credential status records, and trusted identity references. Since operational course-provider processes may vary significantly across

institutions, the MVP will not attempt to standardize all internal educational workflows. Instead, it focuses on the two essential actors that remain stable across contexts: the earner and the issuer.
2. The processes include credential issuance, storage of credential-related references, learner-controlled presentation, verification by third parties, and credential revocation or status update.
3. Smart contract modeling and blockchain integration on a low-cost fallback environment, responsible for credential issuance, verification, status control, and revocation logic. The fallback platform that will be used during piloting is Bloxberg, because a node already exists within the participating institution.
4. The technical MVP also includes UI and UX design, since user acceptance is a critical condition for adoption.

Department-level evaluation – This stage will be conducted within the authors' institution and department, in order to test the platform in a controlled setting.

1. The evaluation will involve three user groups: students as earners, academic staff as issuers, and a selected pool of employers as verifiers.
2. The USE questionnaire (Usefulness, Satisfaction, and Ease of Use) will be employed as the main evaluation instrument for assessing user perception of the platform.
3. Evaluation should provide both quantitative and qualitative feedback regarding the usability of the wallet interface, the clarity of the issuance and verification flows, and the perceived trustworthiness of the system.

The third phase will begin after the first round of evaluation and will focus on platform refinement and institution-wide pilot testing.

1. Necessary modifications will be introduced into both the technical architecture and the user interfaces, based on the feedback.
2. The system should be aligned with the final EBSI registration and integration process, once the relevant governance conditions are stabilized.
3. The institution-wide pilot will also include integration with the existing institutional identity system and with data oracle systems, particularly Moodle and Teams, as operational sources of validated learning evidence.
4. Training activities will be organized for academic staff and students to support adoption and reduce resistance to the new workflow.
5. The second round of evaluation should be conducted with a significantly larger participant group in order to validate the platform under more realistic organizational conditions.

The final phase concerns institutionalization and ecosystem expansion. Once the platform has been validated at the institutional level, the next step is to extend its use toward partner institutions within the university's broader academic network. This

phase is important because the value of micro-credentials depends not only on secure issuance, but also on portability and recognition across organizational boundaries. Beyond technical interoperability, this stage should include work on procedural and regulatory alignment, including proposals for revisions of institutional procedures, internal rulebooks, and governance practices. In the longer term, national bodies should be gradually included in the process, especially where accreditation, quality assurance, and formal recognition frameworks are relevant.

The purpose of this roadmap is not only to test technical feasibility, but also to validate whether the proposed secure-by-design model can function in a real higher education environment shaped by institutional workflows, user expectations, and regulatory constraints.

5 Limitations

The main limitation of this paper is that the proposed model has not yet been empirically evaluated in a real higher education setting. As a result, its usability, institutional fit, and practical effectiveness still need to be validated through pilot implementation and user-centered evaluation. In addition, the practical deployment of such a complex architecture may be challenging for higher education institutions, especially where legacy systems, fragmented identity infrastructures, and limited technical capacity are present.

Further limitations concern institutional readiness for implementation. These include technical conditions for connecting to blockchain infrastructure, onboarding students into a credential wallet ecosystem, defining responsibility for operational issuance costs, and ensuring institutional support for node maintenance and system integration. Finally, the proposed alignment with EBSI may also face delays or constraints related to governance processes and interoperability requirements. Therefore, the proposed platform should be understood as a conceptually mature but still implementation-dependent model whose full validation requires phased deployment and iterative testing.

6 Discussion

The analysis confirms that the current higher education credentialing landscape is characterized by institutional data silos, low interoperability, and slow manual verification procedures, which on average take up to 12 days [2], [5]. While blockchain technology, Decentralized Identifiers (DIDs), and Verifiable Credentials (VCs) offer significant potential for improving authenticity and trust [2], [3], [14], [20], existing solutions often remain technically partial or insufficiently aligned with complex privacy and governance requirements [19], [20]. A key insight is that no single mechanism is sufficient; blockchain improves immutability, but source data reliability remains a critical vulnerability, requiring micro-credential ecosystems to be designed as integrated socio-technical systems [4], [6].

The proposed secure-by-design platform addresses these limitations through a layered architectural response. The LMS / Data Source Layer incorporates a Secure Oracle that responds to the oracle-security gap by introducing multi-source validation before any credential claims are hashed onto the ledger. The Identity and Registry Layer utilizes W3C standardized DIDs to establish a Self-Sovereign Identity (SSI) model, ensuring that learners possess globally unique identifiers and maintain primary control over their academic records without relying on central authorities. Leveraging the EBSI-based Infrastructure Layer [8] provides a governmental trust anchor, supporting cross-border interoperability and digital sovereignty within the European education area.

The most significant enhancement is the Privacy-Preserving Verification Layer. Unlike existing systems that often require full document sharing, the proposed model integrates Zero-Knowledge Proofs (ZKPs) to enable selective disclosure. This allows a learner to prove possession of a valid micro-credential or the fulfillment of a requirement (e.g., "GPA > 8.0") without exposing a full transcript or unrelated personal data, thereby satisfying GDPR data minimization principles. At a systemic level, this platform provides a blueprint for a trustworthy digital ecosystem where secure data input, trusted identity, and privacy-preserving verification are architecturally intertwined to support the future of lifelong learning.

7 Conclusion

This paper examined how micro-credential systems in higher education can be designed to better support trust, privacy, and interoperability. The review of the literature and existing platforms showed that, while current solutions have improved digital issuance and verification, they still face important limitations. These are most visible in privacy protection, governance, integration with institutional systems, and cross-border interoperability.

In response, the paper proposed a five-layer secure-by-design architecture aligned with the EBSI ecosystem. The model combines a Secure Oracle for validated input data, decentralized identity and trusted registries, an EBSI-based infrastructure layer, a Privacy-Preserving Verification Layer based on selective disclosure and zero-knowledge proofs, and an application layer that supports learners, issuers, and verifiers. The architectural proposal is further extended through an implementation and evaluation roadmap, positioning the model not only as a conceptual framework, but also as a practical direction for future development.

The study also has clear limitations. The proposed model has not yet been empirically tested in a real institutional setting, and its implementation depends on technical readiness, governance capacity, and the maturity of the surrounding trust infrastructure.

The main contribution of the paper lies in framing micro-credential management as more than a technical issue of issuing digital certificates. It shows that trust, privacy, interoperability, and governance must be addressed together. Although the model still requires empirical validation, it provides a clear conceptual and practical foundation

for the future development of learner-centered and privacy-aware credential ecosystems in higher education.

Acknowledgments. AI-use disclosure statement. During the preparation of this manuscript, AI-based tools were used solely for language editing and improving readability. All scientific content, methodology, experiments, and conclusions were developed and verified by the authors. The authors take full responsibility for the integrity and accuracy of the work.

References

1. T. Naumović, M. Simić, M. Suvajdžić, M. Vulić, and A. Bjelica, "Roles, Responsibilities, and Value Flows: A Stakeholder Map of Micro-Credential Processes in Higher Education Ecosystem," *E-bus. Technol. Conf. Proc.*, vol. 4, no. 1, 2025, [Online]. Available: https://www.ebt.rs/journals/index.php/conf-proc/issue/archive
2. R. Jopate *et al.*, "Blockchain-Based Secure Identity and Credential Management for Decentralized Education Ecosystems," *Int. J. Oo Recent Trends Technol. Eng.*, vol. 4, no. 4, 2025, doi: 10.5281/zenodo.17862602.
3. H. A. Alsobhi, R. A. Alakhtar, A. Ubaid, O. K. Hussain, and F. K. Hussain, "Blockchain-based micro-credentialing system in higher education institutions: Systematic literature review," *Knowledge-Based Syst.*, vol. 265, p. 110238, Apr. 2023, doi: 10.1016/J.KNOSYS.2022.110238.
4. L. Wei, M. Al-harthi, and R. Chen, "Smart Contract-Enabled Micro-Credentialing for Competency-Based E-Learning : Architecture , Empirical Analysis , and Design Guidelines," *Trends Rev. Technol. Eng. Educ.*, vol. 3, no. 2, pp. 1–27, 2025, doi: https://doi.org/10.63646/trtee.2025.030201.
5. D. L. Fekete and A. Kiss, "Toward Building Smart Contract-Based Higher Education Systems Using Zero-Knowledge Ethereum Virtual Machine," *Electron.*, vol. 12, no. 3, pp. 1–39, 2023, doi: 10.3390/electronics12030664.
6. Z. Z. Li, K. Liu Joseph, J. Yu, and D. Gasevic, "Blockchain-based Solutions for Education Credentialing System: Comparison and Implications for Future Development," *Proc. - 2022 IEEE Int. Conf. Blockchain, Blockchain 2022*, pp. 79–86, 2022, doi: 10.1109/Blockchain55522.2022.00021.
7. J. A. Berrios Moya, J. Ayoade, and M. A. Uddin, "A Zero-Knowledge Proof-Enabled Blockchain-Based Academic Record Verification System," *Sensors*, vol. 25, no. 11, pp. 1–24, 2025, doi: 10.3390/s25113450.
8. "EBSI Verifiable Credentials - EBSI -." https://ec.europa.eu/digital-building-blocks/sites/spaces/EBSI/pages/600343491/EBSI+Verifiable+Credentials (accessed Apr. 27, 2026).
9. T. Naumović and D. Kostić, "Upravljanje Mikrokredencijalima u Visokoškolskom Obrazovanju Zasnovano na Blokčejn Tehnologijama," Seminar on Computer Science and Applied Mathematics, Beograd, Srbija, 2025. [Online]. Available: https://www.mi.sanu.ac.rs/novi_sajt/seminars/programs/seminar2.may2025.php
10. J. Komljenovic, "The future of value in digitalised higher education: why data privacy should not be our biggest concern," *High. Educ.*, vol. 83, no. 1, pp. 119–135, 2022, doi: 10.1007/s10734-020-00639-7.

11. C. Mazzocca, A. Acar, S. Uluagac, R. Montanari, P. Bellavista, and M. Conti, “A Survey on Decentralized Identifiers and Verifiable Credentials,” *IEEE Commun. Surv. Tutorials*, vol. 27, no. 6, pp. 3641–3671, 2025, doi: 10.1109/COMST.2025.3543197.
12. M. Polychronaki, M. G. Xevgenis, D. G. Kogias, and H. C. Leligou, “Decentralized Identity Management for Metaverse-Enhanced Education: A Literature Review,” *Electron.*, vol. 13, no. 19, 2024, doi: 10.3390/electronics13193887.
13. C. N. Butincu and A. Alexandrescu, “Design Aspects of Decentralized Identifiers and Self-Sovereign Identity Systems,” *IEEE Access*, vol. 12, pp. 60928–60942, 2024, doi: 10.1109/ACCESS.2024.3394537.
14. A. A. Ayare, V. A. Jadhav, M. K. Banatwala, S. V Changlere, A. Mote, and P. Joshi, “A Systematic Review on Blockchain-Based Framework for Storing Educational Records Using InterPlanetary File System,” *Cureus J. Comput. Sci.*, pp. 1–28, 2025, doi: 10.7759/s44389-024-00745-y.
15. T. Berichte Nr, C. Meinel, M. Galbas, and D. Hagebölling, *Digital sovereignty: insights from Germany’s education sector*, 157th ed. Potsdam: Potsdam University Press, 2023. doi: 10.25932/publishup-59772.
16. X. Li, M. Tan, and W. Tian, “A Secure and Efficient Sharing Framework for Student Electronic Academic Records: Integrating Zero-Knowledge Proof and Proxy Re-Encryption,” *Futur. Internet*, vol. 18, no. 1, p. 47, 2026, doi: 10.3390/fi18010047.
17. D. Schumm, K. O. E. Muller, and B. Stiller, “Are We There Yet? A Study of Decentralized Identity Applications,” *IEEE Access*, vol. 13, pp. 125232–125259, 2025, doi: 10.1109/ACCESS.2025.3588170.
18. “Art. 25 GDPR – Data protection by design and by default - General Data Protection Regulation (GDPR).” https://gdpr-info.eu/art-25-gdpr/ (accessed Apr. 27, 2026).
19. M. Základsník, V. Konečná, O. Faldík, O. Trenz, and A. Gono, “Issuing Micro-Credentials With Bloxberg Blockchain,” pp. 172–181, 2025, doi: 10.11118/978-80-7701-047-4-0172.
20. A. Mohammad and S. Vargas, “Barriers Affecting Higher Education Institutions’ Adoption of Blockchain Technology: A Qualitative Study,” *Informatics*, vol. 9, no. 3, 2022, doi: 10.3390/informatics9030064.
21. E. Tan, E. Lerouge, J. Du Caju, and D. Du Seuil, “Verification of Education Credentials on European Blockchain Services Infrastructure (EBSI): Action Research in a Cross-Border Use Case between Belgium and Italy,” *Big Data Cogn. Comput.*, vol. 7, no. 2, 2023, doi: 10.3390/bdcc7020079.
22. C. Allen, “The Path to Self-Sovereign Identity - Life With Alac...” https://www.lifewithalacrity.com/article/the-path-to-self-soverereign-identity/ (accessed Apr. 27, 2026).
23. S. El Haddouti and M. D. Ech-Cherif El Kettani, “Analysis of identity management systems using blockchain technology,” *Proc. - 2019 Int. Conf. Adv. Commun. Technol. Networking, CommNet 2019*, Apr. 2019, doi: 10.1109/COMMNET.2019.8742375.
24. M. Martin and P. Van der Hijden, “Short courses, micro-credentials, and flexible learning pathways: a blueprint for policy development and action: policy paper,” UNESCO IIPE, 2023. Accessed: Apr. 20, 2026. [Online]. Available: https://unesdoc.unesco.org/ark:/48223/pf0000384326
25. A. Jušić, S. Fuks, P. Kochovski, and V. Stankovski, “Blockchain-Enabled Micro Credentialing: Enhancing Trust, Mobility, and Interoperability in Lifelong Learning,” *EDULEARN25 Proc.*, vol. 1, pp. 5382–5388, 2025, doi: 10.21125/edulearn.2025.1342.
26. V. Holotescu, R. Vasiu, A. Ternauciuc, C. Holotescu, and D. Andone, “Building a Decentralized Education Ecosystem : Politehnica University of Building a Decentralized Education Ecosystem : Politehnica University of Timisoara ’ s Pioneering Blockchain

Initiatives," in *EDEN 2024 Research Workshop & PhD Masterclass*, Timisoara, Romania, 2024. [Online]. Available: https://www.researchgate.net/publication/385416650_Building_a_Decentralized_Education_Ecosystem_Politehnica_University_of_Timisoara's_Pioneering_Blockchain_Initiatives

27. Apoorva Jain, "Blockchain-Based Decentralized Identity Systems in Education," *Sci. J. Artif. Intell. Blockchain Technol.*, vol. 2, no. 4, pp. 30–38, 2025, doi: 10.63345/sjaibt.v2.i4.104.
28. M. Chen, B. Chen, Z. Ling, P. Zhang, and S. Zhang, *A Study of Managing University Course based on Decentralized Identifiers*, vol. 1, no. 1. Association for Computing Machinery, 2023. doi: 10.1145/3651655.3651657.
29. X. Xu, "Zero-knowledge proofs in education: a pathway to disability inclusion and equitable learning opportunities," *Smart Learn. Environ.*, vol. 11, no. 1, 2024, doi: 10.1186/s40561-024-00294-w.
30. A. Rustemi, F. Dalipi, V. Atanasovski, and A. Risteski, "A Systematic Literature Review on Blockchain-Based Systems for Academic Certificate Verification," *IEEE Access*, vol. 11, pp. 64679–64696, 2023, doi: 10.1109/ACCESS.2023.3289598.
31. P. S. Rani, S. B. Priya, and A. Kala, "Digital Certificate Verification in Educational Sector Using Non-Interactive Zero Knowledge Proof and Blockchain," *J. Inf. Sci. Eng.*, vol. 41, no. 2, pp. 467–480, 2025, doi: 10.6688/JISE.202503_41(2).0012.

Original Research Article *https://doi.org/10.5281/zenodo.23476259*

Artificial Intelligence for Digital Transformation in Sport: An Integrated Conceptual Model for Elite Sport, Recreational Sport, and Public Health

Aleksandar Đurić[0009-0002-5102-5243]

University of Belgrade, Faculty of Organizational Sciences
e-mail: djalexandar@gmail.com

The application of artificial intelligence (AI) in sport has been driven primarily by the needs of elite and professional athletes, while its use in recreational sport remains limited due to technical, organizational, and regulatory constraints. This paper examines opportunities for the digital transformation of sport in the Republic of Serbia through a combination of normative analysis and conceptual modeling.

The main contribution of the study is the proposed 360° Integrated AI Model for Sport, which seeks to connect elite sport, recreational sport, and public health systems within a common digital framework. The model is designed to support the integration of sport-related data and services, improve decision-making processes, and enable more personalized and evidence-based approaches. It also creates opportunities for research, innovation, and collaboration among sport organizations, public institutions, and the academic community.

The analysis identifies a number of technical, strategic, regulatory, and ethical challenges related to the broader adoption of AI in sport. The findings suggest that better alignment of strategic frameworks, greater interoperability between existing information systems, and the use of shared digital infrastructure could support a more integrated and sustainable approach to the digital transformation of sport.

Keywords: artificial intelligence, sport, recreational sport, digital transformation, public health, integrated model

1 Introduction

The application of artificial intelligence (AI) has significantly influenced modern sport, while also contributing to a growing technological imbalance between professional and recreational sport environments. Advanced algorithms, sensor systems, and digital infrastructures are increasingly used in elite sport for performance optimization and injury prevention, whereas recreational sport and public health systems remain only partially integrated into these developments. This gap is associated with fragmented IT infrastructures, insufficient strategic alignment, and the lack of a broader interoperable

framework capable of connecting physical activity data with public health systems and sport-specific analytics.

To overcome these challenges, a more integrated approach to design of sport and health information systems is needed. The main contribution of this paper is the proposal of a conceptual "360° integrated AI model for sport," designed to support the integration of existing technical, scientific, and institutional resources within the Republic of Serbia. Through a centralized digital infrastructure, the proposed model enables bidirectional data exchange between sport and healthcare systems, supporting personalized physical activity and broader public health initiatives.

This paper is primarily conceptual and strategic rather than empirical in nature. Its objective is to establish theoretical and architectural foundations for the future development of integrated AI-supported sport ecosystems. Accordingly, the paper follows a top-down approach by first examining key infrastructural and regulatory challenges, followed by the presentation of the main components and integration mechanisms of the proposed 360° AI model.

2 Research Hypotheses

The study is structured around one main hypothesis and four supporting hypotheses. Together, they examine the strategic, technical, and organizational conditions required for the broader adoption of artificial intelligence in sport and explore the potential of the proposed 360° Integrated AI Model for Sport to support the digital transformation of the sport ecosystem.

Main Hypothesis (H_0): The architectural framework of the proposed "360° integrated AI model for sport" conceptually supports the integration of elite sport, recreational sport, and public health systems by addressing key technological and systemic challenges, including structural fragmentation and limited interoperability within the national sport ecosystem.

Supporting Hypotheses 1 (H_1): IT Infrastructure & Interoperability
The proposed model suggests that interoperability between fragmented recreational platforms and institutional health systems could be supported through a centralized AI-enabled data exchange layer.

Supporting Hypotheses 2 (H_2): Resource Integration
The scalability and practical applicability of the proposed model depend on the integration of existing national scientific, institutional, and cloud computing resources within a shared digital environment.

Supporting Hypotheses 3 (H_3): Regulatory Frameworks
Legal and ethical issues related to data privacy, protection, and ownership represent important constraints for the practical implementation of integrated AI systems in sport.

Supporting Hypothesis 4 (H_4): Strategic & Policy Alignment
The long-term sustainability and broader adoption of the proposed model require greater alignment between public policies, digital strategies, and sport governance frameworks.

3 Research methodology

This study adopts a conceptual-analytical approach and combines three complementary methods: a structured literature review, normative analysis of national strategic documents, and conceptual modeling. Together, these methods provide the basis for proposing a systemic framework that addresses current challenges related to sport development and the application of artificial intelligence.

3.1 Literature Collection and Evaluation Protocol

The literature review was conducted through a structured search of major scientific databases, primarily Scopus and Web of Science. The search focused on studies addressing the application of artificial intelligence in sport, including performance analysis, talent identification, injury prevention, personalized recovery, and sport management.

Keywords such as “artificial intelligence in sport” and related terms were used to identify relevant publications. The analysis was limited to studies published between 2024 and 2026 to reflect recent research trends. Attention was given to review articles because they offer a broader perspective on recent developments in the field.

The selection and screening of studies followed the general principles of the PRISMA framework, adapted to the scope of this research.

During the review, it was observed that many studies rely on relatively small and homogeneous samples, often focused on professional athletes. This limitation may affect the broader applicability of findings and increase the risk of bias.

3.2 Normative Analysis and Conceptual Modeling

In the second phase of the research, a normative analysis was conducted to examine the strategic and regulatory framework of the Republic of Serbia. By combining insights from relevant academic literature with an analysis of key national policy documents, an institutional gap in the application of artificial intelligence in sport was identified.

Based on these findings, a “360° integrated AI model for sport” was developed using synthesis and conceptual modeling methods. The model incorporates key findings into a practical framework, based on bidirectional communication, interoperability, and responsible data governance.

4 Literature Review: Technological Framework and Dominant Applications of Artificial Intelligence in Sport

Most recent studies on artificial intelligence in sport focus on the use of AI-enabled systems in elite and professional sport environments. Current applications are primarily based on Machine Learning (ML), Deep Learning (DL), Computer Vision (CV), and Natural Language Processing (NLP) techniques and have been widely implemented across various sports disciplines [2, 4, 12, 16, 23].

These technologies are mainly applied in biomechanical and tactical analysis, where sensor data and video footage are utilized for training optimization, load management, and injury prevention [1, 7, 9, 17–19, 21, 23]. Furthermore, AI-driven systems are increasingly employed for predictive analytics, real-time performance tracking, and automated refereeing, providing direct support to coaches and sports analysts at the professional level [3, 5, 11, 21]. Beyond athlete-centered applications, AI is used in sport event management to support scheduling, resource allocation, crowd management, and operational decision-making [4].

A review of the available literature indicates that advanced AI systems remain largely concentrated in elite sport, with relatively limited attention given to recreational sport and its integration with public health initiatives. While professional sport increasingly relies on personalized analytics, wearable technologies, and dedicated digital infrastructures [16, 18–21], recreational athletes primarily use consumer-oriented wearable devices and fitness applications, with limited integration into broader sport and health ecosystems [5, 12, 15]. At the same time, existing studies mostly focus on isolated performance-oriented applications, while integrated frameworks connecting elite sport, recreational physical activity, and public health systems remain largely unexplored [12, 15, 20, 22]. This gap represents the primary motivation for the development of the proposed “360° integrated AI model for sport”.

5 Technical Capacities and Data Availability in the Republic of Serbia

The implementation of advanced AI models within a national sports system depends on the availability of adequate digital infrastructure and high-quality data. Contemporary AI systems rely on processing large volumes of heterogeneous data, ranging from structured databases to unstructured sources such as video recordings, sensor data, and textual information.

In this context, existing national digital infrastructure provides a foundation for the broader adoption of AI in sport.

5.1 Infrastructure for Large-Scale Data Processing: The State Data Center

The State Data Center represents an important component of the digital infrastructure that could support the digital transformation of sport in Serbia. According to official guidelines from the Office for Information Technologies and eGovernment, this center serves as the backbone of the National Platform for Artificial Intelligence, providing the computational power required for training complex deep learning models on large-scale datasets [16].

From both legal and security perspectives, reliance on national cloud infrastructure is of particular importance for the sports system. It reduces dependence on foreign providers and allows sensitive biometric, medical, and training data to be managed within the national regulatory framework, while ensuring compliance with applicable data protection regulations.

5.2 Research and Development Potential and Technology Validation

In addition to developed hardware infrastructure, the Republic of Serbia also has relevant institutional and research capacities in the fields of information technology and sport sciences. Expertise in AI is supported through study programs at leading national faculties.

In this context, an important role is played by the Artificial Intelligence Research and Development Institute of Serbia (established in 2021), which serves as one of the central institutions for the development and application of AI technologies. However, despite its role in connecting academia, industry, and the public sector, the application of AI in sport has not yet been systematically developed.

The growing number of publications by domestic researchers reflects increasing interest in the application of artificial intelligence in sport. Recent studies have addressed diverse areas, including markerless motion capture and computer vision techniques for non-invasive movement analysis [2], machine learning models for heart rate prediction based on data collected from commercial wearable devices [15], and the use of predictive analytics and heart rate variability monitoring for fatigue detection and injury prevention [8]. These studies illustrate the availability of relevant scientific expertise and highlight opportunities for the broader application of AI within the national sport ecosystem.

Despite the availability of relevant expertise and technological resources, these capacities continue to operate largely within isolated academic and technological silos, without clear strategic integration into a unified national sport system. This fragmentation limits the broader use of existing capacities and highlights the importance of stronger coordination between research, technology, and sport institutions.

5.3 National and Regional Records as AI Data Sources

The effective use of supercomputing infrastructure and domestic AI models depends on access to high-quality and well-structured data. According to the Law on Sport

(Article 168) and the accompanying Rulebook on National Records in Sport, key institutions, such as the Republic and Provincial Institutes for Sport and Sports Medicine, continuously collect and maintain various types of data.

These datasets include information on athletes (age, sport discipline, competition level), as well as databases on sports injuries and their treatment. From a data science perspective, such longitudinal datasets represent a valuable resource for the developing machine learning models.

Alongside structured records, modern sport environments generate large volumes of unstructured and semi-structured data, including video recordings, wearable sensor data, and data from digital applications. However, these data are often fragmented and insufficiently integrated, which limits their broader applicability.

With appropriate anonymization techniques, which remove personal identifiers while preserving relevant medical, biomechanical, and training parameters, a portion of these datasets could be used for the development of domestic predictive AI models.

Such an approach could support the development of injury prevention models tailored to the characteristics of the local population, thereby improving their applicability within the national sport ecosystem.

6 Strategic Framework: Global Challenges and Local Fragmentation

The effective use of technological resources depends on a clearly defined and well-aligned strategic framework. However, many digital sport systems face a form of institutional lag, where public policies do not evolve at the same pace as technological development.

This misalignment can limit the effective implementation of artificial intelligence, particularly in complex systems such as sport, where multiple stakeholders and data sources must be coordinated.

6.1 Comparative Analysis of Institutional Challenges: China and the United Kingdom

Institutional challenges related to the integration of artificial intelligence in sport are evident even in technologically advanced countries. In China, despite major investments in AI and sports analytics, research points to limited coordination between national institutions, research centers, and local sport organizations, resulting in fragmented data ecosystems and reduced interoperability [4, 20].

In contrast, the United Kingdom has adopted a more coordinated approach through centralized data infrastructures that support the integration of data from multiple sport-related systems. Such approaches support more comprehensive monitoring of athlete performance, health, and long-term development.

This comparison indicates that the successful application of AI in sport depends not only on technological resources, but also on institutional coordination, interoperability, and strategic alignment.

6.2 Strategic Asymmetry and Educational Gap in the Republic of Serbia

Building on the findings of the comparative analysis, the situation in the Republic of Serbia reveals persistent strategic and institutional fragmentation. A clear misalignment exists between two key national policy documents. While the draft *Ex-ante Analysis of the Sports Development Strategy in the Republic of Serbia (2026–2035)* acknowledges the importance of digitalization, it does not address the broader application of advanced AI technologies [14]. Conversely, the *Strategy for the Development of Artificial Intelligence in the Republic of Serbia (2025–2030)* does not recognize sport as a domain of application, focusing instead on industry, public administration, and healthcare [6].

An additional challenge lies in the lack of strategic continuity within the sports sector. The last officially adopted Sports Development Strategy covered the period 2014–2018, followed by a prolonged policy gap without a new overarching framework. This discontinuity hinders the systematic introduction of emerging technologies, including artificial intelligence.

Insufficient coordination at the national level is also reflected in the academic and professional domains. The same pattern can be observed in the organization of professional education for sports practitioners, which is often conducted through ad hoc initiatives, lacking structured collaboration with relevant scientific and technological institutions.

An illustrative example of sector-specific digitalization without adequate coordination is the *Zdravitas* platform [12]. Developed within the education sector, the system was designed to collect data on students' anthropometric characteristics and motor abilities. Despite the relevance of such data for both sport and public health, its broader application has remained limited due to the lack of integration with other systems. In particular, the absence of connectivity with healthcare information systems has resulted in the creation of isolated datasets with restricted usability.

Furthermore, the governance of sensitive data presents an additional challenge. The lack of clearly defined data collection purposes, combined with insufficiently transparent communication with users, has led to implementation difficulties. This highlights the absence of a coherent data governance framework, which is essential for the responsible development and deployment of AI systems.

A similar pattern can be observed within the healthcare sector. The national e-health platform, *eZdravlje*, provides access to medical data through services such as *eKarton* (electronic health records) and "*Način života*" (lifestyle monitoring) modules [13]. However, despite capturing certain aspects of physical activity, the system remains primarily oriented toward health monitoring and disease prevention, rather than supporting sport participation or performance development.

As a result, systems operating across health, education, and sport remain largely disconnected. The absence of integration with sports institutions and training processes significantly limits the potential for meaningful data utilization. Consequently, data remains confined within isolated systems, reinforcing existing fragmentation across sectors.

This fragmentation hinders the development of interdisciplinary approaches and delays the establishment of an effective AI governance framework, ultimately limiting the ability of sports professionals to participate in the digital transformation of their field.

7 Regulatory Framework and Governance of Artificial Intelligence: Ethical Implications and Data Protection

The application of artificial intelligence in sport is not solely a technological challenge; it also raises legal and ethical concerns. The use of AI-driven analytics requires the continuous collection and processing of sensitive data, including biometric, health, and location information, which necessitates a clearly defined governance framework.

Without appropriate rules and oversight mechanisms, such systems may create risks, including unauthorized access, data misuse, and discriminatory outcomes. These challenges become particularly relevant in integrated digital environments where data are exchanged between sport, healthcare, and education systems.

7.1 Institutional Frameworks, Privacy Protection, and the Risk of Discrimination

The processing of personal data in sport entails risks that extend beyond simple data breaches. Recent studies highlight extreme, yet plausible scenarios in which predictive models could be misused by clubs or sponsors, for example, to exclude athletes from competition or deny contractual opportunities based solely on algorithmically generated injury risk assessments, even in the absence of clinical evidence [22].

In response to such challenges, the global regulatory landscape is evolving rapidly. Responsible AI deployment is supported by general data protection frameworks, such as the General Data Protection Regulation, as well as international guidelines, including the OECD Principles on Artificial Intelligence (2024) and the European Commission's Ethics Guidelines for Trustworthy AI (2019).

Within the sports domain, international organizations also play an important role. The International Olympic Committee (IOC), through its Olympic AI Agenda (2024), identifies key ethical challenges, while the World Anti-Doping Agency (WADA) provides guidelines for managing sensitive data in sport [22].

In this context, national AI governance frameworks should align with these principles, with particular emphasis on transparency, accountability, and interoperability. Such an approach reduces the risk that user data, from elite athletes to recreational participants, remains confined within fragmented and incompatible systems or becomes subject to misuse.

These challenges point to the need for domain-specific governance mechanisms for AI in sport, capable of addressing its unique organizational, ethical, and societal dimensions.

7.2 Technological Safeguards: Federated Learning and Differential Privacy

In addition to legal and institutional measures, protecting privacy in sport requires the application of advanced data science techniques. One such technique is Differential Privacy, which involves introducing controlled statistical noise into datasets, thereby enabling model training without revealing individual user characteristics.

Another relevant approach Federated Learning, which allows models to be trained directly on user devices, such as smartwatches or mobile phones, while only updated model parameters, rather than raw data, are transmitted to a central server [22]. This approach helps reduce the risk of misuse of sensitive biometric and health-related information.

These methods are particularly relevant in the context of mass and recreational sport, where large numbers of users continuously generate and share data related to their health status and physical activity.

7.3 The Black-Box Problem and Explainable Artificial Intelligence

The complexity of deep learning models often leads to the so-called black-box problem, where systems produce highly accurate predictions without providing clear insight into how those decisions are made. Recent studies emphasize the importance of Explainable Artificial Intelligence (XAI) as a means of addressing this challenge [9, 22].

In sport, XAI can provide greater transparency into decision-making processes by illustrating how specific variables, such as movement patterns or physiological stress indicators, contribute to predictions, including injury risk assessments.

This level of transparency is essential for preserving user autonomy. Through the human-in-the-loop principle, the athlete, whether professional or recreational, remains an active participant who can understand, interpret, and critically evaluate system recommendations.

In this context, collaboration between engineers, coaching staff, and athletes is essential for the responsible use of artificial intelligence in sport. Such an approach helps ensure that AI systems function as decision-support tools rather than substitutes for human judgment and experience.

8 Conceptual “360° Integrated AI Model for Sport”: Integration of Professional, Recreational, and Healthcare Systems

In response to the technological asymmetry and institutional fragmentation identified in previous chapters, this paper proposes a national *360° Integrated AI Model for Sport*. The model draws on existing practices, including centralized data management systems used in advanced sport ecosystems, and responds to challenges related to coordination and data exchange.

Unlike the prevailing siloed approach, where data remains confined within isolated organizations or applications, the proposed model aims to establish an integrated, interoperable, and controlled system. Such an approach facilitates the exchange of data and functionalities across multiple levels while maintaining compliance with ethical principles and data protection standards.

The model brings together professional sport, recreational sport, healthcare, and digital technologies within a common framework. In this sense, the model can be viewed as a framework for the digital transformation of sport, aligned with broader public health objectives.

Table 1 presents the proposed model, which comprises five interconnected layers. The elite sport, recreational sport and public health layers represent the primary data domains, while a centralized data and AI research and development layer provides interoperability and analytical capabilities. The strategic and policy layer establishes governance, regulatory compliance and long-term sustainability.

Table 1. Core Components, Stakeholders and Functions of the Proposed 360° Integrated AI

Core Component	System Connected	Key Stakeholders	Primary Function/Value
Elite Sport Tier	Professional clubs, National sports institutes, Biomechanical labs	Professional athletes, Coaches, National Sports Federations	Advanced predictive analytics, performance optimization, and injury prevention using specialized IT infrastructure.
Recreational Tier	Consumer wearables (smartwatches), Fitness apps, IoT sports equipment	Recreational athletes, General public, Amateur sports clubs	Monitoring daily physical activity, tracking fitness goals, and maintaining long-term motivation for an active lifestyle.
Public Health Tier	E-health records (e.g., eKarton, Zdravitas), Clinical databases, Lifestyle monitoring modules	Healthcare providers, Public health institutes, General patients	Utilizing lifestyle and activity data for disease prevention, risk assessment, and bridging sports with medical care.
Centralized AI Data and Research and Development (R&D) Layer	Cloud computing networks, National databases, University research centers	Data scientists, IT administrators, Academic institutions, Faculties, R&D Institutes	Ensuring data interoperability, algorithm development, scientific validation, and secure data exchange across all tiers.

Strategic & Policy Layer	Legal frameworks, Sports governance, Public health policies	Government ministries, Policymakers, Sports Federations, Ethical committees	Harmonizing regulations, ensuring data privacy, and providing institutional support for ecosystem integration.

Figure 1 presents the overall architecture of the proposed model and illustrates the relationships between its main layers.

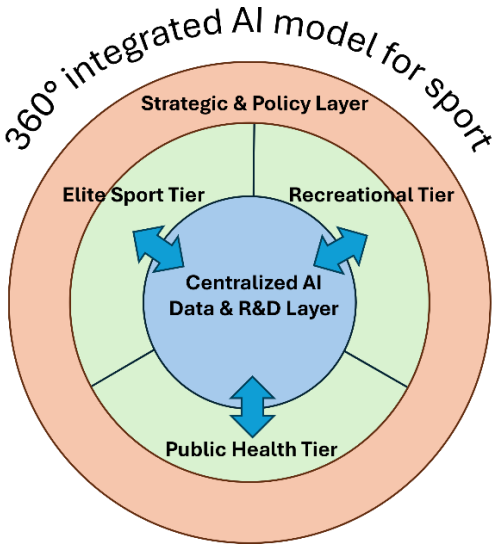


Fig. 1. Conceptual architecture of the proposed 360° integrated AI model for sport.

As illustrated in Figure 1, the model integrates elite sport, recreational sport, and public health systems through a centralized AI data and R&D layer, while the strategic and policy layer provides governance and regulatory support.

8.1 Bidirectional Interface and a Holistic Approach to Performance

In contrast to traditional systems, where the athlete is primarily treated as a passive source of biometric data collected through wearable sensors, the proposed model is based on bidirectional interaction between the user and the system. To improve transparency and user trust in AI-driven processes, the platform is designed to support active user participation.

Algorithms dynamically adapt based on user-provided inputs, particularly those that cannot be reliably captured through sensors. These include subjective measures such as rate of perceived exertion (RPE), daily stress levels, and sleep quality [7, 22]. As a result, the proposed model allows for the integration of objective measurements and individual perceptions of load and recovery.

The importance of such an approach is supported by contemporary research. Several studies suggest that isolated biometric data are insufficient for accurate modeling and that metabolic parameters must also be considered [7]. In this context, the bidirectional interface enables the integration of additional user inputs, including nutrition and hydration data across different phases of activity, before, during, and after training.

The combination of objective and subjective data may support more personalized approaches to training load and recovery management. Such an approach could help reduce the risk of overtraining and contribute to long-term health preservation. It also encourages a more active role for the athlete in the decision-making process.

8.2 Multi-Level Talent Identification and Lateral Sport Transitions

At the professional, youth, and recreational levels, the proposed 360° Integrated AI Model for Sport is intended to support dynamic, longitudinal monitoring rather than relying solely on static observations. A key element of the proposed model is the concept of lateral sport transitions, whereby data-driven analysis could support the redirection of an individual from one discipline to another that better aligns with their biomechanical or physiological profile. Depending on individual goals, this may contribute to competitive success in elite sport, continued participation in recreational activities, or the maintenance of long-term health and well-being.

In elite contexts, AI systems can synthesize large-scale datasets, genetic markers, physiological output, and biomechanical efficiency to pinpoint where an athlete's inherent traits might yield a higher performance ceiling in an alternative discipline [4, 10]. This is particularly vital in youth systems like Serbia's *Zdravitas*, where young athletes often move between sports based on trial-and-error. The proposed model could support a more proactive, data-driven approach to talent development and sport selection, potentially reducing dropout rates associated with a mismatch between an individual's characteristics and the demands of a particular sport.

Beyond competitive domains, this "lateral transfer" logic has significant implications for recreational sports and public health. For the general population, AI-driven analysis may indicate when an individual's physical profile or cumulative injury history suggests that another activity could be more appropriate for long-term participation and health.

The concept of data-supported sport transitions is consistent with the findings of the Ex-ante analysis of sports development in Serbia [14], which highlights the need for more structured monitoring of physical activity across all age groups. Integrating these insights with national health platforms like eZdravlje [13] could support a more preventive approach to healthcare by encouraging long-term participation in appropriate forms of physical activity.

8.3 Mass Sport and the Public Health Dimension

One of the most socially relevant component of the *360° Integrated AI Model for Sport* is its focus on reducing the technological gap and supporting the broader population, particularly recreational participants. In practice, generic algorithms embedded in

commercial wearable devices often fail to account for the heterogeneity of users across different age groups, fitness levels, and health conditions, thereby necessitating a more personalized and adaptive approach.

Several studies indicate that many AI-driven applications are primarily designed for elite sport, while recreational and aging populations remain significantly underrepresented [7]. In response, the proposed model seeks to overcome the limitations of closed, proprietary systems by promoting interoperability and enabling the integration of diverse applications and sensor ecosystems.

For such a system to operate safely and sustainably, integration with healthcare information systems would be essential, particularly with electronic health records (EHR). Subject to appropriate legal frameworks and informed user consent, bidirectional data exchange would allow the AI system to incorporate relevant medical information, including clinical history and laboratory results, before generating personalized physical activity recommendations [17].

Furthermore, integration with existing digital health services, such as lifestyle monitoring modules within national e-health platforms, enables the inclusion of data on daily activity patterns, behavioral habits, and baseline physiological indicators. This could improve the quality and contextual relevance of personalized recommendations.

By combining health data with information derived from everyday physical activity, the proposed model could support a more individualized calibration of training intensity and activity selection. Such an approach may help reduce the risk of injury and adverse health events while strengthening the preventive role of physical activity in public health.

In this way, the proposed model extends beyond performance optimization and contributes to a broader societal objective: the promotion of long-term, sustainable physical activity and the improvement of population health outcomes.

9 Conclusion

This paper examined opportunities for the digital transformation of sport through the proposal of a conceptual 360° Integrated AI Model for Sport tailored to the national ecosystem of the Republic of Serbia. By shifting the focus from isolated, performance-oriented applications toward a broader multi-stakeholder framework, the proposed model outlines a possible approach for connecting sport-related data, public health systems, and institutional stakeholders.

The analysis provides conceptual support for the Main Hypothesis (H_0), indicating that reducing fragmentation within the sport ecosystem requires a more integrated and interoperable digital environment. The analysis also supports the Supporting Hypotheses by suggesting that technical interoperability and resource integration (H_1, H_2) should be accompanied by appropriate regulatory frameworks (H_3) and greater alignment among public policies, digital strategies, academic initiatives, and sport governance structures (H_4).

As a primarily conceptual and strategic study, this research is limited by the absence of empirical validation. The development of a functional prototype and its pilot

implementation involving selected sport organizations and recreational user groups would help assess the practical feasibility, scalability, and broader applicability of the proposed model.

Acknowledgments. The author used AI-based language tools to support translation, grammar and spelling checks, and limited stylistic refinement intended to improve readability during manuscript preparation. AI tools were not used for the generation of research ideas, methodology, analysis, interpretation of results, or conclusions. The author reviewed, edited, and approved the final manuscript and take full responsibility for its content.

Disclosure of Interests. The author have no competing interests to declare that are relevant to the content of this article.

References

1. Adebisi, E. et al.: Leveraging Artificial intelligence (AI) for Stress Management in Peak Athletic Performance: An Integrative Review. Sci. J. Eng. Technol. 2, 2, 94–106 (2025). https://doi.org/10.69739/sjet.v2i2.999.
2. Aleksic, J. et al.: Advancing Field-Based Vertical Jump Analysis: Markerless Pose Estimation vs. Force Plates. Life. 14, 1641 (2024). https://doi.org/10.3390/life14121641.
3. Baladaniya, M., Choudhary, A.K.: Artificial Intelligence in Sports Science: A Systematic Review on Performance Optimization, Injury Prevention, and Rehabilitation. J CLIN MED KAZ. 22, 3, 64–72 (2025). https://doi.org/10.23950/jcmk/16412.
4. Du, T., Bi, N.: Application of Artificial Intelligence Advances in Athletics Industry: A Review. Concurrency and Computation: Practice and Experience. 37, e8372 (2025). https://doi.org/10.1002/cpe.8372.
5. Fadare, S.A. et al.: Machine Learning in Physical Education, Sports, and Recreation: Opportunities, Challenges, and Ethical Considerations – A Systematic Review. Salud, Ciencia y Tecnología. 5, 2251–2251 (2025). https://doi.org/10.56294/saludcyt20252251.
6. Government of the Republic of Serbia: Strategy for the Development of Artificial Intelligence in the Republic of Serbia (2025–2030). Government of the Republic of Serbia, Belgrade (2025).
7. Grivas, G.V., Safari, K.: Artificial Intelligence in Endurance Sports: Metabolic, Recovery, and Nutritional Perspectives. Nutrients. 17, 20, 3209 (2025). https://doi.org/10.3390/nu17203209.
8. Kimi, M.M. et al.: Optimization of sports injury treatment through artificial intelligence: Methods for effective prevention, diagnosis and rehabilitation. Sport - nauka i praksa. 14, 1–2, 1–16 (2024). https://doi.org/10.5937/snp14-1-2-54833.

9. Kranzinger, S. et al.: A scoping review of explainable artificial intelligence in sports science. Discov Artif Intell. 6, 1, 5 (2025). https://doi.org/10.1007/s44163-025-00709-8.
10. Li, W. et al.: A review of artificial intelligence for sports: Technologies and applications. Intelligent Sports and Health. 1, 3, 113–126 (2025). https://doi.org/10.1016/j.ish.2025.05.001.
11. Millington, B. et al.: Sport and the Promise of Artificial Intelligence: Human and Machine Futures. Sociology of Sport Journal. 42, 4, 341–350 (2025). https://doi.org/10.1123/ssj.2024-0150.
12. Ministry of Education of the Republic of Serbia: Zdravitas Platform for Monitoring Physical Development and Motor Skills of Students in Serbia, https://admin.zdravitas.gov.rs/, last accessed 2026/04/24.
13. Ministry of Health of the Republic of Serbia: eZdravlje (eHealth) Platform of the Republic of Serbia, https://hepor.e-zdravlje.gov.rs, last accessed 2026/04/24.
14. Ministry of Sport of the Republic of Serbia: Ex-ante Analysis of the Sports Development Strategy in the Republic of Serbia for the Period 2026–2035. Ministry of Sport of the Republic of Serbia (2025).
15. Namazi, A. et al.: Comparative Analysis of Machine Learning Techniques for Heart Rate Prediction Employing Wearable Sensor Data. Sports. 13, 3, 87 (2025). https://doi.org/10.3390/sports13030087.
16. Office for Information Technologies and eGovernment: State Data Center in Kragujevac (National AI Infrastructure of the Republic of Serbia), http://www.ite.gov.rs/tekst/sr/5273/drzavni-data-centar.php, last accessed 2026/04/24.
17. Palermi, S. et al.: Unlocking the potential of artificial intelligence in sports cardiology: does it have a role in evaluating athlete's heart? Eur. J. Prev. Cardiol. 31, 4, 470–482 (2024). https://doi.org/10.1093/eurjpc/zwae008.
18. Pisaniello, A.: The Game Changer: How Artificial Intelligence is Transforming Sports Performance and Strategy. Geopolitical, Social Security and Freedom Journal. 7, 75–84 (2024). https://doi.org/10.2478/gssfj-2024-0006.
19. Puce, L. et al.: The role of artificial intelligence in sports training: opportunities, challenges and future applications for competitive swimming. Biol. Sport. 43, 1, 355–367 (2025). https://doi.org/10.5114/biolsport.2026.152352.
20. Wang, X.: AI-Based Big Data Platform for Sports Training Construction and Application. International Journal of Information System Modeling and Design. 16, 1, 1–23 (2025). https://doi.org/10.4018/IJISMD.384305.
21. Wu, J. et al.: Artificial intelligence assisted wearable flexible sensors for sports: research progress in technology integration and application. International Journal of Smart and Nano Materials. 16, 3, 510–548 (2025). https://doi.org/10.1080/19475411.2025.2519582.
22. Zhou, D. et al.: Artificial intelligence in sport: A narrative review of applications, challenges and future trends. Journal of Sports Sciences. 1–16 (2025). https://doi.org/10.1080/02640414.2025.2518694.
23. Zhou, Y.: Application of Artificial Intelligence Technology in Sports Training. JOC. 36, 6, 51–60 (2025). https://doi.org/10.63367/19911599202512360600 4.

THEMATIC SECTION

Human Resource Management for Building Resilient Employees

Resilient and Intelligent Organizations
in Times of Permanent Disruption

SYMORG 2026 • CONFERENCE PROCEEDINGS

Original Research Article *https://doi.org/10.5281/zenodo.23476534*

A Mixed Methods Study of Learning Organizations: Implications for Health Care Practice

Jelena Pokimica[1][0009-0009-8597-3672], Jelena Ruso[2][0000-0003-4120-8759], Indiana Laub[1][0009-0005-7729-4224], and Aleksandar Milovanovic[3][0000-0002-9505-6124]

[1]Boise State University, Boise, ID, USA
[2]University of Belgrade – Faculty of Organizational Sciences, Belgrade, Serbia
[3]University of Belgrade – Faculty of Medicine, Belgrade, Serbia

Abstract. The purpose of this study was to quantitatively diagnose seven learning organization (LO) dimensions through the DLOQ measure and the practice of three ISO standards (ISO 9001, ISO 14001, and ISO 45001) in public health care in Serbia. We also tested group differences among doctors, medical staff, and administrative staff on the dimensions and standards (n=85). We later extended the study to include interviews for a broader understanding of learning organizations in the country. We conducted six extensive qualitative interviews with doctors in various medical specialties at five secondary and tertiary health centers across Serbia, in 2025 and 2026, regarding their experiences with LO and recommendations. In quantitative analyses, we included graphical representations and descriptive statistics, and Kruskal-Wallis tests comparing median ranks, followed by pairwise comparisons. Overall, the results revealed that strategic leadership, systems connections, and applications of ISO standards were more present, while other LO dimensions had shortcomings, such as embedded systems and dialogue and inquiry. Thematic analyses of qualitative data revealed three emergent themes: 1) Diagnostic advancements, continuous learning, and informal knowledge of ISO; 2) Multidimensional barriers to organizational learning: Power dynamics, interpersonal communication, and divergent approaches to medical problem-solving; and 3) Embedded systems and training and development needs. This study contributes to drawing attention to several affected LO dimensions in public health care. Implications for structural, human process/HRM, and strategic interventions are drawn, as are linkages to past studies.

Keywords: Learning Organizations, DLOQ, Health Care, Mixed Methods, Serbia

1 Introduction

Health systems are intense and complex organizations, with high demands and time-critical situations, which call for employee readiness, adaptability, and vigilance. In order to address these realities, employee learning and knowledge sharing become crucial for promoting quality care for patients and aligning with organizational mission, vision, and goals [1]. In this sense, organizational learning combines intentional and unintentional processes of formal and informal learning and an organization's capacity

to create, share, and use knowledge in response to health care events and environmental changes [1, 2]. Learning culture creates conditions for learning in a health system, which can gradually transform into a learning organization, if such culture is strong. Nowadays, internal and external cooperation in organizations, through the mechanisms of learning and resilience, is well-remarked, and it particularly contributes to organizational creativity or elasticity in handling crisis [3]. Learning organizations (LO) excel at organizational learning [4]. They can also be considered a product of organizational learning [5]. Per Peter Sange, five key principles characterize these organizations: personal mastery - with learning and change-adaptive mindset, systems thinking, mental models, shared vision, and team learning. These principles are cultivated through building trust in organizations through open communication about successes and failures, investing in training and development, rewarding employee learning new skills, leading by example, and building knowledge-sharing platforms. Marsick and Watkins [6] built on this and related work and developed the Dimensions of the Learning Organization Questionnaire (DLOQ), which is a validated instrument used across the world in various sectors, such as health care, in both long and short versions of the scale and translated into various languages [7-11]. The instrument contains seven dimensions: 1) continuous learning, 2) dialogue and inquiry, 3) team learning and collaboration, 4) embedded systems, 5) empowerment, 6) systems connections, and 7) strategic leadership.

Per Alrashidi et al. [7] and Leufvén et al. [10], continuous learning implies learning on the job and ongoing employee growth and development opportunities, while dialogue and inquiry accommodate for expressing various perspectives, listening, inquiring into others' views, receiving feedback, and experimenting. Team learning and collaboration are recognized as a valuable component of organizational culture whereby group/team members learn from each other. Embedded systems represent knowledge management systems, as part of organizational memory, for creating, organizing, and distributing knowledge, as well as for tracking past and future performance gaps (via intranets, digital knowledge repositories, and other HRM platforms). Empowerment means that people partake in a common vision, take initiative, and have a sense of ownership and accountability. Systems connections imply an organization as an open system, connected to its external environment (e.g., community), with performance effects visible across the organization. Finally, strategic leadership signifies that leaders support organizational learning and use it strategically to achieve organizational effectiveness [7, 10-11]. Organizational learning and knowledge management are important parts of organization development interventions [12]. Of note is that employees may learn in an organization, but such learning may not always constitute organizational learning in a broader sense (beyond individual level) if not shared among others in the organization; if not embedded in organizational culture, structure, and systems; and if not achieving particular organizational purposes [13]. The inverse is also possible – organizations may learn, while individual learners may not be updated with such learning, or there may be a lack of adequate training.

Literature points out that growing a health care system into a learning organization could be beneficial for health care outcomes [1]. However, this is not a simple endeavor. Leadership, power dynamics, and organizational structures matter for the promotion and propagation of LO, as well as a learning culture orientation. For example, in a study

in Greece, bureaucratic procedures and top-down decision-making in public health care were associated with the lowest scores in empowerment from the seven DLOQ dimensions [9]. In the same study, transactional leadership was found to inhibit learning, and the authors recommended a shift toward transformational leadership [9]. Some variations were noted across specific dimensions in terms of gender and age. For example, older participants were more likely to score higher on inquiry and dialogue. Chughtai et al. [14] found that vertical hierarchies diminished the likelihood of open dialogue and creative risk-taking in the pharmaceutical sector and, thus, suggested adaptive leadership. Ju et al. [15] found a moderating effect of power distance on the relationship between DLOQ and organizational performance (i.e., overall innovative performance, financial performance, and knowledge performance/products, services, intellectual capital). In a study from Nepal, DLOQ data were obtained from healthcare workers in different roles (nurses, doctors, students) demonstrating notable differences on empowerment and strategic leadership dimensions of LO [10]. Nurses scored low on empowerment, while doctors scored lower on strategic leadership.

Interestingly, one of the studies used DLOQ not just as a diagnostic tool to assess the presence of learning culture, but as a framework for performance enhancement, that is, how practicing the seven dimensions could lead to improved outcomes [7]. This approach may offer valuable guidance and inform practical strategies for development and reform in health care institutions. Due to variations in national (and organizational) cultures and in order to reduce bias, it is recommended to supplement DLOQ with other performance measures and include qualitative data collection, such as interviews, peer assessments, or document reviews [11]. Across cultures, DLOQ has been applied in various health care settings (e.g., small and large hospitals), in Nepal, Romania, Saudi Arabia, Greece, and other countries, showing its versatility [7-11]. Goula et al. [9] included six general public hospitals in the Attica region, capturing a mix of hospital sizes and departments. The DLOQ instrument has also been implemented and validated in smaller private practices as seen in Romania where the tool was used in a small private ophthalmology practice [8]. Studies from Saudi Arabia featured multiple government and private hospitals [7] and those with large bed capacity of over 700 beds, with inpatient and outpatient care [1]. Alrashidi et al. [7] point out that the health care institutions in Saudi Arabia have been massively developing over the last decade and a half, aiming to become learning organizations, as part of the Saudi Vision 2030. The DLOQ has also been used in the National Capital Region of India, in a tertiary health care center, which is a teaching hospital for nursing students, as well as in a dental institution [16, 17]. Results from the studies implied that the DLOQ tool could be used for the assessment of multidimensional individual and organizational learning capacity. Watkins et al. [18] employed the DLOQ in four public health care departments in the Mid-Western United States. The continuous learning dimension was rated the highest, with the most room for growth on the empowerment dimension. In 2025, Bostwick et al. [19] psychometrically refined the Learning Organization Survey (LOS-27) [20], with an added open-ended question, when measuring learning culture in a pediatric hospital delivery room, among medical professionals in diverse roles. There is an overlap between some of the LOS sub-scales (e.g., leadership, supporting learning environment) and DLOQ, although performance monitoring, time for reflection, experimentation, and items on

forums for meetings with experts from outside the organizations, other divisions, or patients/families are more specific, as are items on training when shifting to a new position. In other words, there are variations in scales depending on the specific organizational (diagnostic) focuses. There are also some overlaps between research on high-reliability organizations for the prevention of harmful events, in the context of Veterans Health Administration [21] and LOS and DLOQ-based research, relative to the system connections dimension and learning from others in the community. The literature highlights vibrant developments in learning organization measurement.

1.1 The Serbian Context

Serbia has an established comprehensive national public health care system financed by payroll taxes (10.3%), although it also has private care, which is more expensive. Health insurance is compulsory. The public health care system is maintained by the National Health Insurance Fund, the main purchaser of health services, while governed and organized by the Ministry of Health [22]. Primary care is free, although patients could incur higher costs for specialized care and medications, and they may experience longer waiting times for complex procedures/surgeries in more advanced centers. The public health care system combines primary care health centers, secondary and tertiary institutions (with smaller and larger hospitals respectively) and clinical centers, across regions. Public health policy restructuring has introduced registering with a doctor of choice in primary care [22].

As seen in the literature, previous studies employed the DLOQ instrument in different countries and health settings, and there is an opportunity to fill a gap in the Serbian context and explore the presence of learning culture in its public health care, as well as how medical professionals experience the LO dimensions and if they have any recommendations. Our study aim is to contribute in this regard. We also add three international standards as relevant to health care, namely, ISO 9001 International standard for quality management systems, ISO 14001 International standard for environmental management systems, and ISO 45001 International standard for occupational health and safety management systems, in an exploratory way. There have been other explorations of DLOQ expansions [23]. We find that the three standards are applicable to our study because they help hospitals/health centers ensure quality and safety of medical care and protection of internal and external environment.

ISO 9001 focuses on operational effectiveness, risk management-related policies, continuous improvement, and patient-centered health care [24]. Recent studies found that ISO 9001 increased inter-departmental collaboration and managerial safety expectations, but it did not reduce punitive measures for errors or administrative certification burdens [24]. In yet another study [25], ISO 9001 was used for surgical quality and reduction of surgical site infection rates. The standards necessitated committed leadership; ISO 9001 specifies requirements but offers flexibility with the range of quality management system tools that could be used. ISO 14011 uses the Plan-Do-Check-Act approach for preventing pollution, continuous improvement, and legal compliance. It compels organizations to perform gap analysis among other major stages for certification. ISO 45001 helps prevent injuries, death, illnesses, and workplace accidents. It uses

similar approach as ISO 14001 and calls for committed management and staff. The three mentioned ISO standards closely relate to several DLOQ dimensions, such as continuous learning, team learning and collaboration, embedded systems, and strategic leadership. It could be informative to juxtapose compliance versus organizational learning aspects of ISO and DLOQ. It is even possible that health centers with the same ISO standards may have different DLOQ outcomes based on the nature of their learning cultures.

From the presented background, the concept of learning organization seems to be an important factor in health care, and it is with that grounding that we ask the following three research questions in our study:

RQ1: What is the current diagnostic status of the DLOQ dimensions and ISO standards in public health care settings in Serbia?
RQ2: Are there significant differences among doctors, medical staff, and administrative staff on DLOQ dimensions and ISO standards?
RQ3: In what ways do doctors describe their experiences with DLOQ dimensions in the context of their work at public health centers/hospitals? How familiar are they with ISO standards?

1.2 Integrated Theoretical Framework of Organizational Learning

We present an integrated framework of organizational learning (see Fig. 1). First, we draw from the 4I framework of organizational learning [26], which describes how learning moves through organizational levels (individual, group, organization) and enters the realm of institutional knowledge - in other words, how individual and group insights are transformed into organizational learning. This framework focuses on the psychological and social processes of intuiting, interpreting, integrating, and institutionalizing (4Is). In the 4I framework, intuiting and interpreting arise at the individual (e.g., medical doctor) level at a health center; interpreting and integrating take place at the group/team level; and integrating and institutionalizing occur at the highest (organization) level. Intuiting is the process of gaining insight on a pre-conscious level, when an individual recognizes past patterns or future possibilities, rooted in personal experiences. Individuals who engage in this process could influence others through interactions. In the process of interpreting, individuals explain ideas and insights to self or others in words/actions. Integrating involves dialoguing, shared understanding, coordinated action, and group adjustment. Integration usually starts informally, but if it evolves into recurring coordinated action, learning can be institutionalized. That means that it can be formally embedded into knowledge management systems (KMS), using knowledge bases, learning management systems, document management software, collaboration tools, and similar, in the form of procedures, lessons learned, strategies, policies, and structures. In sum, interpreting is the binding agent between individual and group/team levels, and integrating between the middle and top level.

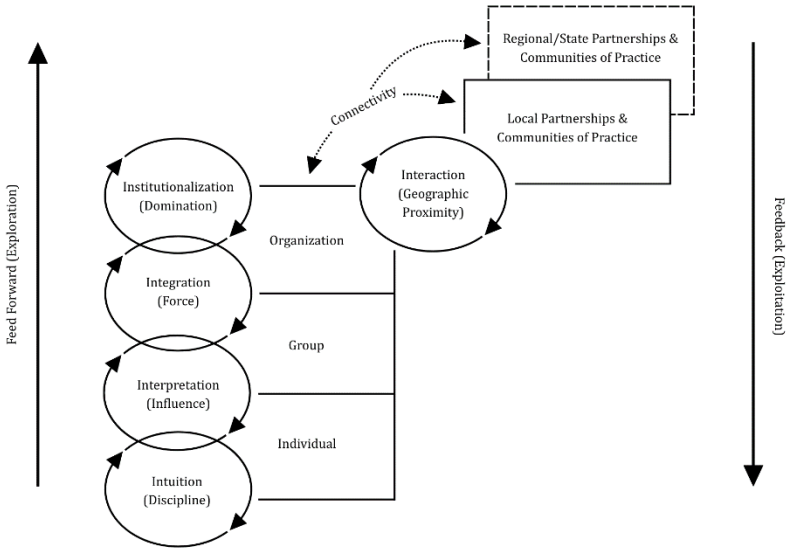


Fig. 1. Integrated framework of organizational learning adapted from Crossan et al. (1999), Lawrence et al. (2005), March (1991), Wiewiora et al. (2020), and Wodnik et al. (2024)

The interlinked nature of these recursive processes and levels emphasizes that learning may not immediately leap from the individual level into the lines of institutional policy, but that it moves within and across organizational levels. This holistic view of organizational learning highlights the complex interplay among micro, meso, and macro forces. Organizational learning allows organizations to achieve strategic renewal through exploration (the assimilation of new learning) and exploitation (the use of what has already been learned) [26, 27], as represented by the feed forward and feedback arrows in our figure.

Lawrence et al. [28] delineated four forms of power, specifically linking discipline to intuition, influence to interpretation, force to integration, and domination to institutionalization. These forms additionally frame our thinking into why some organizational learning becomes integrated or institutionalized and some does not. The forms of power that influence interpretation and integration occur in the episodic mode from bottom-up (via discrete/autonomous acts by individual self-interested actors), while institutionalization occurs in a systemic mode from top-down (power is diffused through the system). Interestingly, Lawrence et al. [28] cite past research on how the systemic mode surprisingly primarily operates in the intuiting process, as well, even though we may think of individual imaginations being only internal to self. This is because procedures, structures, and systems of individuals' working lives largely affect their experiences, behavior, and perception. Power dynamics in health care are, thus, important mechanisms for unfolding of the 4I processes.

Recently, Wodnik et al. [29] extended the 4I framework by including two more pathways for inter-organizational learning at the global level: 1) Interaction - informal partnerships and communities of practice and 2) Incorporation - macro-level coordinating bodies (health organizations). Arguably, health centers within a given country, such as Serbia, could also capitalize on interaction, that is, share knowledge via partnerships and communities of practice locally and across regions of Serbia (including virtually), so we adapted interaction to locales and regions/state. This exchange of knowledge is facilitated by geographic proximity, as local-level connections tend to be stronger and yield more frequent knowledge transfer [30], being first-points of interaction, followed by regional or state ones when local resources/interactions are exhausted. Lastly, Wiewiora et al. [30] identified the connectivity of senior leaders as a resource that influences the ability of health institutions to benefit from the learning exchange that stems from partnership with other organizations at varying levels of geographic proximity. Connectivity, power, and proximity can all thus be understood as factors that facilitate organizational learning processes.

2 Method

This is a mixed methods study whereby we first administered the DLOQ instrument (21 items, translated into Serbian, 1-6 rating scale: almost never – almost always) to 85 respondents, followed by six qualitative interviews meant to probe more deeply into quantitative findings, in secondary and tertiary health centers across the country. Interview participants also took the DLOQ survey prior to interview. The following subsections explain our sample, mixed method design, and data collection and analyses in more detail.

2.1 Samples

For the quantitative portion, our sample was comprised of 85 respondents: doctors (41%), medical staff (38%), and administrative staff (21%). Seventy nine percent were female. Respondents' age ranged from 45-54 years (27%), 55-64 (25%), 35-44 (22%), and 25-34 (19%); the rest did not disclose. The following is a breakdown of their academic status: specialized high school (32%), specialization beyond regular MD (20%), college degree (20%), MD/Bachelor's degree (17%), Master's degree (7%) and Doctoral (3.5%). One fifth of them had 31 or more years of experience, one fifth also had one to five years of experience, and the rest were relatively evenly distributed across years of experience, mainly from 10 to over 20 years; only five percent had less than one year of experience. Their time spent on learning monthly ranged as follows: 1-10h (34%), 11-20h (33%), 36+h (18%), 0h (7%) and the rest (8%) spent anywhere from 21h to 35h. We used maximum variation purposive sampling [10] in order to reach a diverse set of cases for the study of the LO phenomenon in Serbian health care. All data were collected in the Serbian language. For the qualitative portion, we interviewed six doctors from the following medical specialties: internal medicine, general surgery, anesthesiology/reanimation/intensive therapy, pediatrics, and obstetrics and gynecology. Their ages ranged from early 30's to early 40's (four) and early 50's to early 60s (two).

Two were females and four males, all with a specialization in their disciplines and one PhD. They spent about 11-20h on learning per month, on average.

2.2 Data Collection, Analytic Techniques, and Mixed Methods Design

All quantitative data were collected in May of 2025 and qualitative data in the fall of 2025 and early spring of 2026 (including six additional DLOQ surveys, pre-interview). Most of the semi-structured interviews were conducted in-person and two online. We also took field notes. Example interview protocol questions related to decision making for requesting additional resources to meet work needs, successes and barriers to organizational learning, and recommendations for any of the LO dimensions; these questions were also accompanied by a series of relevant probes. We additionally inquired about ISO applications as a unique addition to DLOQ, in an exploratory way. All participants in both qualitative and quantitative stages consented to participate and remained in the study until the end of each stage.

Quantitative data were analyzed in SPSS 30 using descriptive statistics and non-parametric Kruskal-Wallis tests for significant differences among independent groups, with pairwise comparisons. Qualitative inquiry centers on meaning in the health practice context [31]. The mixed method design combines quantitative and qualitative data in order to present a more complete picture of LO realities in public health care in Serbia. More precisely, we followed an explanatory sequential design type of mixed methods [31] and collected quantitative data first, then qualitative, for the purpose of a richer understanding of our quantitative findings (from closed-ended to open-ended items). Qualitative research is particularly suited for probing more deeply into significant findings from surveys, especially intriguing, curious, and surprising findings, revealing the ”why” behind certain occurrences and including how a given phenomenon (e.g., a learning organization) manifests. Although mixed methods often take more time to complete and necessitate separate sets of analysis, we could capitalize on the structured questions in the quantitative DLOQ instrument and still capture unanticipated aspects of the LO dimensions through interviews, which help with quantitative interpretations in a holistic way [32]. Qualitative data were thematically analyzed [33] using constant comparisons and data coding, from open coding to collapsing similar codes into categories and ultimately generating aggregate themes in spreadsheets suitable for the size of our project. We plan to add several more interviews to achieve complete saturation into thematic insights in the summer of 2026; nonetheless, we still believe that presenting preliminary qualitative findings will be informative as to some major directions to which they point and the kind of initial learning organization pattern they trace.

3 Results

From the descriptive analyses of quantitative data (n=85), item 19 on Dimension 7 for strategic leadership (“At my organization, leaders mentor and coach those they lead”) had the highest mean (4.26). ISO standards were also rated highly, especially ISO 9001 for quality management systems [“Formalized knowledge in a form of an international standard for quality management system (ISO 9001) is applied at my organization;”

mean=4.38]. Item 3 on Dimension 1 for continuous learning ("At my organization, people are rewarded for learning") and item 10 on Dimension 4 for embedded systems ("My organization creates systems to measure gaps between current and expected performance") were scored the lowest (mean=3.29 for both). Mean values for all the dimensions per each group (doctors, medical staff, administrative staff) are presented below (see Fig. 2). From our analyses of six additional surveys just prior to interviews, less familiarity with ISO applications was noted, although new leadership was positively rated in terms of new equipment acquisition and acting in line with organizational values. Promoting opportunities for new learning was lacking. More rewards for self-initiated learning on the job were desired in pediatrics. Dimension 3 for team learning was more prominent - such as team discussions and revising thinking as a result of those - but formally established knowledge management systems (Dimension 4) were not present. These results imply that integration from the 4I theoretical framework was present at the team level, but that it did not rise to the level of institutionalization via embedded systems.

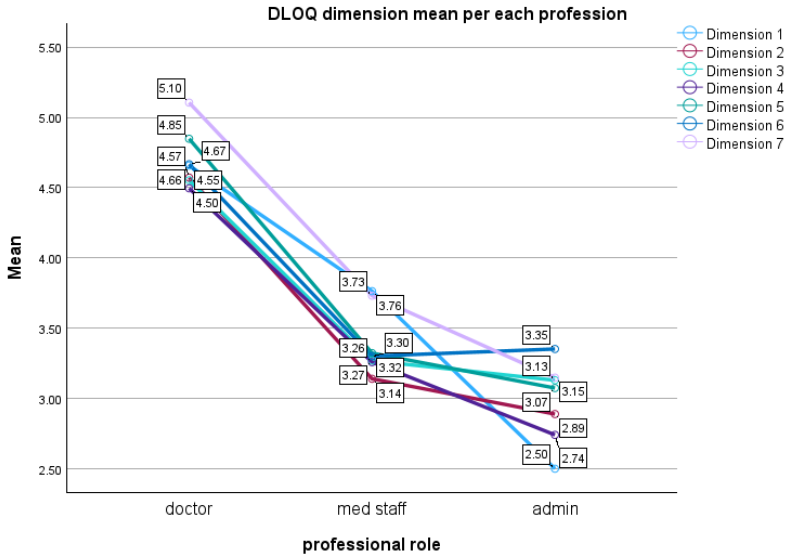


Fig. 2. Dimension mean per professional group (n=85)

ISO standard means also ranged in a similar fashion: 5.10 (doctors), 3.96 (medical staff), and 3.06 (administrators). Group difference analyses are presented in Table 1.

Table 1. Group differences

Professional group	Dimension	1	2	3	4	5	6	7	ISO standards
Doctors	N	35	35	35	35	35	35	35	35
	Mean	4.66	4.57	4.55	4.50	4.85	4.67	5.10	5.10
	St. Dev.	1.16	1.27	1.18	1.14	1.29	1.13	1.11	1.15
Medical staff	N	32	32	32	32	31	31	31	31
	Mean	3.70	3.08	3.21	3.20	3.32	3.30	3.73	3.96
	St. Dev.	1.06	1.37	1.26	1.26	1.51	1.29	1.40	1.55
Administrators	N	18	18	18	18	18	18	18	18
	Mean	2.50	2.89	3.13	2.74	3.07	3.35	3.15	3.06
	St. Dev.	1.45	1.04	1.18	1.12	1.17	1.27	1.22	1.23
	Kruskal-Wallis test (df) & p-value	25.2 (2) <.001	24.7 (2) <.001	22.3 (2) <.001	25.1 (2) <.001	22.2 (2) <.001	19.7 (2) <.001	27.0 (2) <.001	22.9 (2) <.001

Pairwise comparisons were conducted and significant differences were found between doctors and medical staff and doctors and administrators on all dimensions and ISO standards. For Dimension 1 (continuous learning), a statistically significant difference was additionally found between medical staff and administrators.

3.1 Thematic Findings

The following are qualitative findings across three main emergent themes.

Diagnostic Advancements, Continuous Learning, and Informal Knowledge of ISO. Across six interviews at five secondary and tertiary health centers in the country, doctors elaborated on their experiences with several DLOQ dimensions, especially team learning and collaboration, empowerment, embedded systems, and strategic leadership. Although suggestions for additional resources could be made by medical doctors, they could only suggest but not decide – for example, a request for an ionogram must be approved by the State Health Insurance Fund (sr., RFZO). Still, with an investment of close to a half a million euros, one of the secondary level health centers now has new endoscopic surgery procedures with reduced surgery and recovery time (fast-track), plus less time away from work for patients. Also, at that hospital, "medical cadre is more motivated to stay because of learning on the job and up-to-date education where they are globally competent" (per the General Surgery Specialist). Yet another tertiary health center now has more sophisticated diagnostics for liver tissue status, plus ERCP as a minimally invasive procedure with no anesthesia for bile and pancreatic duct treatment, avoiding surgery. With new leadership at one of the smaller hospitals, more types of diagnostics are now carried out (e.g., Doppler, Holter, Roentgen). In that setting, ISO 9001, 14001, and 45001 are applied, although primarily for compliance and legal reasons (especially ISO 45001 standard for occupational health and safety management system). In one of the centers with a large hospital, ISO 45001 is highly applied, especially regarding instrument sterilization and equipment, standardized patient reanimation measures, and annual physical exam for employees; however, there is less familiarity with how the other ISO standards are used and if certifications are secured. In

general, ISO applications and adherences were perceived as part of the management/administrative role in all five health care centers. In one of the tertiary care centers, continuous learning at conferences (e.g., Italy) and medical education points for intensive care were practiced, and a new pain management center was opened. Attending scientific congresses within Serbia was financially supported (up to RSD 15000 and 5 days of absence from work for doctors). International conference attendance was more self-initiated, per data across multiple health centers.

Multidimensional Barriers to Organizational Learning: Power Dynamics, Interpersonal Communication, and Divergent Approaches to Medical Problems. Major barriers to organizational learning in public health care were: lack of enough time to engage in all medical activities and organizational learning, discrepancies between Eastern and Western approaches to medical problem-solving, power dynamics, interpersonal communication and relations between more senior doctors and early-career ones, and lack of considering initiative-taking as an important component of empowerment.

From the expanded 4I framework relative to power dynamics, some tensions were present between senior and early-career doctors, which pointed to empowerment differentials on Dimension 5. For example, one of the doctors explained:

> We should give a chance to younger doctors, improve supervisor-employee relations and the management structure, and use a humanistic approach to treating others; we have enough stress already and our work is hard, so no additional conflict and problems are desired.

Empowering early career doctors for initiative-taking is highly desired since new ideas quickly get "cut at the root", per the same participant, as well as boosting dialogue and inquiry (Dimension 2). In this sense, intuiting, interpreting, integrating, and empowering would need to operate together (i.e., 4I components and Dimensions 2 and 5). Influence, as a power form, operates in the interpretation process when those with domain-specific expertise and social skills take control or pursue an idea [28]. Even more so, supervisors with formal authority may use force as a power form and impose restrictions on alternative perspectives and redirect agendas [29]. This power form is also the most prominent when attempting to move from integration to institutionalization in the integrated theoretical model, that is, from team to organization level. Domination is said to operate at the top level, trickling down to the rest of the levels. Regarding Eastern and Western medical approaches, one of the internal medicine specialists, with about 30 years of experience, stressed the importance of expert recognition in complex medical decision-making even if a problem is not in their direct job description (Eastern approach) rather than a sharp separation between roles (Western approach). It seems that adaptations between the two approaches may be useful to accommodate for cultural context particularities and experience. There was a deeper meaning implied in the attestation that "I guarantee for a young and rising doctor that my word/approval means something." Several doctors across health centers shared that more competent professionals in certain medical and management positions were needed, as well as better pay rates for employees.

Embedded Systems and Training and Development Needs. This theme captured opportunities for systematizing organizational learning into formal knowledge management systems (Dimension 4) for tracking gaps between current and desired performance and institutionalizing learning (organization level of the integrated framework). Evidence of shared knowledge existed at the team level, for example, but mainly in oral form, with potentially more chance for such knowledge to get lost versus if it were in written/electronic form. Continuous learning about various areas of the health system is widely recognized as important for keeping current with health care innovations, which links to interactions at the regional/state levels (and beyond), as presented in Figure 1. Regular meetings with management and medical personnel during which reviews of the provision of services are discussed and updated were highly desired in one of the secondary care centers.

Piloting a formation of specialized committees at the national level, along with consultants from other relevant institutions (see connectivity in Fig. 1) to periodically evaluate and (re)direct doctors toward further training, if needed, as frequently done in the West, was recommended. Some resistance toward this proposal by various hospitals was also anticipated. To mitigate both advantages and disadvantages of this approach, borrowing from the integrated framework would be beneficial, such as via partnerships and communities of practice platforms [29] which could connect health centers across regions of Serbia, so that they could share organizational learning if this pilot program were to take place. These findings link to Dimension 6 of the DLOQ on systems connections. In addition, one of the doctors shared how integrating good practices, guidance, and health care protocols from the state guides could be an area of improvement in secondary care for general surgery. Specialized training and development interventions were needed in some of the tertiary centers, specific to endoscopy in internal medicine. For example,

> You study everything on your own; there is no transfer of knowledge; more practical [endoscopic] training is needed; hiring more medical professionals would help to allow time for specialists' training… The only lectures that are available are through pharmaceutical firms.

This finding points to the need for structural interventions to allow for transfers of individual to group- and to system-level learning if training for internists were to take place, as well as possibilities for multi-directional knowledge sharing flows – feedback and feed-forward, based on the integrated theoretical framework. Hospital leaders and more senior doctors could also engage in mentoring and supporting new doctors besides directing or delegating, especially if certain knowledge or skills need to be honed for early-career doctors, via situational leadership.

4 Conclusion

This mixed methods study brought to light the status of DLOQ dimensions at public secondary and tertiary health centers across a suite of medical specialties in Serbia. We were ambitious in our research pursuits, as this kind of work is just beginning to paint

a broader picture of organizational learning in Serbian health care and to highlight opportunities for further development. We wanted to gain a diverse understanding of organizational learning realities in these institutions specific to strategic leadership, embedded systems, employee empowerment, continuous learning, systems connections, dialogue and inquiry, and team learning, as well as any familiarity with ISO. Overall, quantitative results revealed that strategic leadership, systems connections, and applications of ISO standards were more present, while other LO dimensions had shortcomings, such as embedded systems and dialogue and inquiry. Compared to quantitative results from other studies from countries in South-East Europe, mean dimensional values from our study were higher than those in the Romanian study [8] on the 21-item scale. Organizational (and national) cultures in these two countries are similar. However, strategic leadership (Dimension 7) was dropped upon a pilot test in the Romanian study as the participants had less familiarity with it. Accounting for health setting sizes and types in these two studies is also important. The Greek study used the longer DLOQ scale [9] and most of the dimensional mean values were lower than 3 on a 5-point scale (Dimension 2 mean was the only one higher than 3).

Moving beyond the region, the dimensional values from Alrashidi et al.'s [7] study from Saudi Arabia were markedly higher than in the other studies [1, 9, 18]. Interestingly, this and our study were the only ones in which the strategic leadership dimension crossed the mean value of 4. This could be the case because the quantitative sample in our study was drawn from a more prominent, specialized, and larger tertiary health and teaching institution, while the qualitative sample was more dispersed across the country. The overall status of the DLOQ dimensions in our study was higher than in the Indian studies, except more noticeably on dialogue and inquiry (Dimension 2), which was higher in the study by Kumar et al. [16], based in a tertiary center and simultaneously a teaching hospital for nurses and medical interns. The same could be said when our study is compared to the Egyptian one [3]. These comparative studies also used the 21-item scale. Developing Dimension 2 more in Serbian health care would be valuable. This dimension is about an institution's capacity for generating a culture of inquiry, feedback, and experimentation, and it may be tied to different value systems and health system structures across countries. Our qualitative findings probed more deeply into these nuances and painted a more detailed picture across tertiary and secondary health centers in Serbia.

Three emergent themes were found in qualitative analyses: 1) Diagnostic advancements, continuous learning, and informal knowledge of ISO; 2) Multidimensional barriers to organizational learning: Power dynamics, interpersonal communication, and divergent approaches to medical problem-solving; and 3) Embedded systems and training and development needs. Relative to the first theme and concerning ordering new equipment and instruments and updating diagnostic and surgical procedures, there has been a remarked progress across all health centers in the study. In terms of continuous learning, scientific conferences were highlighted, especially sponsored conferences within the country, with paid work leave and financial coverage. Sponsorship was less present for international conferences and specialized training abroad; that kind of training had to be self-financed, which hindered career development goals of some doctors. There were no systematized support and reward systems for self-initiated continuous

learning among doctors (Dimension 1), which was a prominent finding in pediatrics and internal medicine. There was an exception in one tertiary health center, where medical education fairs were accredited and where receiving education points was practiced. Advances in pain management therapy figured in some of the most recent educational innovations, and that center sent its doctors to specialized training in Italy. Rotational scientific congress attendance in intensive care was also practiced, so that more doctors could have those experiences in an equitable way. As far as other congresses, those are financed by pharmaceutical companies, which was a recurring finding from the analyses. Overall, doctors were mainly informally familiar with ISO standards across health centers, with less insight into ISO certifications, so they resorted to management, finding it beyond their scope.

Regarding the second theme, doctors desired more decisional input and leadership support for employee development. Power differences were present in these findings relative to how different processes from the integrated framework played out from one organizational level to the next, and especially in terms of inter-generational gaps and seniority. Early-career doctors had a difficult time being recognized by their supervisors for initiative-taking and offering alternative perspectives. Possibly, hierarchical leadership structures played into these realities, with various forms of power (influence, force, domination) at more than one organizational level. Consequently, there is an opportunity for further development of Dimension 3 - team building between supervisors and employees and engaging in joint interpreting and integrating from the integrated framework of organizational learning. We also concur with Sasaki et al. [34] who studied organizational learning from incidents in three hospitals in Japan, positing that empowering behavior of supervisors helps employees better deal with issues through bolstered motivation to learn. Our study implies strategic, transformational, and empowering leadership innovations that support employee development and promote a learning culture, going beyond primarily equipment/logistical supports or operational needs. The empowerment dimension in DLOQ deals with support for initiative-taking, having control over resources to accomplish work, and collective visioning, as some examples. Transformational leadership imparts interest for others, as well, beyond self-interest. Such innovations could also enhance organizational commitment [35]. Probing further into leadership structures in Serbian health care is suggested in future studies. Regarding ISO standards, doctors were only informally familiar with them, with less insight into ISO certifications; so, they resorted to management since they found it beyond the scope of their role.

Concerning the third theme, major implications from our findings point to: structural interventions for formalized embedded knowledge management systems for tracking performance across services; sharing lessons learned; identifying and addressing training/employee development needs (Dimension 4); and increasing rewards for employee learning new skills and enhancing knowledge (Dimension 1), across all interview participants. There were gaps in leaders' recognizing training needs of medical specialists, especially in endoscopy. Interestingly, we found that dual medical and administrative (leadership) roles created continuous learning opportunities (Dimension 1) in secondary level care. For example, once-a-month, one of the internists, who is at the same time in charge of his unit as an administrator at a secondary institution, travels to a larger

(tertiary) health center in another region of the country, gaining exposures to other high-level specialists (e.g., nephrologists) and knowledge exchange. This is one of the crucial examples of how connectivity from our integrated theoretical framework functions, accounting for geographic proximity. Having better systems and formal mentorship transfer would facilitate the spread of this knowledge back to the local setting. Knowledge, skills, and attitudes (KSAs) that employees gain in organizations with high learning culture orientation contribute to a resource pool in organizational memory to which others could turn for information when they need it [36]. We relate this finding to embedded systems (Dimension 4) and systematic knowledge management approaches. Based on qualitative findings, there also seem to be some variations in adoption of good practices and protocols from national guides, but those findings have only surfaced, so future studies should investigate such adoption in more detail.

4.1 Limitations

In our qualitative portion of the study, we have focused only on doctors, so it is possible to include other medical personnel in interviews for more comprehensiveness in the future. Regarding our qualitative stage, we plan to continue interview data collection for more analytic saturation across LO insights, and this stage of our study is very recent. Our research is also paving the way for studying LOs in Serbia.

Acknowledgments. This study was conducted through independent collaboration of researchers from the United States and Serbia. We acknowledge medical point persons in all health centers who facilitated our data collection processes in adherence to ethical science.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Alatawi, S. R. F., Salem, O. A., Nazik, Z. M. A.: Nurses' perception of learning organization dimensions. Int. J. Med. Health Sci. Res., 9(1), 10–17 (2022). https://doi.org/10.18488/9.v9i1.2953
2. Marsick, V.J., Watkins, K.E.: Informal and Incidental Learning in the Workplace (Routledge Revivals). Routledge, London (2015)
3. Elsayed, W.A., Abdel-ghani, A.M.: Learning organization and its influence on organization resilience and creativity in Mansoura Oncology Center. Egypt. J. Health Care, 11(2), 485-499 (2020). https://doi.org/10.21608/ejhc.2020.147928
4. Tsang, E.W.K.: A preliminary typology of learning in international strategic alliances. J. World Bus. 34(3), 211–229 (1999). https://doi.org/10.1016/S1090-9516(99)00016-4
5. McGuire, D.: Human Resource Development. 2nd edn. Sage, Thousand Oaks (2014)
6. Marsick, V.J., Watkins, K.E.: Demonstrating the value of an organization's learning culture: the Dimensions of the Learning Organization Questionnaire. Adv. Dev. Hum. Resour. 5(2), 132–151 (2003). https://doi.org/10.1177/1523422303005002002

7. Alrashidi, N.A., Lagura, G.A.L., Celdran, M.C.B.: Utilization of the dimensions of learning organization for enhanced hospital performance. Front. Commun. 8, 1189234 (2023). https://doi.org/10.3389/fcomm.2023.1189234
8. Gheorghe, C.-M., Purcărea, V.L., Gheorghe, I.-R., Popa-Velea, O.: Investigating the dimensions of learning organizations questionnaire (DLOQ) in a Romanian private ophthalmology organization. Rom. J. Ophthalmol. 62(4), 288–295 (2018). https://doi.org/10.22336/rjo.2018.44
9. Goula, A., Stamouli, M.-A., Latsou, D., Gkioka, V., Sarris, M.: Validation of Dimensions of Learning Organization Questionnaire (DLOQ) in health care setting in Greece. J. Public Health Res. 9(3), 517-522 (2020). https://doi.org/10.4081/jphr.2020.1962
10. Leufvén, M., Vitrakoti, R., Bergström, A., KC, A., Målqvist, M.: Dimensions of Learning Organizations Questionnaire (DLOQ) in a low-resource health care setting in Nepal. Health Res. Policy Syst. 13, 6 (2015). https://doi.org/10.1186/1478-4505-13-6
11. Nguyen-Duc, T., Phuong, T.T., Le, T.T.B., Nguyen, L.T.T.: Validation of dimensions of the learning organization questionnaire (DLOQ) in a Vietnamese context. Learn. Org. 30(5), 511–531 (2023). https://doi.org/10.1108/TLO-04-2022-0041
12. Cummings, T.G., Worley, C.G.: Organization Development and Change. Cengage, Boston (2019)
13. Snyder, W.M.: Organizational learning and performance: an exploration of the linkages between organizational learning, knowledge, and performance. Doctoral Dissertation, University of Southern California (1996)
14. Chughtai, M.S., Syed, F., Naseer, S., Chinchilla, N.: Role of adaptive leadership in learning organizations to boost organizational innovations with change self-efficacy. Curr. Psychol. 43, 27262–27281 (2024). https://doi.org/10.1007/s12144-023-04669-z
15. Ju, B., Lee, Y., Park, S., Yoon, S.W.: A meta-analytic review of the relationship between learning organization and organizational performance and employee attitudes: using the Dimensions of Learning Organization Questionnaire. Hum. Resour. Dev. Rev. 20(2), 207–251 (2021). https://doi.org/10.1177/1534484320987363
16. Kumar, J. K., Gupta, R., Basavaraj, P., Singla, A., Prasad, M., Pandita, V., et al.: An insight into health care setup in National Capital Region of India using dimensions of learning organizations questionnaire (DLOQ) - A cross-sectional study. J. Clin. Diagn. Res., 10(6), ZC01 (2016). https://doi.org/10.7860/JCDR/2016/16186.7898
17. Kumar, J. K., Patthi, B., Singla, A., Gupta, R., Prasad, M., Pandita, V.: Application of dimensions of learning organization questionnaire in a dental institution in National Capital Region of India: A cross-sectional study. J. Indian Assoc. Public Health Dent., 14(2), 175-181 (2016). https://doi.org/10.4103/2319-5932.181898
18. Watkins, K. E., Milton, J., Kurz, D.: Diagnosing the learning culture in public health agencies. Int. J. Cont. Educ. Lifelong Learn., 2(1), 65-81 (2009). https://hdl.voced.edu.au/10707/51261
19. Bostwick, A., Ades, A., Rodriguez-Paras, C., Dombroski, M., Lim, C., Ordoñez Paredes, M., et al.: A survey of team culture and learning organization in the resuscitation of neonates with congenital anomalies: A single center experience. Resusc. Plus, 22, 100877 (2025). https://doi.org/10.1016/j.resplu.2025.100877
20. Singer, S. J., Moore, S. C., Meterko, M., Williams, S.: Development of a short-form learning organization survey: The LOS-27. Med. Care Res. Rev., 69(4), 432-459 (2012). https://doi.org/10.1177/1077558712448135
21. Ford, J., Isaacks, D. B., Anderson, T.: Creating, executing and sustaining a high-reliability organization in health care. Learn. Org., 31(6), 817–833 (2024). https://doi.org/10.1108/TLO-03-2023-0048

22. Richardson, E.: Health systems in action 2024: Serbia. Technical report, European Observatory on Health Systems and Policies (2024). https://eurohealthobservatory.who.int/publications/i/health-systems-in-action-serbia-2024
23. Kim, M.K., Kim, S.M., Bilir, M.K.: Investigation of the dimensions of workplace learning environments (WLEs): development of the WLE Measure. Perf. Improv. Q. 27, 35–57 (2014). https://doi.org/10.1002/piq.21170
24. Kyriakeli, G., Georgiadou, A., Lithoxopoulou, M., Tsimtsiou, Z., Kotsis, V.: The impact of ISO certification procedures on patient safety culture in public hospital departments. Healthcare. 13(6), 661 (2025). https://doi.org/10.3390/healthcare13060661
25. Slakey, D.P., George, J.S., Anderson, E., Willeumier, D., Guglielmi, K.: Applying international organization for standards 9001 to create an effective surgical quality committee. Am. J. Surg. 221, 598–601 (2021). https://doi.org/10.1016/j.amjsurg.2020.11.014
26. Crossan, M.M., Lane, H.W., White, R.E.: An organizational learning framework: from intuition to institution. Acad. Manag. Rev. 24(3), 522-537 (1999). https://doi.org/10.2307/259140
27. March, J.G.: Exploration and exploitation in organizational learning. Org. Sci. 2, 71–87 (1991). https://doi.org/10.1287/ORSC.2.1.71
28. Lawrence, T.B., Mauws, M.K., Dyck, B., Kleysen, R.F.: The politics of organizational learning: integrating power into the 4I framework. Acad. Manag. Rev. 30(1), 180–191 (2005). https://doi.org/10.5465/amr.2005.15281451
29. Wodnik, B.K., Andiappan, M., Di Ruggiero, E., Lavery, J.V.: The 6I model: an expanded 4I framework to conceptualise interorganisational learning in the global health sector. BMJ Open. 14(5), 1-8 (2024). https://doi.org/10.1136/bmjopen-2023-083830
30. Wiewiora, A., Chang, A., Smidt, M.: Individual, project and organizational learning flows within a global project-based organization: exploring what, how and who. Int. J. Proj. Manag. 38, 201-214 (2020). https://doi.org/10.1016/j.ijproman.2020.03.005
31. Merriam, S.B., Tisdell, E.J.: Qualitative Research: A Guide to Design and Implementation. 4th edn. Jossey-Bass, San Francisco (2016)
32. Wasti, S.P., Simkhada, P., Van Teijlingen, E., Sathian, B., Banerjee, I.: The growing importance of mixed-methods research in health. Nepal J. Epidemiol. 12(1), 1175–1178 (2022). https://doi.org/10.3126/nje.v12i1.43633
33. Merriam, S.B.: Qualitative Research: A Guide to Design and Implementation. Jossey-Bass, San Francisco (2009)
34. Sasaki, M., Hirota, C., Nakazawa, K., Yonekura, Y., Ogata, Y.: Relationship between empowering leadership, resilience and organisational learning from incidents: a cross-sectional study. J. Adv. Nurs. 1-17 (2025). https://doi.org/10.1111/jan.70364
35. Shabalala, Z.A., Chitamba, A.: From leadership to well-being: the mediating role of employee resilience: an integrative review. Int. J. Res. Bus. Sch. 14(9), 158–165 (2026). https://doi.org/10.20525/ijrbs.v14i9.4629
36. Moschovopoulou, A., Papavassiliou-Alexiou, I.: The effect of perceived organizational support, empowering leadership and learning organization on Greek psychiatric nurses' employee resilience: a cross-sectional study. Int. J. Workplace Health Manag. 18(4), 371–390 (2025). https://doi.org/10.1108/IJWHM-04-2024-0088

Original Research Article *https://doi.org/10.5281/zenodo.23476660*

Unravelling the Employee Resilience concept: The supremacy of effective learning at work

Vuk Mirčetić[1[0000-0001-6675-5025], Ivana Kovačević[2a,b[0000-0002-7601-1000], and Marko Slavković[3[0000-0002-2604-1228]

[1,3] University of Kragujevac, Faculty of Economics, Kragujevac, [2a] Serbia University of Belgrade, Faculty of Organizational Sciences, Belgrade, Serbia; [b]East-European Center for Research in Economics and Business,West University of Timisoara, Faculty of Economics and Business Administration, Romania

Abstract. Employee resilience has become an increasingly important construct in contemporary work and organisational research, yet questions remain regarding its behavioural structure and practical measurement in specific empirical contexts. The present study examines the internal structure and reliability of a nine-item measure of employee resilience, understood as adaptive functioning in the work context. There were 242 respondents sampled from a number of different demographic and occupational types. The results of a principal component analysis determined that all items converged on a single underlying dimension, confirmed through parallel analysis, internal consistency, and reliability in an employed subsample. The analysis produced one component, which explained 58.39% of the variance of the total and had strong factor loadings (ranging from .713–.845) and a very high internal reliability (Cronbach's $\alpha = .910$). Of all the items, the greatest degree of loading was associated with the statement pertaining to learning from one's mistakes and improving one's work, indicating that, in this sample of employees, reflective learning may be an important contributor to one's level of adaptive resilience. The one-component structure was found to be stable across the employed subsample, indicating that the construct is a reliable and valid measure concerning resilience in a workplace setting. Research showed no statistically significant differences in total resilience scores across the majority of respondent characteristics; only gender and work modality demonstrated negligible and nearly statistically significant effects. The findings of the study support characterising employee resilience as a coherent behaviour that integrates collaborative behaviours, workload management, crisis management, learning behaviour, feedback-seeking behaviour, relationships with others who provide assistance, and orientation toward constructive change.

Keywords: Employee Resilience, Principal Component Analysis, Adaptive Work Behaviour, Workplace Learning, Psychological Safety.

1 Introduction

The contemporary business world is more challenging than ever. Organisations today operate in business environments characterised by persistent uncertainty, accelerated change, intensified performance demands, and growing expectations of continuous adaptation. Faced with the pervasive crisis, employers are turning to various solutions, ranging from excessive resource exploitation to more sustainable strategies. Recent work has also shown that contemporary human resource management (HRM) practice is increasingly shaped by complex and interdependent environmental pressures, requiring organisations to respond with more adaptive and strategically grounded people-management approaches [1]. From this point, the concept of resilient organisation emerges, describing an adaptive workplace effective in overcoming crisis by going beyond survival in times of instability and uncertainty but also capitalising and turning it to its favour [2]. In that setting, new employment deals are imposed on the workforce, expecting them to cope with the situation effectively and beneficially for the organisation, while preserving their own welfare. This capacity relies on the resilience of the workforce and their capability to benefit from organisational support [3].

Mental health studies usually refer to resilience as the main individual resource to deal with adverse influences, implying that it is a trait, capacity or process [4] that enables a person to endure and sustain when circumstances are not favourable. Recently, scholars have acknowledged that resilience is context-specific and contingent, varying not only between individuals but also within the same person, depending on the sphere of life, a finding that aligns with an ecologically valid approach to the phenomenon [5]. The workplace is the main arena for expressing resilience, enabling employees not only to deal with critical issues but to outgrow the situation and negative influences.

Over the past few decades, there have been endeavours to develop tools for identifying, assessing and developing resilient potential among employees. Näswall and colleagues [6] operationalised and validated their idea of an Employee resilience measure with a 9-item scale that portrays it as continuous proactive adaptation to everyday workplace situations.

Nevertheless, the content of the notion is still novel and might vary among researchers. Our idea is to examine the internal structure of the dimension and to identify the most prominent aspects of the resilience that occur in the specific context of a contemporary business setting. In the context of imposing on organisations to reconfigure the HRM function to be more sustainable and employee-supportive [7], it is valuable to recognise the main levers and manifestations of resilient behaviours.

1.1 Theoretical considerations

In general, there are three perspectives of individual resilience: as a relatively stable personality trait, as the capacity that can be developed, and as the process that emerges as the effect of positive adaptation [8]. Although in the past resilience was often defined as a personal attribute closely related to the character trait of hardiness [9], scholars are increasingly turning to more manageable conceptions. Acknowledging the circumstantial effects and significance of particular setting features, they started to study resilience in domain-specific terms [4].

Due to changes in the business, labour market, and professional landscape, work has become the main area of life's challenging events, imposing the necessity of dealing with emerging issues. As resilience makes employees more open to changes and highly productive despite the uncertainty [5], the concept comes into focus for researchers and practitioners.

Kuntz [10] sees workplace resilience as a multifaceted, multilevel concept that involves using and developing available resources to cope with and flourish in challenging work situations. So, it seems that this relation between organisational and employee resilience is reciprocal. Organisational assets enable employees to develop their resilience potential, while enriching organisational capacity to provide employees with an adequate setting for growing [11]. Näswall and colleagues [12] identify at least three organisational resilience leverages: supportive leadership, learning culture and supportive team and organisational environment.

Malik and Garg [13] describe resilient employees as those who, in adverse situations, sustain and express adaptive capacity, as well as the ability to use resources at their disposal to accomplish personal, as well as organisational goals. It is expected from the resilient employee to have the capacity to actively use organisational resources for facing challenges rather than to wait passively for aid [3]. Further, there is evidence that resilient employees are more engaged and committed, offering higher work performance [6], [14], [15].

That is the reason, among others, that studies are covering the HRM practices that promote employee resilience and complement each other [16], [17]. The main inspiration of the mentioned studies is to identify HRM enablers to support the emergence of resilient employees that would foster the work environment's resilience reciprocally. These HRM conceptions are usually defined as sustainable HRM, designed around opportunities for employees' development and empowerment, and that offers a healthy and safe work environment [18].

There are at least two theoretical foundations of the concept [3]. One derives from positive psychology and its support of well-being by regulating emotions, fostering growth despite challenges, and learning from mistakes. The other root is in the more recent conservation resource theory. Scholars often postulate this individual variable within the JD-R model [19] as the representative of the personal resource, or so-called psychological aspect, opposite to job demands and in line with organisational resources that support and sustain resilience and employee well-being [15]. Sustainable HRM practices might be seen as one of those organisational resources [17].

1.2 Research problem

The study's framework defines employee resilience as a proactive, manageable capability that can be developed and shaped in a suitable, fertile environment and with available resources [20]. These capabilities are evident in employees' resilient behaviours, including proactiveness, learning capacities, and networking strategies, which are bolstered by organisational assets [6]. This ecological perspective [5] that respects the specific context of resilience expression is operationalised in more than one instrument.

According to Näswall and coauthors [5], researchers predominantly use Psychological Capital (PsyCap; [21]) and the Workplace Resilience Inventory (WRI; [22]). Nevertheless, it is found to be less work setting-focused. Näswall and colleagues operationally defined employee resilience in a behavioural rather than a dispositional manner (as it had previously been practised). There is explicit consideration of organisational resources that nurture and contribute to (possible) development of personal resilience. As well as the existence and occurrence of the resilient behaviours independently of potential stressors and pressure.

The mentioned instrument is becoming increasingly popular in research, demonstrating strong metric properties. So, the main goal of our study is to examine the structural content of one emerging measurement model of the employee resilience concept defined in the previously described manner. We are using principal component analysis to corroborate the construct's one-dimensional nature and identify the most prominent feature of resilience in the workplace.

2 Methods

2.1 Sample

The empirical research included 242 respondents. The sample was evenly distributed by gender: 121 respondents were male, and 121 were female, with each group accounting for 50.0% of the sample. The largest age group was respondents aged 25 or younger (n = 78; 32.2%). This was followed by respondents aged 31 to 40 years (n = 57; 23.6%) and 41 to 50 years (n = 48; 19.8%). In terms of education, the largest share of respondents held a bachelor's degree (n = 82; 33.9%), followed by secondary school education (n = 70; 28.9%) and a master's degree (n = 49; 20.2%). Regarding employment status, most respondents worked in domestic private companies (n = 111; 45.9%). A further 61 respondents worked in foreign-owned private companies (25.2%), 35 in the public sector (14.5%), and 35 reported that they were not employed (14.5%). With respect to organisation size, the largest group of respondents came from organisations with more than 250 employees (n = 83; 34.3%). 55 respondents worked in organisations with 1-9 employees (22.7%), 37 with 50-249 employees (15.3%), and 34 with 10-49 employees (14.0%). On-site work was the most frequent work modality (n = 118; 48.8%), followed by hybrid work (n = 68; 28.1%) and work

from home (n = 22; 9.1%). The sample characteristics are presented systematically in Table 1.

Table 1. Sample characteristics

Gender	Male	121	50.0
	Female	121	50.0
Age	Under 25 years	78	32.2
	25 to 30 years	32	13.2
	31 to 40 years	57	23.6
	41 to 50 years	48	19.8
	Over 50 years	27	11.2
Education	Secondary school	70	28.9
	Higher/vocational school	34	14.0
	Bachelor's degree	82	33.9
	Master's degree	49	20.2
	Doctoral degree	7	2.9
Employment status	Domestic private company	111	45.9
	Foreign-owned private company	61	25.2
	Public sector	35	14.5
	Not employed	35	14.5
Organization size	1-9 employees	55	22.7
	10-49 employees	34	14.0
	50-249 employees	37	15.3
	More than 250 employees	83	34.3
	In career transition	33	13.6
Work modality	On-site work	118	48.8
	Hybrid work	68	28.1
	Work from home	22	9.1
	In career transition	34	14.0

2.2 Instrument

The study used a nine-item scale designed to assess individual work resilience, understood here as adaptive functioning in the work context. The scale is constructed and validated by Näswall and colleagues and called the Employee Resilience Scale (ER). The items cover collaboration, workload management, crisis resolution, learning from experience, re-evaluation of one's own performance, responsiveness to feedback, help-seeking, managerial support, and the interpretation of change as an opportunity for growth. Responses were given on a five-point Likert-type scale; higher scores indicate stronger agreement with the statement and, consequently, a higher level of the construct. The internal reliability is $\alpha(9)=.91$.

2.3 Statistical Analysis

Before conducting principal component analysis, the data were screened for missing values, item distributions, descriptive statistics, and the suitability of the correlation matrix for dimensionality reduction. No missing values were found in the nine analysed items.

Principal component analysis was conducted on the correlation matrix of the nine items. Categorical variables, including gender, age, education, employment status, organisation size, and work modality, were not entered as active variables in the PCA. They were used to describe the sample and to conduct additional comparisons of the total score.

The suitability of the data for PCA was assessed using the Kaiser-Meyer-Olkin measure of sampling adequacy, Bartlett's test of sphericity, inspection of the correlation matrix, and the ratio of respondents to variables. The number of components was evaluated using Kaiser's criterion, the percentage of explained variance, the scree criterion, and parallel analysis based on 5,000 randomly generated datasets. Because only one component was retained, no rotation was applied. Internal consistency was evaluated using Cronbach's alpha, corrected item-total correlations, and alpha values if an item was deleted.

3 Results

As it is stated in Table 2, item means were high overall, ranging from M = 3.94 to M = 4.39. The highest mean was observed for the item "I learn from mistakes and improve the way I do my job", whereas the lowest mean was obtained for "I use change at work as an opportunity for growth". Even the lowest mean, however, was above the theoretical midpoint of the scale.

The item distributions were negatively skewed, indicating that responses tended to cluster toward the upper end of the scale. This pattern is not unusual for self-reports of positively valued work behaviours, but it should be kept in mind when interpreting the high average scores.

Table 2. Descriptive statistics of the items

Item	Statement	M	SD	Median	Min-max
S1	I effectively collaborate with others to handle challenges at work	4.26	1.01	5	1-5
S2	I successfully manage a high workload for long periods of time	4.10	0.99	4	1-5
S3	I resolve crises competently at work	4.12	1.00	4	1-5
S4	I learn from mistakes and improve the way I do my job	4.39	0.94	5	1-5
S5	I re-evaluate my performance and continually improve the way I do my work	4.21	0.97	4	1-5
S6	I effectively respond to feedback, even	4.16	0.97	4	1-5

	criticism				
S7	I seek assistance at work when I need specific resources	4.21	1.03	5	1-5
S8	I approach managers when I need their support	4.15	1.07	5	1-5
S9	I use change at work as an opportunity for growth	3.94	1.04	4	1-5

The items were positively intercorrelated. Inter-item correlations ranged from r = .401 to r = .700, with an average inter-item correlation of approximately r = .531. This pattern indicates sufficient shared variance among the items without suggesting that the items are merely redundant.

The Kaiser-Meyer-Olkin measure of sampling adequacy was KMO = .900, which indicates excellent suitability for PCA. Individual MSA values ranged from .871 to .934. Bartlett's test of sphericity was statistically significant, $\chi^2(36) = 1228.10$, $p < .001$, confirming that the correlation matrix differed significantly from an identity matrix. The respondent-to-item ratio was 242:9, or 26.9 respondents per item.

The PCA identified one dominant component. The first component had an eigenvalue of 5.255 and explained 58.39% of the total variance. The second component had an eigenvalue of .901 and therefore fell below Kaiser's criterion. All remaining components also had eigenvalues below 1. The eigenvalues and percentages of explained variances of components are given in Table 3.

Table 3. Eigenvalues and explained variance

Component	Eigenvalue	% of variance	Cumulative %
1	5.255	58.39	58.39
2	0.901	10.01	68.41
3	0.728	8.09	76.50
4	0.449	4.99	81.49
5	0.421	4.68	86.16
6	0.395	4.39	90.55
7	0.339	3.77	94.32
8	0.283	3.14	97.46
9	0.228	2.54	100.00

Parallel analysis supported the one-component solution. Only the first empirical eigenvalue exceeded the 95th percentile of the corresponding random eigenvalue; the second and subsequent empirical eigenvalues were lower than their random-data criteria (See Table 4).

Table 4. Parallel analysis

Component	Empirical eigenvalue	95th percentile random eigenvalue
1	5.255	1.399
2	0.901	1.265
3	0.728	1.175
4	0.449	1.101

As it is presented in Table 5, all loadings on the first component were positive and high, ranging from .713 to .845. The strongest loading was observed for the item "I learn from mistakes and improve the way I do my job". The lowest, although still clearly satisfactory, was observed for "I approach managers when I need their support".

Table 5. Component loadings and communalities

Item	Statement	Loading	Communality
S1	I effectively collaborate with others to handle challenges at work	0.771	0.595
S2	I successfully manage a high workload for long periods of time	0.761	0.580
S3	I resolve crises competently at work	0.750	0.563
S4	I learn from mistakes and improve the way I do my job	0.845	0.714
S5	I re-evaluate my performance and continually improve the way I do my work	0.795	0.632
S6	I effectively respond to feedback, even criticism	0.768	0.589
S7	I seek assistance at work when I need specific resources	0.741	0.549
S8	I approach managers when I need their support	0.713	0.508
S9	I use change at work as an opportunity for growth	0.725	0.526

Communalities ranged from .508 to .714, meaning that the one-component solution accounted for roughly 51% to 71% of the variance in individual items. No item showed a weak loading or low communality that would justify its removal from the scale.

Internal consistency was high, Cronbach's alpha = .910. Corrected item-total correlations ranged from .636 to .787. Alpha values if an item was deleted ranged from .893 to .903, indicating that removing any single item would not improve the reliability of the scale.

Based on the PCA and reliability analysis, a total score was computed as the mean of the nine items. The mean total score was M = 4.17, SD = 0.76, 95% CI [4.07, 4.27]. This indicates a relatively high level of self-reported individual work resilience in the sample.

Because a subset of respondents reported not being currently employed, the structure of the scale was also examined in the employed subsample (n = 207). The results remained stable: KMO = .890, Bartlett's test $\chi^2(36) = 953.65$, $p < .001$, first-component eigenvalue = 5.046, explained variance = 56.07%, loading range = .684-.844, and Cronbach's alpha = .900. These results support the stability of the one-component structure in the work context.

Additional analyses, presented in Table 6, examined whether the total score differed across basic respondent characteristics. In most cases, no statistically significant differences were found. The gender difference was borderline, with female respondents showing a slightly higher mean score. For work modality, the standard ANOVA was significant, but Welch's ANOVA, which is more appropriate when group sizes

and variances differ, was borderline; this result should therefore be interpreted cautiously.

Table 5. Additional comparisons of the total work resilience score

Variable	Finding	Test	Effect
Gender	Male M = 4.07 (SD = 0.79) Female M = 4.27 (SD = 0.73)	t(238.71) = -1.95, p = .052	d = 0.25
Age	No significant difference across age groups	F(4, 237) = 1.09, p = .363	η^2 = .018
Education	No significant difference by level of education	F(4, 237) = 0.78, p = .537	η^2 = .013
Employment status	No significant difference by employment status	F(3, 238) = 2.05, p = .108	η^2 = .025
Organisation size	No significant difference by organisation size	F(4, 237) = 1.65, p = .163	η^2 = .027
Work modality	The standard ANOVA was significant, but Welch's ANOVA was borderline	F(3, 238) = 4.18, p = .007; Welch F(3, 64.11) = 2.66, p = .056	η^2 = .050

4 Discussion and conclusion

4.1 Discussion

The results provide clear empirical support for the one-component structure of the scale. The nine items converge around a single dominant component, suggesting that collaboration, workload management, crisis resolution, learning from mistakes, responsiveness to feedback, help-seeking, managerial support, and a constructive view of change are not isolated behaviours in this dataset. Rather, they appear to represent connected aspects of a broader pattern of adaptive functioning at work.

This pattern is best interpreted as individual work resilience. The construct should not be reduced to simply tolerating pressure or "bouncing back" after difficulties. In the present item set, resilience also involves learning, reflection, adjustment, and the ability to mobilise social and organisational resources. The strongest loading was on the item about learning from mistakes, suggesting that reflective learning is central to the measured dimension. As Malik and Garg [23] corroborate in their study, learning culture and knowledge sharing evoke positive emotional states among employees and nurture their resilience potential. There is also strong evidence that a positive learning culture in an organisation fosters employee resilience [24], implying that these settings are ideal for displaying the learning aspect of resilient behaviour. As a matter of fact, Näswall and colleagues [25] see resilience as a transformative process that leads to conquering new territories of personal and professional experience.

Although social support in general was found to be in positive correlation with resilience [26], [27], the item "I approach managers when I need their support" had the lowest loading, but the loading was still satisfactory. This may reflect the fact that seeking support from managers is partly shaped by organisational culture, managerial

availability, and the quality of supervisory relationships. For example, Sommer and colleagues [28] found that transformational leadership matters when it comes to the resilience of subordinates. Also, studies of career resilience did not find the link with supervisor support [29], so it probably also depends on the perspective. Still, the item contributes both statistically and conceptually to the breadth of the scale.

The high reliability coefficient indicates that the items measure the same underlying construct consistently. At the same time, the content of the items is not overly narrow: the scale captures individual and social elements of adaptive functioning, ranging from self-reflection and feedback use to seeking help and support. The sensitivity analysis in the employed subsample further supports the conclusion that the one-component structure is not an artefact of the full sample composition.

The high average scores suggest that respondents generally see themselves as capable of handling work demands, challenges, and change. However, because the response distributions were skewed toward the upper end of the scale, this finding should be interpreted with some caution. The pattern may reflect genuinely high levels of adaptive work behaviour, but it may also partly reflect socially desirable responding, which is a common issue in self-report measures of positive workplace characteristics.

Although some research suggests that some demographic variables might be relevant, in our sample, there are no indications of these effects. We did not find differences between age groups, contrary to studies that follow resilience over the career lifespan, which implies that older workers might have better-developed resilience mechanisms compared with younger employees and those at the beginning of their careers. Nevertheless, these studies are conducted in different contexts and with a wider age spread among respondents [30].

4.2 Practical implications

The findings have several practical implications for organisations and human resource professionals. First, the scale can be used as a brief and reliable indicator of individual work resilience. The total score is easy to compute and can be useful in internal employee surveys, evaluations of development programs, and assessments of where additional support may be needed.

Second, the central role of learning from mistakes suggests that resilience development should not be limited to stress-tolerance or endurance training. Organisations may benefit more from practices that encourage reflection, constructive learning from errors, and open discussion of work experience. As a matter of fact, learning from failure is at the core of resilience, as it increases the sense of responsibility [31], followed by introspective phases in order to grow stronger from disruption [25]. These insights are close to sense-making behaviours that enable a person to proactively face challenges [32] and to reframe work stressors [10].

Third, the items related to feedback and performance re-evaluation point to the importance of regular, constructive, development-oriented feedback. Performance systems should not only evaluate outcomes but also provide employees with clear guidance on how to improve their work. It aligns with research showing that feedback is

positively related to resilience, especially among promotion-focused and prevention-focused employees [27]. Seeing and responding effectively to feedback while constantly monitoring one's own performance for improvement are in the very core of resilience [6], as well as acquiring personal accountability for one's own development [15].

Fourth, the results highlight the social side of resilience. Seeking assistance, resources, and support should not be interpreted as weakness, but as part of adaptive functioning. Organisations should therefore invest in psychological safety, managerial accessibility, and clear channels for support. Resilience flourishes in the positive social context [33] that provides social support [26]. Although strong interpersonal relations are the basis of employee resilience in challenging times, they depend on trust,, and sometimes it is not easy to approach managers when employees lack sufficient autonomy and feel insufficiently empowered in the context [20].

Finally, the lowest item mean was observed for viewing change at work as an opportunity for growth. This suggests that change management remains a particularly important area for intervention. Employees are more likely to respond constructively to change when they understand its purpose, have some opportunity to participate in the process, and can connect the change to their own professional development. As studies often emphasise, the sense of control, as well as the understanding and reframing of the meaning of the change [10] are crucial for a growth mindset while experiencing crisis [34].

4.3 Limitations of the study

There are several limitations to consider when making inferences from these findings. One is that this study employed a convenience sample, which limits external validity. While the sample size was adequate for dimensionality analysis as well as being sufficiently heterogeneous (demographically and in terms of work-related exposures), strong generalisations beyond the sample of the study cannot be made. Therefore, the observed psychometric pattern should be viewed as bounded evidence (specific to our context only), not as definitive proof that the scale will operate in the same way across all occupations, sectors or nations.

The second limitation was that this study relied solely on self-report data, which is important because the behaviours being analysed were all desirable in the workplace, including: collaborating, learning from mistakes, being responsive to feedback, seeking help, and engaging constructively with change.

The third limitation of the present study relates to the cross-sectional design. Although the analyses support the internal consistency and reliability of the measure in the current sample, they do not allow for conclusions regarding temporal stability, developmental pathways or causal relationships. This is particularly relevant when considering employee resilience, which is often described as a dynamic, context-dependent ability rather than a fixed trait. Therefore, on the one hand, this research presents results which provide evidence of structural coherence at a single point in time, but the data do not provide evidence regarding how resilience becomes activated, maintained or increased over time during an employee’s work experience.

Fourth, the scale is originally constructed in English and is translated into Serbian, which might undermine the strength and validity of the conclusions. And finally, this is one of the first applications of the scale on the population at the territory of Serbia (the first one was conducted on 45 employees in 2025 [35]).

The inclusion of respondents in career transitions adds another layer of conceptual variation, as noted in the previous examination. It would benefit future studies testing the scale on samples of workers in the labour force on an active basis, by occupational group, sector of the economy, and different workforce arrangements.

Although the mentioned limitations do not invalidate the findings from this study, they do limit how the results of the research can be interpreted. The results support that the nine-item measure, as tested in the research, has been shown to be valid and reliable, however, claims of the measure's construct validity, contextual generalisability and overarching psychometric merit must be established through further testing. Future studies should build upon this study by using confirmatory factor analysis, longitudinal designs, subgroup invariance testing and validation against theoretically relevant organisational constructs and outcomes.

4.4 Conclusion

The results of the study showed statistically significant evidence supporting the comprehensive, reliable use of a single hierarchical factorial tool to measure employee resilience, and all items loaded positively and significantly onto the extracted component, which explained a large percentage of total variance. These results collectively provide additional evidence for the scale's acceptability as a valid behavioural indicator of an employee's adaptive capacity at work. Consequently, the used tool can be seen as providing a reasonable estimate of employee resilience, with higher scores representing stronger collaboration, workload management strategies, crisis resolution abilities, responsiveness to feedback, support-seeking behaviour, and constructive coping through adapting to change within the work environment.

The study provides a solid psychometric framework, but it also makes a substantial contribution to understanding the nature of resilience among employees. The relative importance of items referring to learning from one's errors and to improvement in performance suggests that we should not view resilience in an organisational context solely as a type of endurance, recovery, or resistance to stress. Instead, the study's findings suggest a view of resilience that is more developmental in nature, linking together reflective learning, self-correction, and the ability to use challenging experiences as a basis for improvement. The principal component analysis confirms that the nine items form a clear, stable, and reliable one-component scale. Therefore, the findings of the study conclude that resilient functioning in the workplace appears to involve not only performing well under stress, but also learning from the challenges faced and modifying one's behaviour to sustain future performance.

From a theoretical perspective, the mentioned assertion reinforces the notion that employee resilience should be viewed as an active, integrated capacity rather than a reactive, passive trait that provides protection. As constructed here, employee resilience appears to represent the person's self-regulatory, relational and developmental

adaptation in one behavioural construct. Within this configuration, learning has an important role in establishing a link between difficulty and growth, and establishing resilience as a developmental process, as opposed to pure survival. Seen from this viewpoint, resilience is demonstrated by how employees cope with the demands of work and also by how they interpret setbacks, utilise feedback, seek support, and evolve their behaviour over time.

The present research contributes to the existing literature by providing empirical evidence that a brief (9-item) measure can capture employee resilience as a valid construct in the current context. This study also suggests that one of the most meaningful expressions of resilience may be through learning. Therefore, if organisations are seeking to improve the resilience of their employees, interventions that focus solely on coping or stress tolerance may be overly limited. Instead, promoting psychologically safe, feedback-oriented and learning-centric workplaces may provide employees with opportunities not just to endure challenge, but also to grow through the challenge.

Acknowledgments.

AI Disclaimer: Artificial intelligence tools were used solely for language editing and proofreading purposes. No AI-generated content, analysis, or interpretations were included in the manuscript.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Mirčetić, V., Popović, G., Vukotić, S., Mihić, M., Kovačević, I., Đoković, A., Slavković, M.: Navigating the Complexity of HRM Practice: A Multiple-Criteria Decision-Making Framework. Mathematics 12(23), 3769 (2024)
2. Youssef, C.M., Luthans, F.: Positive organizational behavior in the workplace: The impact of hope, optimism, and resilience. Journal of Management **33**(5), 774–800 (2007)
3. Kuntz, J.R., Malinen, S., Näswall, K.: Employee resilience: Directions for resilience development. Consulting Psychology Journal: Practice and Research **69**(3), 223–242 (2017)
4. Hartmann, N. B., Hüffer, T., Thompson, R. C., Hassellöv, M., Verschoor, A., Daugaard, A. E., Rist, S., Karlsson, T., Brennholt, N., Cole, M., Herrling, M. P., Hess, M. C., Ivleva, N. P., Lusher, A. L., Wagner, M.: Are we speaking the same language? Recommendations for a definition and categorization framework for plastic debris. Environmental science & technology, **53**(3), 1039–1047 (2019)
5. Tonkin, K., Malinen, S., Näswall, K., Kuntz, J.C.: Building employee resilience through wellbeing in organizations. Human Resource Development Quarterly **29**(2), 107–124 (2018)
6. Näswall, K., Malinen, S., Kuntz, J., Hodliffe, M.: Employee resilience: Development and validation of a measure. Journal of Managerial Psychology **34**(5), 353–367 (2019)
7. Lu, Y., Zhang, M.M., Yang, M.M., Wang, Y.: Sustainable human resource management practices, employee resilience, and employee outcomes: Toward common good values. Human Resource Management **62**(3), 331–353 (2023)

8. Kossek, E.E., Perrigino, M.B.: Resilience: A review using a grounded integrated occupational approach. The Academy of Management Annals **10**(1), 729–797 (2016)
9. Fisher, D.M., Ragsdale, J.M., Fisher, E.C.: The importance of definitional and temporal issues in the study of resilience. Applied Psychology **68**(4), 583-620 (2018)
10. Kuntz, J.C.: Resilience in times of global pandemic: Steering recovery and thriving trajectories. Applied Psychology **70**(1), 188–215 (2021)
11. Shin, J., Taylor, M.S., Seo, M.G.: Resources for change: The relationships of organizational inducements and psychological resilience to employees' attitudes and behaviors toward organizational change. Academy of Management Journal **55**(3), 727–748 (2012)
12. Näswall, K., Kuntz, J., Hodliffe, M., Malinen, S.: Employee resilience scale (EmpRes) measurement properties. Resilient Organizations Research Programme, Christchurch, New Zealand, pp. 1–4 (2015)
13. Malik, P., Garg, P.: Learning organization and work engagement: The mediating role of employee resilience. The International Journal of Human Resource Management **31**(8), 1071–1094 (2020)
14. Plimmer, G., Berman, E.M., Malinen, S., Franken, E., Näswall, K., Kuntz, J., Löfgren, K.: Resilience in public sector managers. Review of Public Personnel Administration **42**(2), 338–367 (2022)
15. Plimmer, G., Kuntz, J., Berman, E., Malinen, S., Näswall, K., Franken, E.: The negative relationships between employee resilience and ambiguity, complexity, and inter-agency collaboration. Australian Journal of Public Administration **82**(2), 248–270 (2023)
16. Shao, D., Ji, Y., Zhou, R., Liu, T.: Unpacking the mixed influences of sustainable human resource management on employee resilience: Effects of challenge–hindrance appraisal and trait mindfulness. Asia Pacific Journal of Human Resources **63**(1), e12433 (2025)
17. Trunk Širca, N., Riaz, Z., Hamid, S., Žurauskė, G., Stankevičiutė, Ž.: The interplay among employee resilience, sustainable HRM and work engagement. Human Systems Management **43**(4), 573–587 (2024)
18. Stankevičiūtė, Ž., Staniškienė, E., Daunorienė, A., Ramanauskaitė, J.: Hindrance Job Demands as Factors Undermining Employee Resilience. Sustainability, 18(6), 2692. (2026).
19. Bakker, A.B., Demerouti, E., Sanz-Vergel, A.: Job demands–resources theory: Ten years later. Annual Review of Organizational Psychology and Organizational Behavior **10**(1), 25–53 (2023)
20. Malik, G., Singh, P.: Fostering social sustainability: Unveiling HR's power in enhancing employee resilience via social exchange and broaden-and-build theories. Employee Relations: The International Journal **46**(3), 675–701 (2024)
21. Luthans, F., Youssef, C.M., Sweetman, D.S., Harms, P.D.: Meeting the leadership challenge of employee well-being through relationship PsyCap and health PsyCap. Journal of Leadership & Organizational Studies **20**(1), 118–133 (2013)
22. McLarnon, M.J., Rothstein, M.G.: Development and initial validation of the workplace resilience inventory. Journal of Personnel Psychology **12**(2), 63–73 (2013)
23. Malik, P., Garg, P.: The relationship between learning culture, inquiry and dialogue, knowledge sharing structure and affective commitment to change. Journal of Organizational Change Management **30**(4), 610–631 (2017)
24. Caniëls, M.C., Baaten, S.M.: How a learning-oriented organizational climate is linked to different proactive behaviors: The role of employee resilience. Social Indicators Research **143**(2), 561–577 (2019)
25. Näswall, K., Kuntz, J., Hodliffe, M., Malinen, S.: Employee resilience scale (EmpRes): Technical report (Resilient Organisations Research Report No. 2013/06). Resilient Organisations Research Programme, Christchurch, New Zealand (2013)

26. Todt, G., Weiss, M., Hoegl, M.: Mitigating negative side effects of innovation project terminations: The role of resilience and social support. Journal of Product Innovation Management **35**(4), 518–542 (2018)
27. Kuntz, J., Connell, P., Näswall, K.: Workplace resources and employee resilience: The role of regulatory profiles. Career Development International **22**(4), 419–435 (2017)
28. Sommer, S.A., Howell, J.M., Hadley, C.N.: Keeping positive and building strength: The role of affect and team leadership in developing resilience during an organizational crisis. Group & Organization Management **41**(2), 172–202 (2016)
29. Kidd, J.M., Smewing, C.: The role of the supervisor in career and organizational commitment. European Journal of Work and Organizational Psychology **10**(1), 25–40 (2001)
30. Scheibe, S., De Bloom, J., Modderman, T.: Resilience during crisis and the role of age: Involuntary telework during the COVID-19 pandemic. International Journal of Environmental Research and Public Health **19**(3), 1762 (2022)
31. Luthans, F., Vogelgesang, G.R., Lester, P.B.: Developing the psychological capital of resiliency. Human Resource Development Review **5**(1), 25–44 (2006)
32. Mallak, L.A., Shank, C.: Workplace resilience and performance: Workload and organizational constraints. Journal of Organizational Psychology **21**(6), 1-13 (2021)
33. Meneghel, I., Borgogni, L., Miraglia, M., Salanova, M., Martinez, I.M.: From social context and resilience to performance through job satisfaction: A multilevel study over time. Human Relations **69**(11), 2047–2067 (2016)
34. Siu, O.L., Yang, Y., Li, A., Wang, H., Ng, T.K.: Leveraging a dual-focused growth mindset to boost employee resilience and work well-being: Evidence from a two-wave survey and an intervention study. Stress and Health **41**(4), e70093 (2025)
35. Kovačević, I., Bétourné, N., Stankevičiūtė, Ž., Màthé, I., Anagnostopoulos, A., Abrudan, D., Anđelković Labrović, J.: Employee resilience and career perception across countries and occupations. In: ITAIS 2025 Conference, Castellanza, Italy, 17–18 October 2025 (2025)

Original Research Article *https://doi.org/10.5281/zenodo.23476939*

Exploring the Relationship of Self-Directed Learning and Employee Resilience in the AI Era

Jeremić, Emilija [1][0000-0002-4697-2791]

[1] Faculty of Organizational Sciences, 11000 Belgrade, Serbia
emilija.jeremic@gmail.com

Abstract. The growing integration of artificial intelligence (AI) in the workplace is reshaping job demands and requiring employees to continuously adapt. In this context, employee resilience has become a key capability for maintaining performance, while self-directed learning (SDL) is increasingly viewed as a central mechanism for adapting to technological change. This exploratory study explores the relationship between AI exposure and employee resilience, focusing on the mediating role of SDL. Drawing on social cognitive theory and conservation of resources theory, the study tests whether AI exposure contributes to resilience indirectly through proactive learning behaviours. Preliminary results indicate that AI exposure is positively associated with SDL, while SDL is a significant predictor of employee resilience, providing also some evidence of a mediating role of SDL in this relationship. The findings suggest that resilience in the AI era may depend less on the technology itself, and more on the employees' ability to adapt through learning. Despite some discussed limitations, this study provides a useful basis for future research in the area of employee adaptation to AI-induced change.

Keywords: AI, Human Resource Management, Self-Directed Learning, Employee Resilience, Digital Transformation, AI Exposure.

1. Introduction

The previous decade, which has accumulated the effects of the Industry 4.0 and COVID-19 pandemic, has largely been marked by the increased Internet usage, expansion of automatized learning and artificial intelligence, all of which has resulted in digital transformation becoming an imperative for the organizations wishing to remain competitive in the market, thus pushing them to acquire technological innovation as quickly as possible [5, 16, 20, 27]. Digital transformation is, therefore, fundamentally shaping the nature of work, organizational structures and the demands placed in front of the employees [5, 24].

In the context of the fourth industrial revolution, however, artificial intelligence (AI) no longer represents a mere technological innovation, playing instead the role of the catalyst of radical changes in the way our work is being effectuated, including related decision-making processes, as well as human resource management [16, 48]. The organizations are forced to adapt indefinitely, and with them also the employees, who need to keep developing new competencies and grasping new tools, living in a state of

constant uncertainty brought on by these technological changes [11]. Human resource management (HRM) in the digital era is faced with new challenges which go beyond its traditional functions of recruitment, learning & development and performance evaluation. Its focus is now increasingly being moved towards digital transformation navigation and leading comprehensive change management processes on the global level [33, 50]. This demands from the HR professionals a whole new set of digital skills and competencies in order to help them address these challenges in the newly-digitalized, data-driven workplace context, which simultaneously represents an opportunity as well as a major disruption for the HRM system itself, promoting rapid automatization of all the core HR processes [9, 14, 30, 50, 52]. Some authors recognize as the key drivers of this transformation (aside from the general digital transformation on the industry-level) factors such as higher customer expectations, competition, digital innovation governance tendencies and increasing geopolitical uncertainty [37, 43, 57]. Recent research has further underlined the importance of understanding this wider context of adaptation to digital transformation, especially in the HRM area. Bansal [5], for example, has pointed out the multidimensionality of the HR digital transformation construct, where the success of innovation within an organization largely depends on the integration of both digital and individual factors, as well as the constant dedication to upgrading organizational digital capability and creativity.

In such a turbulent environment, protective factors such as employee resilience emerge as the key elements of organizational sustainability [3, 6, 29]. Traditionally, individual resilience has been conceptualized as one's ability to bounce back after facing challenges [8]. Over the years, various studies have confirmed that resilient individuals are more successful in tackling diverse obstacles, adapting to change and performing at a similar level regardless of the surrounding pressure [34, 56]. It is important, however, not to regard resilience as an innate tendency available only to the luckier ones among us. Contemporary research shows resilience can be understood rather as an adaptive capacity of the person which can be developed and strengthened throughout one's lifetime, both through organizational resources and the employee's effort when it comes to the workplace setting [29]. Employee resilience can, therefore, be defined as an adaptive behavioural capacity for collecting, integrating and utilizing organizational resources, typically demonstrated by the employee's engagement in proactive and support-seeking actions, suggesting the need for the establishment of a supportive organizational environment where such behaviours (especially learning) are actively fostered [25, 26, 39]. A group of authors has even developed and validated the Employee Resilience Scale (EmpRes) which measures the presence of resilient organizational behaviours, including effective cooperation, workload and crisis management, learning from one's mistakes, feedback-seeking and regarding change as opportunity for growth, connecting it with factors such as job satisfaction and work engagement [38, 39].

Bearing all of this in mind, it is not surprising that learning & development is increasingly being researched as one of the central HR processes for building employee resilience and, subsequently, organizational sustainability, with the potential to ensure stability in the midst of the ever present changes, especially those brought on by the rapid AI expansion [17, 29, 41, 42, 54]. Learning can be seen as one of the mechanisms

of adaptation to AI-related changes [7], largely in line with Hobfoll's resource conservation theory [21] which points out the importance of using one's existing resources to develop new ones necessary for providing an adequate response to the growing demands of the environment. While the negative effects of technostress are undeniable [49], it is important not to disregard the developmental potential of AI exposure in the workplace, which can motivate various proactive learning behaviours among employees whilst engaging with these new technologies. According to previously-developed models of new technology acceptance, frequent interaction with new tools not only increases familiarity and reduces uncertainty, but also enhances employees' perceptions of their usefulness and relevance for professional development, where greater previous exposure and perceived usefulness of these tools positively predict individual willingness to adopt them in everyday work and learning routines [12, 53]. AI represents a particularly interesting case in this context, due to its ability to support adaptive learning by providing personalized feedback options, skill gap analysis and independent learning flow organization [47], which can in turn strengthen learners' resilience.

All of this puts into spotlight the role of self-directed learning (SDL) in the adoption of new skills and capacity development, a strategy defined as taking on individual responsibility for setting and attaining one's learning objectives independent of others. According to some of the original models, the three core elements of this learning process comprise self-direction, self-monitoring and motivation [19], consistent with Bandura's social cognitive theory which underlined the importance of self-efficacy and self-regulation in the context of learning [4]. In the professional environment, SDL can be conceptualized as a proactive process in which the employees recognize the need for further learning on their own, set their goals and oversee their own learning progress over time [47].

Various authors have confirmed the benefits of employing self-directed learning practices in the work context, as a part of the wider learning organization environment. Some of these include greater work engagement and higher perceived employability of professionals, closely related to adaptability to change, and therefore, resilience [13, 32, 36, 46]. In most of the studies, employee resilience had a mediating role between learning opportunities and proactive work engagement, where a constructive learning organization environment enhanced employee initiative and risk-taking [10, 31, 32, 36], while in the others, self-directed learning tendencies acted as mediators explaining the relationships between resilience and other constructs such as optimism and emotional intelligence [22].

Contemporary research has shown that various AI tools (such as intelligent tutoring systems, chatbots and other platforms) can significantly support and further enhance self-directed learning [40]. While many authors focused on the implementation of AI in learning practices within schools and higher academic settings [2, 18, 23], as well as testing its efficacy in the context of language acquisitions [28, 44], a smaller number of researchers has also started analyzing the nature of this relationship in the professional context [40]. Studies almost consistently demonstrate that AI-based learning interventions (regardless of the context) can positively affect learners' self-directed capacities (notably planning and autonomy), all the while reducing their cognitive load, and in-

creasing self-reflection and motivation, mainly by providing adequate and timely interactive feedback as well as greater opportunity for personalization, with most of the benefits growing with the length of the AI learning tool implementation [1, 35, 51, 55].

Building on all of the above, the scope of the present study is to further examine and attempt to explain the relationship between AI-driven human resource transformation and employee resilience, by assessing the role of self-directed learning in resilience strengthening under these circumstances. Specifically, the research aims to test whether regular exposure to AI in the workplace may enhance self-regulated learning tendencies among employees, which could in turn promote their resilience levels, as this connection seems to be underrepresented in existing literature. Special attention will be given to HR professionals, who, for the above-mentioned reasons, have the potential to pioneer this practice.

2. Methodology

2.1. Research Design and Hypotheses

This study adopts a quantitative cross-sectional survey design. Due to the fact that the integration of AI exposure, self-directed learning and employee resilience is still largely under-explored, especially in the organizational setting, this exploratory research aims to establish initial empirical patterns and test theoretically grounded assumptions based on previously presented theoretical framework [4, 21], in which SDL may represent one of the key behavioural mechanisms through which AI exposure can enhance employee resilience. This means that the employees might be more likely to respond to increased technological demands by engaging in proactive learning behaviours, which will in turn strengthen their adaptive capacity. Accordingly, and in line with previously analyzed literature, the following hypotheses are tested:

H1: AI exposure is positively related to self-directed learning.

H2: Self-directed learning is positively related to employee resilience.

H3: Self-directed learning mediates the relationship between AI exposure and employee resilience.

These hypotheses reflect the assumption that AI exposure does not directly lead to resilience, but instead operates through employees' active engagement in learning processes, consistent with earlier findings on the links between learning environments, proactive behaviours and resilient outcomes [10, 36].

2.2. Method

The research primarily targets HR professionals, as their unique positioning, as previously discussed, which is at the intersection of technological systems and human capital management, at the same time makes them more vulnerable to the consequences of technological change and more likely to engage in proactive resilience-building activities [5, 45]. The study also included professionals from other fields, to capture broader variability in AI exposure across industries. All the participants filled in an online ques-

tionnaire consisting of 25 items. All the constructs were measured using a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree), allowing for greater sensitivity in capturing individual differences.

Employee resilience is conceptualized here as a dynamic, behavioural capacity to adapt and recover in response to workplace challenges, rather than a fixed personality trait. It is measured using the validated Employee Resilience Scale [38, 39], capturing behavioral indicators such as learning from mistakes, feedback-seeking, adaptive coping and support-seeking behaviour. This scale demonstrates strong internal consistency, (Cronbach's $\alpha > .80$) and construct validity in prior organizational samples, and its reliability was re-examined in the present sample. The other two constructs (AI Exposure and Self-Directed Learning), for the sake of efficiency, are measured using short self-developed scales with descriptive items capturing the frequency of interaction with AI tools, perceived AI-related changes in work tasks and decision-making processes (for AI Exposure), and engagement in proactive learning behaviours, independent learning goal-setting, resource-seeking and outcome monitoring (for SDL). Given the exploratory nature of the study, these scales were not subjected to formal validation procedures, and their psychometric properties should be interpreted with appropriate caution. These explorative operationalizations were inspired by the perspectives highlighting familiarity with new technologies as one of the main indicators of technological exposure and acceptance [12, 53], as well as the existing conceptualizations of SDL both in and out of the workplace setting [19, 47, 15]. Control variables such as age, gender, education level and professional domain were also included.

2.3. Data Analysis

Quantitative data was analyzed using descriptive statistics and group comparisons (e.g. linear regression) to test the three hypotheses. Given the exploratory nature of the study and the modest sample size, no factor analysis or scale validation was attempted at this stage.

3. Results

3.1. Sample Description

The first analyzed sample (data collection is ongoing) consisted of 73 employed adults from Serbia, 83.6% of which were female, and 16.4% male. The age range varied from 23 to 74 years ($M = 35.08$, $SD = 11.19$), and their reported work experience ranged from 1 to 48 years ($M = 10.68$, $SD = 10.29$), and the majority of participants held a master's degree (56.2%), followed by a bachelor's degree (39.7%), while a small proportion held a doctoral degree (4.1%). Approximately one quarter of the sample were employed in the HR sector, while the remaining participants worked in other professional domains, including management, education, IT, finance, and healthcare. This distribution is consistent with the study's primary focus on HR professionals while still capturing variability in AI exposure across different industries.

3.2. Descriptive Statistics

Among the constructs, Employee Resilience showed the highest mean value ($M = 5.83$, $SD = 0.67$), followed closely by Self-Directed Learning ($M = 5.73$, $SD = 0.93$), suggesting that participants generally reported high levels of proactive learning behaviors and adaptive capacity at work. AI Exposure was, on average, near the theoretical midpoint of the scale ($M = 4.05$, $SD = 1.64$), with a substantially wider spread of responses, indicating considerable variability in the extent to which participants encountered AI tools and AI-related changes in their daily work.

Normality was assessed using the Shapiro-Wilk test together with skewness and kurtosis indices. While the Shapiro-Wilk test indicated mild deviations from normality for AI Exposure and SDL, all skewness and kurtosis values fell within the conventional +/- 1 range, with the exception of a slightly platykurtic distribution for AI Exposure. Given that linear regression and Pearson correlations are robust to mild deviations from normality, the distributions were treated as approximately normal for the purpose of subsequent analyses.

3.3. Quantitative Data Analysis

Bivariate Pearson correlations supported the basic pattern proposed by the hypotheses. AI Exposure was positively associated with SDL ($r = 0.39$, $p < 0.001$) and with Employee Resilience ($r = 0.31$, $p = 0.007$), and SDL was, in turn, strongly associated with Employee Resilience ($r = 0.50$, $p < 0.001$). All three correlations were in the predicted direction, providing initial support for H1 and H2. To verify that these results were not driven by mild non-normality in the AI Exposure and SDL distributions, Spearman's rank-order correlations were computed and showed an almost identical pattern ($\rho = 0.39$, 0.55, and 0.32; $p < 0.01$).

To formally test H1, a linear regression was conducted with AI Exposure as the predictor and SDL as the outcome. The model was significant, $F(1, 71) = 12.42$, $p < 0.001$, with AI Exposure explaining approximately 14.9% of the variance in SDL ($R^2 = 0.149$). AI Exposure positively predicted SDL ($\beta = 0.39$, $p < 0.001$), thus supporting H1. Similarly, H2 was tested through a second linear regression with SDL predicting Employee Resilience. This model was also significant, $F(1, 71) = 24.05$, $p < 0.001$, with SDL accounting for 25.3% of the variance in Employee Resilience ($R^2 = 0.253$). SDL positively predicted Employee Resilience ($\beta = 0.50$, $p < 0.001$), also supporting H2.

In order to examine H3, AI Exposure and SDL were tested together as predictors of Employee Resilience in a multiple regression analysis. The overall model was also significant, $F(2, 70) = 12.90$, $p < 0.001$, $R^2 = 0.27$, Adjusted $R^2 = 0.25$, indicating that the two predictors jointly explained approximately 27% of the variance in Employee Resilience. SDL remained a strong, significant predictor ($\beta = 0.45$, $p < 0.001$), while AI Exposure was no longer a significant predictor once SDL was included in the model ($\beta = 0.14$, $p = 0.216$). This pattern could be seen as informative, as although AI Exposure was significantly associated with Employee Resilience when considered on its own ($\beta = 0.31$, $p = 0.007$, $R^2 = 0.10$), this direct association led to non-significance once SDL was included in the equation, which could point to a SDL-mediate relationship, providing preliminary support for H3.

Lastly, independent samples t-tests were conducted to examine whether HR professionals differed from the other employees on the three core variables. HR participants reported significantly higher levels of AI Exposure ($M = 4.87$, $SD = 1.37$) than non-HR participants ($M = 3.76$, $SD = 1.64$), $t(71) = 2.65$, $p = 0.010$, $d = 0.71$, indicating a medium-to-large effect. By contrast, no significant differences were observed between HR and non-HR participants on SDL or on Employee Resilience, suggesting that greater AI exposure of HR professionals from this sample in their daily work does not render their AI-SDL-Employee Resilience pathway any different that in the other professional groups. Additional t-tests revealed no significant gender differences on any of the tested variables, and Pearson correlations indicated that age was not significantly related to AI Exposure, SDL, or Resilience ($p > 0.23$).

Cronbach's α values indicated good to excellent internal consistency and reliability across the three scales, with $\alpha = 0.89$ for AI Exposure, $\alpha = 0.90$ for Self-Directed Learning and $\alpha = 0.84$ for Employee Resilience, and item-rest correlations supported the inclusion of each item in its respective scale. Given the exploratory nature of this study, these values were considered sufficient for the present analyses.

4. Discussion

The aim of this exploratory study was to examine the relationship between AI exposure in the workplace and employee resilience, with self-directed learning (SDL) hypothesized as a key behavioural mechanism linking the two. Drawing on social cognitive theory [4] and the conservation of resources perspective [21], it was proposed that exposure to AI in the daily work environment could activate proactive learning behaviours, which can in turn strengthen employees' adaptive capacity. The findings provide preliminary empirical support for all three hypotheses and, taken together, suggest that the developmental potential of AI exposure for employee resilience is largely realized through the active engagement of employees in self-directed learning processes.

In line with H1, it can be observed that the employees who reported higher exposure to AI in their work also reported higher levels of self-directed learning. This finding aligns with technology acceptance frameworks, which propose that frequent interaction with new tools can increase familiarity and perceived usefulness, motivating further engagement [12, 53], and with the conservation of resources theory positing that employees mobilize existing resources to acquire new ones in response to growing job demands [21]. It also resonates with recent evidence linking technological change to learning demands and proactive skill development among employees [7]. The supplementary observation that HR professionals reported significantly higher AI exposure than employees from other domains, while showing comparable levels of SDL, further suggests that exposure translates into proactive learning at similar rates across professional groups, at least in this sample. H2 was also supported, with the data showing SDL as a strong and positive predictor of employee resilience, accounting for approximately one quarter of its variance. This is consistent with a growing body of research linking learning-oriented organizational climates and proactive learning behaviours to adaptive workplace outcomes [10, 32, 36], as well as with conceptualizations of resilience as a developable, behaviourally enhanced capacity rather than a stable trait [25,

29, 39]. Employees who set their own learning goals, monitor their progress, and proactively seek new knowledge appear to be better equipped to learn from setbacks, manage demanding workloads, and reframe change as an opportunity, which are precisely the behavioural markers captured by the Employee Resilience Scale [39].

When it comes to H3, the support extracted from this sample can be described as theoretically informative. While AI Exposure was significantly associated with Employee Resilience when considered on its own, this association was no longer significant once SDL was included as a predictor in the multiple regression model. SDL, by contrast, remained a strong predictor of Employee Resilience after controlling for AI Exposure. This pattern is consistent with the classical mediation logic, suggesting that AI Exposure does not directly enhance employee resilience, but rather it appears to do so by activating proactive learning behaviors that, in turn, strengthen adaptive capacity. This finding may help to better nuance the existing negative narrative around technostress [49] by highlighting a developmental pathway that may co-exist with the stress-related effects commonly documented in the literature. It also aligns with recent evidence that AI-supported learning environments can enhance learners' self-regulatory capacities and, indirectly, their resilience [1, 55].

Building on prior work that has positioned learning as a key mechanism of adaptation to AI-related change [7], the present results can further help specify what kind of learning matters most, i.e. the one that is not externally imposed, but rather initiated by employees themselves, in the form of self-directed engagement with new knowledge and skills. This can carry several practical implications for organizations seeking to build a resilience workforce in the AI era. For example, AI exposure alone is unlikely to increase employee resilience unless it is accompanied by proactive, self-directed learning. Organizations can support this through enhancing learning autonomy, accessible resources and growth-oriented communication about AI adoption. As previously mentioned, HR professionals could also play a key role as early adopters and promoters of these learning practices across organizations.

5. Limitations

Several limitations should be acknowledged. First, the cross-sectional design limits the possibility of drawing causal inferences, and longitudinal or experimental designs could further be employed to provide more structured empirical insights. Second, the achieved sample (N=73) is somewhat smaller than the planned target of 100, and ongoing data collection may reveal additional currently invisible patterns. Third, for the purpose of efficacy and participant motivation, the AI Exposure and SDL measures were short, self-developed scales designed specifically for exploratory use. Although they demonstrated solid internal consistency in this sample, their dimensionality and validity require formal evaluation in future confirmatory work. Fourth, the sample was predominantly female (83.6%) and drawn from a single national context, which constrains generalizability, so replication in a more gender-balanced and cross-cultural sample (alongside better distribution across professions) is therefore recommended.
Fifth, mediation was examined through regression-based inference rather than more contemporary bootstrapping methods, which would provide a more direct and rigorous

test of the indirect effect in future research. Finally, all measures were self-reported, implicating the need for the implementation of more objective and more nuanced tools in the future, including also additional indicators of resilience such as job performance, perceived well-being etc., to provide a more comprehensive picture of how technological change reshapes employee adaptation.

Disclosure of Interests. The author has no competing interests to declare that are relevant to the content of this article.

References

1. Achuthan, K.: Artificial intelligence and learner autonomy: A meta-analysis of self-regulated and self-directed learning. Frontiers in Education 10, 1738751 (2025)
2. Ali, F., Choy, D., Divaharan, S., Tay, H.Y., Chen, W.: Supporting self-directed learning and self-assessment using TeacherGAIA. Learning: Research and Practice 9(2), 135–147 (2023)
3. Bagga, G.: How to keep the talent you have got. Human Resource Management International Digest 21(1), 3–4 (2013)
4. Bandura, A.: Social foundations of thought and action: A social cognitive theory. Prentice-Hall, Englewood Cliffs (1986)
5. Bansal, A., Panchal, T., Jabeen, F., Mangla, S.K., Singh, G.: Human resource digital transformation. Journal of Business Research 157, 113611 (2023)
6. Bardoel, E.A., Pettit, T.M., De Cieri, H., McMillan, L.: Employee resilience. Asia Pacific Journal of Human Resources 52(3), 279–297 (2014)
7. Blaique, L., Ismail, H.N., Aldabbas, H.: Organizational learning and resilience. International Journal of Productivity and Performance Management 72(6), 1584–1607 (2023)
8. Bonanno, G.A.: Loss, trauma, and human resilience. American Psychologist 59(1), 20–28 (2004)
9. Bresciani, S., Huarng, K.H., Malhotra, A., Ferraris, A.: Digital transformation as a springboard. Journal of Business Research 128, 204–210 (2021)
10. Caniëls, M.C.J., Baaten, S.M.J.: Learning-oriented climate and resilience. Social Indicators Research 143, 561–577 (2019)
11. Cascio, W.F., Montealegre, R.: Technology and organizations. Annual Review of Organizational Psychology and Organizational Behavior 3, 349–375 (2016)
12. Davis, F.D.: Perceived usefulness and user acceptance. MIS Quarterly 13(3), 319–340 (1989)
13. Decius, J., Knappstein, M., Klug, K.: Work-related learning and employability. European Journal of Work and Organizational Psychology 33(1), 24–39 (2024)
14. DiRomualdo, A., El-Khoury, D., Girimonte, F.: HR in the digital age. Strategic HR Review 17(5), 234–242 (2018)
15. Fontana, R.P., Milligan, C., Littlejohn, A., Margaryan, A.: Measuring self-regulated learning. International Journal of Training and Development 19(1), 32–52 (2015)
16. Fregnan, E., Procino, R., Scaratti, G.: HRM 4.0 competencies. Frontiers in Psychology 11, 578251 (2020)
17. Garavan, T.N., Carbery, R., Rock, A.: Talent development architecture. European Journal of Training and Development 36, 5–24 (2012)
18. Garg, A., Rajendran, R.: Generative AI and self-directed learning. In: Sifaleras, A., Lin, F. (eds.) ITS 2024, LNCS, vol. 14798, pp. 1–13. Springer, Cham (2024)

19. Garrison, D.R.: Self-directed learning: Toward a comprehensive model. Adult Education Quarterly 48(1), 18–33 (1997)
20. Gerards, R., van Wetten, S., van Sambeek, C.: New ways of working. Review of Managerial Science 15(7), 2075–2110 (2021)
21. Hobfoll, S.E.: Conservation of resources: A new attempt at conceptualizing stress. *American Psychologist* 44(3), 513–524 (1989)
22. Hwang, E.H., Kim, K.H.: Emotional intelligence and resilience. Frontiers in Public Health 11, 1182689 (2023)
23. Kassymova, G.K., Talgatov, Y.K., Arpentieva, M.R., Abishev, A.R., Menshikov, P.V.: Artificial intelligence in the development of the theory and practices of self-directed learning. *Multidisciplinary Journal of Thought and Research*1(3), 66–79 (2025)
24. Kraus, S., Jones, P., Kailer, N., Weinmann, A., Chaparro-Banegas, N., Roig-Tierno, N.: Digital transformation overview. SAGE Open 11(3), 1–15 (2021)
25. Kuntz, J.C., Malinen, S., Näswall, K.: Employee resilience development. Consulting Psychology Journal 69(3), 223–242 (2017)
26. Kuntz, J.C., Näswall, K., Malinen, S.: Resilient employees. Industrial and Organizational Psychology 9(2), 456–462 (2016)
27. Lanzolla, G., Pesce, D., Tucci, C.L.: Digital transformation in innovation. Journal of Product Innovation Management 38(1), 90–108 (2021)
28. Lashari, A., Umrani, S.: Reimagining self-directed learning language in the age of artificial intelligence: A systematic review. 57, 92–114 (2023)
29. Lengnick-Hall, C.A., Beck, T.E., Lengnick-Hall, M.L.: Organizational resilience through HRM. Human Resource Management Review 21(3), 243–255 (2011)
30. Makridakis, S.: AI revolution and firms. Futures 90, 46–60 (2017)
31. Malik, P.: Learning organization and resilience. Asia-Pacific Journal of Business Administration 15(3), 325–344 (2023)
32. Malik, P., Garg, P.: Learning organization and engagement. International Journal of Human Resource Management 29(8), 1–26 (2017)
33. Manuti, A., De Palma, P.D.: Digital HR: A critical management approach. Palgrave Macmillan, London (2018)
34. Meneghel, I., Salanova, M., Martínez, I.M.: Team resilience and performance. Journal of Happiness Studies 17, 239–255 (2016)
35. Mncube, D.W., Maphalala, M.C., Mkhasibe, R.G.: AI in higher education. Turkish Online Journal of Distance Education 27(1), 275–290 (2026)
36. Mohammad, S., Sağsan, M., Şeşen, H.: Learning organizations and resilience. SAGE Open 14(4), 1-21 (2024)
37. Mosca, M.: Digitalization of HRM: A study of success factors and consequences in the last decade. University of Twente, Enschede (2020)
38. Näswall, K., Kuntz, J., Malinen, S.: Employee Resilience Scale (EmpRes): Technical report. Resilient Organisations Research Report 2015/04. Resilient Organisations, University of Canterbury (2015)
39. Näswall, K., Malinen, S., Kuntz, J., Hodliffe, M.: Employee resilience scale. Journal of Managerial Psychology 34(5), 353–367 (2019)
40. Navas Bonilla, C.d.R., Viñan Carrasco, L.M., Gaibor Pupiales, J.C., Murillo Noriega, D.E.: Self-directed learning with AI. Future Internet 17(8), 366 (2025)
41. Noe, R.A., Clarke, A., Klein, H.: Learning in the workplace. Annual Review of Organizational Psychology and Organizational Behavior 1, 245–275 (2014)
42. OECD: OECD skills outlook 2021: Learning for life. OECD Publishing, Paris (2021)

43. Osmundsen, K., Iden, J., Bygstad, B.: Digital transformation: Drivers, success factors, and implications. In: *MCIS 2018 Proceedings*, pp. 1–10 (2018)
44. Oveisi, M., Nosratinia, M.: Self-directed learning and resilience. Journal of Language and Translation 9(4), 77–91 (2019)
45. Parry, E., Strohmeier, S.: HRM in the digital age: Digital changes and challenges of the HR profession. *Employee Relations* 36(4), 345-361 (2014)
46. Piotrowski, M.: Self-directed learning and resilience. International Journal of Self-Directed Learning 17(1), 19–38 (2020)
47. Sitzmann, T., Ely, K.: Meta-analysis of self-regulated learning. Psychological Bulletin 137(3), 421–442 (2011)
48. Strohmeier, S., Parry, E.: HRM in the digital age: Digital changes and challenges of the HR profession. *Employee Relations* 36, 345–361 (2014)
49. Tarafdar, M., Cooper, C.L., Stich, J.-F.: The technostress trifecta. Information Systems Journal 29(1), 6–42 (2019)
50. Tripathi, R., Kushwaha, P.: Digital HRM practices. Journal of Management Research 9(1), 15–25 (2017)
51. Umar, U., Purwanto, M.B.: AI and self-directed learning. ETERNAL 16(2), 457-465 (2025)
52. Vardarlier, P.: Digital transformation of human resource management: Digital applications and strategic policies of the new era. In: *Digital Business Strategies in Blockchain Ecosystems*, pp. 239–264. Springer, Cham (2020)
53. Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D.: User acceptance of IT. MIS Quarterly 27(3), 425–478 (2003)
54. World Economic Forum: The future of jobs report 2023. World Economic Forum, Geneva (2023)
55. Younas, M., Naveed, M.A., Khalid, S.: AI-driven self-directed learning. IEEE Access 13, 38385–38401 (2025)
56. Youssef, C.M., Luthans, F.: Positive organizational behavior. Journal of Management 33(5), 774–800 (2007)
57. Zhang, J., Chen, Z.: HRM digital transformation. Journal of the Knowledge Economy 15, 1482–1498 (2024)

THEMATIC SECTION

Organizational Strategy and Hybrid Project Management

Resilient and Intelligent Organizations
in Times of Permanent Disruption

SYMORG 2026 • CONFERENCE PROCEEDINGS

Original Research Article *https://doi.org/10.5281/zenodo.23476954*

Integrating Design Thinking into the Innovation Funnel: A Framework for Improving Innovation Performance in Entrepreneurship

Aleksandar Ivković[1][0009-0000-2395-7567] and dr Marko Mihić[1][0000-0003-3994-2268]

[1] Faculty of Organizational Sciences, University of Belgrade, Serbia
aivkovic1997@gmail.com

Abstract. Innovation plays a key role in entrepreneurial success, yet many initiatives fail to achieve sustainable market impact. This paper examines innovation failure and proposes a structured approach to improving innovation outcomes. First, the Diffusion of Innovations framework highlights how early successes often create the illusion of market acceptance, while misunderstanding the pragmatic majority causes the innovation to remain in the "Chasm". Second, the Innovation Funnel illustrates how weak internal innovation management contributes to these failures. Building on these insights, the paper introduces Design Thinking as a practical methodology to improve innovation performance. By emphasizing problem understanding, iterative testing, and user-centered validation, Design Thinking increases the likelihood that ideas passing through the Funnel will become market-successful innovations.

Keywords: Design Thinking, Diffusion of Innovation, Innovation Funnel.

1 Introduction - Why Innovations Fail in Entrepreneurship

Innovation plays a central role in entrepreneurship and economic development. Milutinović et al. identify that strategic innovation is recognized as a key driver of competitiveness and long-term growth, particularly in SMEs [1]. However, not all innovations are successful. Many initiatives fail to create value for users or achieve sustainable market impact. Empirical studies show that a significant proportion of innovation initiatives do not reach commercial success, despite investments in development and marketing [2]. These outcomes are often not the result of a lack of creativity, but rather of deficiencies in the way the innovation management process is executed within organizations.

At the highest level, one of the key reasons for such outcomes lies in the misunderstanding of what innovation actually is. In this regard, Trott [3] emphasizes that innovation should not be understood merely as the generation of an idea. Instead, innovation represents a structured and managed process through which ideas are transformed into practical applications through successful commercialization. Tidd and Bessant [4] support this view by highlighting the importance of managing innovation

from ideation to market entry. Therefore, innovation builds upon ideas, but it extends beyond them.

Idea as such refers to the foundation of the innovation process. According to Birss [5], an idea is not a sudden moment of inspiration, but a new combination of existing elements such as knowledge, experience, and problem - related information. It is formed when these elements are connected in a meaningful way in order to address a specific need or challenge. The existence of ideas alone does not guarantee innovation success.

Beneath this highest level, two key underlying levels of mismanaged innovation can be identified. These levels reveal a more specific mechanism through which promising ideas fail to evolve into successful innovations.

1.1 Superficial User Understanding

Non-viable ideas often arise from an insufficient or incorrect understanding of users. When organizations rely on a superficial view of user needs, they risk defining the problem inaccurately. The process of deeply understanding users such as their habits, behaviors, and real challenges is frequently shortened or treated as secondary. This aligns with studies showing that superficial user research is one of the leading causes of new product failure [6].

As a result, the defined problem may be incorrectly defined. Decisions and ideas developed on the basis of such flawed problem definitions are unlikely to reflect actual user needs. Even well-developed concepts may struggle during implementation, not because of poor execution, but because they were built on an incorrect understanding from the very beginning. As Thomke and von Hippel suggest, misalignment between user understanding and solution design affects the success of innovation [7].

1.2 Premature Attachment to Ideas

At the next level, one of the main causes of innovation failure lies in the premature attachment to ideas. Organizations often become committed to an idea before critically examining its validity. They tend to believe too quickly that the idea will automatically create value, without sufficiently testing whether it truly addresses a real need. This tendency is particularly visible in startup environments, where enthusiasm, speed, and strong personal attachment to an idea can overshadow objective evaluation. Therefore, Ries [8] highlights that startups operate under uncertainty and must test assumptions rather than rely on intuition. However, instead of validating the problem and user demand, attention is often focused on developing and refining the idea itself. As a result, the idea is treated as a solution before its actual usefulness has been confirmed.

Such a way of thinking often leads to decisions based on assumptions rather than evidence. The initial idea is not tested in order to examine its validity. Instead, it is taken for granted. Organizations may subjectively conclude that the idea will create value for users, without verifying whether the underlying problem truly exists or whether the proposed solution is relevant.

Many ideas that seem promising at first fail to deliver value when exposed to real users and market conditions. This gap between belief and actual user need represents one of the fundamental risks in the early stages of innovation management. To address this challenge, this paper contributes to innovation management research by developing a unified innovation framework through the integration of Design Thinking, the Innovation Funnel, and Diffusion of Innovations in order to support organizations in improving innovation performance.

In summary, the failure of innovations in entrepreneurship is rarely the result of a lack of ideas. It stems from ineffective innovation management. Thus, as Cooper points out, the central challenge in entrepreneurship is systematic management of ideas throughout the innovation process [9].

2 Theoretical Framework: Mechanism of Innovation Failure

Research conducted by McKinsey & Company [10] reveals a significant inconsistency in the contrast between the strategic importance of innovation and its real-world results. According to the study, 84% of executives consider innovation essential for achieving competitive advantage and long-term company success. However, only 6% report being satisfied with their organization's innovation performance.

These results reveal a clear gap between the high expectations regarding innovation and its actual impact. In simplified terms, if we consider 100 innovation-oriented companies or initiatives, only a small number would achieve their established innovation goals.

While public and business narratives often emphasize successful innovation stories, they rarely reflect the much larger number of initiatives, products, and startups that fail to achieve sustainable market success. The data therefore suggest that poor innovation management is more widespread than assumed.

This gap, according to Cooper, points toward a systemic weakness in the way innovation processes are organized and executed within organizations [9]. To better understand why it persists, it is necessary to examine it from a theoretical perspective.

2.1 The Phenomenon of Insufficient Market Adoption of Innovations

Many innovations generate initial interest and find practical application among a limited group of users. This early traction may create the impression that the innovation is on a clear path to success. However, such early achievements can be misleading if the subsequent innovation management process is not properly executed.

In such cases, the innovation remains limited to a narrow segment of users and fails to reach the mass market. As a result, problems arise in the process of market adoption. A key underlying cause of this stagnation is an insufficient understanding of different user groups and their specific needs.

This phenomenon can be better understood through Rogers' Diffusion of Innovations theory [11], which explains how innovations spread through a social system over time. Rogers identifies five distinct groups of users based on how quickly they adopt new solutions and in the criteria they use to evaluate whether an innovation is worth adopting.

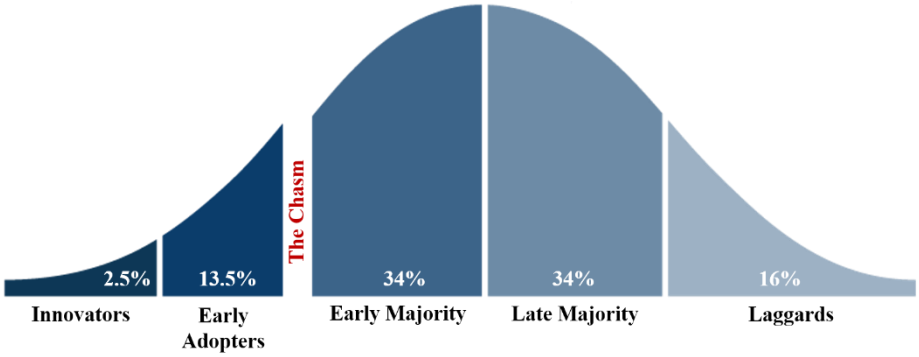


Fig. 1. Diffusion of Innovations model with the Chasm. Source: Author's own illustration based on Rogers (2003) and Moore (2001).

Although the ultimate objective of most innovations is to reach the mass market, represented by the early majority and late majority, adoption typically begins with innovators and early adopters. Rogers explains that innovators are generally willing to experiment with new solutions simply because they are new, demonstrating a high tolerance for uncertainty. Early adopters are more selective, but still open to innovation when they perceive a clear benefit [11].

The most significant differences in mindset emerge when the innovation attempts to transition from early adopters to the early majority. This group is more cautious and pragmatic, being less driven by novelty and more concerned with reliability and the social validation of innovation. At this point, if innovation management is not adapted, the innovation may fall into what is known as "the Chasm" [12], a critical emptiness between early adopters and mass market. From Moor's perspective, when an innovation becomes trapped in the Chasm, it risks failing to reach the mass market [12]. Organizations must therefore recognize this moment and actively manage this Chasm by adjusting the innovation to the expectations of the pragmatic majority.

Unawareness of the need to cross the Chasm initiates the mechanism of innovation failure. Entrepreneurs often assume that the interest of early adopters signals mass demand, as early traction is perceived as confirmation of product - market fit. However, Ries points out that such assumptions must be validated rather than derived from limited early feedback [8]. This leads to premature strategic decisions based on limited feedback.

At the same time, organizations often fail to adapt their value proposition to the expectations of the pragmatic majority, continuing to communicate the innovation in ways that resonate primarily with innovators and early adopters. As a result, what initially appears to be a promising innovation gradually loses momentum.

In conclusion, the mechanism of innovation failure emerges from the way organizations manage the innovation process. A structured perspective for understanding this progression is provided by the Innovation Funnel framework [4].

2.2 The Innovation Funnel: The Core of Innovation Failure

Innovation management literature describes the Innovation Funnel as a visual framework used by organizations to manage new ideas, guiding them from a wide range of concepts through evaluation and testing [4]. Its goal is to filter and develop only the most promising ideas, progressing them to a successful market - ready innovation.

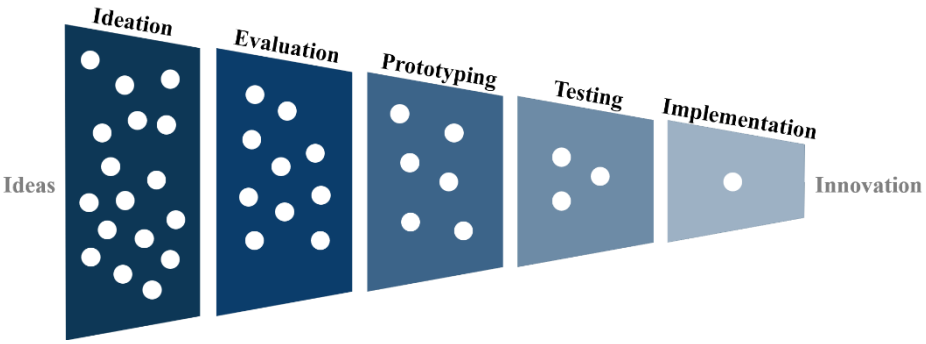


Fig. 2. Innovation Funnel Model. Source: Author's own illustration.

The challenges observed in mass market adoption, such as an innovation's inability to cross the Chasm, often originate from how organizations manage the innovation process internally. While the Diffusion of Innovations illustrates the external effect or symptom on the market, the Innovation Funnel highlights the internal causes that lead to this outcome. These problems often originate in the early phases of the Innovation Funnel, when ideas are still raw and not yet grounded in solid insights, making them insufficiently prepared for successful implementation. This aligns with findings that inefficiencies in organization are often systemic rather than isolated [13].

Ideation. During this phase, organizations aim to generate a large number of ideas from different sources, which provides a diverse starting point for the Innovation Funnel. However, the main challenge is that these ideas are often based on assumptions rather than validated insights. Ries criticizes this approach, emphasizing that such assumptions should be tested early in order to determine whether they are valid, as rejecting weak ideas at this phase is far less costly when only minimal resources have been invested [8]. In practice, however, teams often become attached to their ideas instead of defining a correct problem.

Evaluation. Such unvalidated ideas move into the next phase, where they are not always evaluated using appropriate criteria. Many organizations fail to engage enough with potential users in order to gather insights that could guide the selection of ideas. Therefore, Ulwick emphasizes how important it is to understand the customer as these insights help identify unfulfilled needs, frustrations users experience with existing solutions, and the importance of the problem the idea attempts to solve [6]. Understanding these factors is critical, as the importance of the problem for users determines whether they are willing to adopt and pay for a new solution.

Prototyping. Incorrectly evaluated unvalidated ideas enter the prototyping phase. However, this often occurs without clearly defined hypotheses explaining why the idea is being prototyped. According to Brown, prototyping should serve as a tool for learning and exploration [14]. Thus, in the absence of such hypotheses, the purpose of prototyping is set incorrectly, frequently focusing on testing the feasibility of the idea rather than exclusively presenting it.

Testing. In the final phases of the Funnel, such ideas enter the testing phase, where teams often focus on confirming that the idea itself is sustainable, rather than learning from the testing and validating the hypotheses [8]. As a result of this process, the organization may end up creating a solution that people actually do not want.

The problems identified across the phases of Innovation Funnel are supported by the research of Robert Curedale [15], who states that around 75% of products fail in the market because organizations do not properly understand users and their problems. This indicates that improving innovation performance requires approaches that place users at the center of the innovation process and emphasize a deeper understanding of their needs. One such approach is Design Thinking.

3 Research Methodology

This research examines the mechanism of innovation failure in entrepreneurship, with a focus on the role of the Innovation Funnel and the potential for its improvement through the integration of Design Thinking. The subject of the research is the relationship between innovation process management within the organization and its external outcome in terms of market adoption [4]. The aim of the research is to develop a conceptual framework that explains how integrating Design Thinking into the Innovation Funnel can improve innovation performance by addressing key causes of failure [9].

The research is based on a qualitative and conceptual approach, relying on the analysis and synthesis of relevant literature in the fields of innovation management and Design Thinking. The study does not include primary empirical data, but develops its conclusions through theoretical integration of existing models, focusing on the Innovation Funnel and Design Thinking [14].

The focus of the research is placed on the early stages of the Innovation Funnel, since they represent the critical point at which the main causes of innovation failure emerge [4]. Based on the analysis presented in the previous sections, two key conceptual propositions are identified:

Proposition 1: Superficial user understanding and premature attachment to ideas increase the likelihood of innovation failure in the early stages of the Innovation Funnel.

Proposition 2: Integrating Design Thinking into the Innovation Funnel improves problem definition accuracy and reduces the risk of market adoption failure.

4 Improving Innovation Performance through Design Thinking

4.1 The Design Thinking Approach to Innovation

In the context of innovation, problems can be categorized into three categories [16]:

- Simple
- Complicated
- Complex

Simple problems are fully understood and have straightforward solutions. Complicated problems involve known unknowns. We know what we know, but we also know what we don't know. Certain aspects of the problem are not clear, but solutions exist and can be identified with expert knowledge. Complex problems include unknown unknowns. We know what we know and we know what we don't know, but we also don't know what we don't know. These problems do not have a single solution and are difficult to resolve using conventional linear approaches [16].

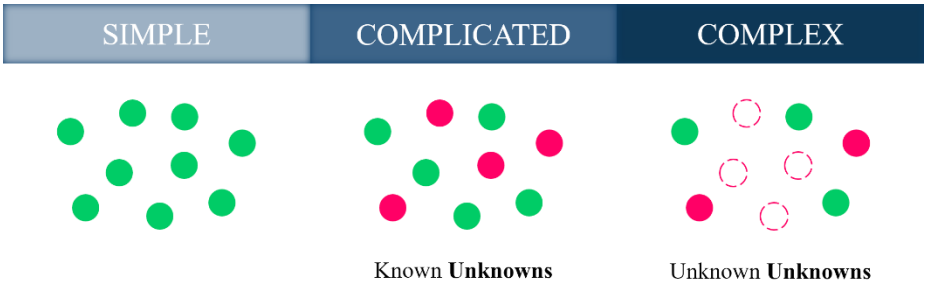


Fig. 3. 3 Categories of Problems. Source: Author's own illustration based on Scrum.org [16].

In today's world, user problems in innovation are typically complex. In such circumstances, when user needs and problems have many unknown unknowns, Design Thinking finds its strongest application and helps generate innovative solutions. This methodology enables organizations to approach problems with empathy toward users, develop prototypes, and continuously test solutions to foster innovation. Importantly, as Brown points out, it helps not only in solving problems effectively but also identifying the right user problems that need to be addressed [14]. The five phases of Design Thinking provide an iterative approach to navigate these complex user problems.

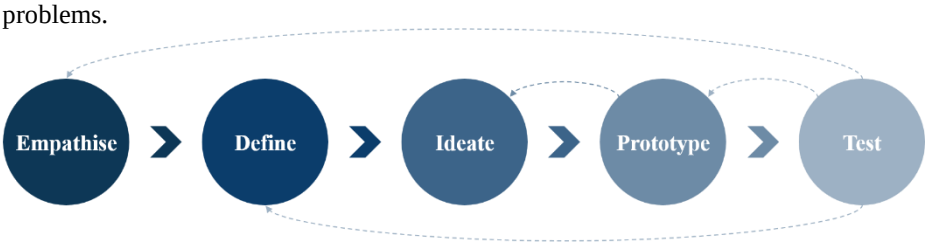


Fig. 4. The five phases of the Design Thinking approach. Source: Author's own illustration.

4.2 Integrating Design Thinking into the Innovation Funnel

Although the Innovation Funnel represents a fundamental framework for managing the innovation process, the discussion in the preceding sections has shown that various problems may arise across its phases, negatively affecting overall innovation performance. Integrating Design Thinking into the Innovation Funnel can improve this performance by aligning the phases of Design Thinking with the corresponding phases of the Funnel, which are conceptually compatible. Such synchronization reflects a shift toward user-centered innovation approach and more effective problem framing [17]. Furthermore, such integration also requires structured governance mechanisms to ensure effective support for innovation process, which aligns with the importance of governance highlighted by Đorđević et al. [18]. By applying methods and techniques from Design Thinking within these phases, organizations are enabled to develop solutions that users actually want.

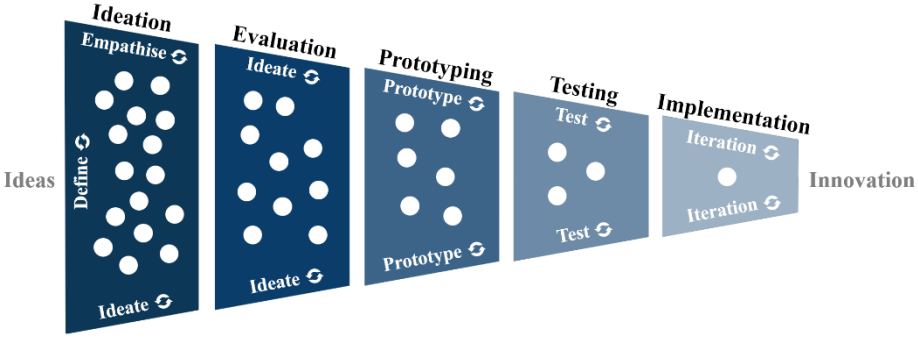


Fig. 5. Integration of Design Thinking phases within the Innovation Funnel. Source: Author's own illustration.

Design Thinking leaves its strongest impact in the first phase of the Innovation Funnel by shifting the organization's mindset and changing the way innovation is approached. For this reason, the first three phases of Design Thinking (Empathize, Define, and Ideate) are integrated into the Ideation phase.

Empathy. The goal of Ideation phase is to generate a wide range of ideas. But they must be well-grounded. Therefore, it is essential to accurately define the problem the user faces and evaluate its importance, as this determines whether they will truly value the solution and be willing to pay for it. Liedtka describes that value is created by aligning solutions with deeply understood user needs [19]. Those problems that cause negative experiences or frustrations for the user are the ones that hold the greatest significance. Only with this understanding can the team generate ideas that are relevant and solve a real problem. And empathy is the key to achieving this understanding. It allows us to put ourselves in the user's position and understand how they perceive and experience the world. Because in today's world, things are often not seen as they truly are, which is why observations and conclusions should be treated as assumptions until

the user is fully understood. Methods such as ethnography, observation, and interviews can be used for developing empathy and gaining deeper insights into users [17].

Users experience problems with different levels of intensity, which raises the question of which users to empathize with? And the answer is: extreme users. These are users who have intense needs and for whom the problem may be of high importance. The advantage of such users is that they are real, so when it comes time to test the prototype and gather feedback, it is clear who they are and where to find them.

Defining the problem. After collecting user information through the empathy, the next step is to synthesize the data to define the user's problem. As Curedale suggests, synthesis can be conducted using methods such as empathy maps, affinity diagrams, user journey mapping, and user experience mapping [15].

Problems arise from unmet user needs. In other words, a problem represents the gap between what users need and what existing solutions provide. Therefore, it is important to identify key needs, as they represent the most valuable problems to solve. Liedtka states that this perspective reflects a core principle of Design Thinking, focusing on problem framing and a user-centered approach [19]. Additionally, bad user experiences reveal unmet needs and represent the root cause of problem, as users develop needs when their experience does not meet expectations. Often, they tend to use a solution to achieve a desired experience, rather than valuing the solution itself.

A well-defined problem ensures that innovation is focused on real user needs rather than assumptions.

Ideation Phase. The transition moment from the problem domain to the solution domain. Once the problem is well-defined, the focus shifts to generating potential solutions. The goal of Ideation phase within the Innovation Funnel is to generate a wide range of ideas in order to identify the most promising ones. According to Curedale, Design Thinking supports this process by offering various brainstorming techniques that encourage idea generation [15]. In parallel, it fosters a mindset in which ideas are not immediately evaluated, allowing teams to think more freely and increase the diversity of generated ideas, as emphasized by the Hasso Plattner Institute [20].

Ideate within Evaluation Phase. All generated ideas now go through a screening process, where they are filtered based on set criteria. Tidd and Bessant note that if these criteria are not properly set, there is a risk that promising ideas may be prematurely rejected [4]. Furthermore, Brown highlights that Design Thinking shifts the mindset toward a user-centered perspective. Through fostering empathy and a clearly defined problem, the evaluation criteria are set from the user's point of view [14].

Prototype within the Prototyping Phase. Ideas that have successfully passed the Evaluation phase now require prototyping. Curedale explains that the goal of the prototyping phase is to create a prototype that simulates the idea. And the purpose of the prototype is to present the idea, not to test its feasibility which is a common mistake organization often make [15]. Furthermore, organizations should be aware of the reason for prototyping, which can be classified into three categories:

1. To visually explain the idea more effectively
2. To interact with users
3. To demonstrate customer value

In addition, Curedale offers several types of prototypes, and organizations should choose the one that best fits them [15].

- Paper prototypes
- Interactive prototypes
- Video prototype
- Physical prototype
- Pinocchio prototype
- The wizard of Oz
- Minimum Viable Products (MVP)

Test within the Testing Phase. Ready prototypes are now introduced to extreme users for testing to validate the idea. And idea validation follows a process supported by Ries [8], who emphasizes testing assumptions through the following steps.

1. Defining assumptions

The first step is to clearly define assumptions about the solution and its users so they can be tested.

2. Identifying critical assumptions

Critical assumptions are those on which the idea is based. As such, they represent a priority and cannot remain untested. They must be correct, because if they are not, it indicates that the solution is not the right fit for users. These assumptions are essential to the MVP prototype.

3. Analyzing test results

Analysis of the test results shows whether users receive enough value to solve their problem, which is determined in how much their experience improves by using the solution.

When targeting extreme users for testing, they are selected from the Diffusion of Innovations user group that is most likely to adopt the solution first [11]. The organization then chooses from a variety of testing methods that Design Thinking supports, depending on the reasons for prototyping mentioned in the previous phase, .

Implementation Phase. In the final phase of the Funnel, the prototype undergoes further development to become a real product, with the aim of entering the market and becoming an innovation. However, as Cooper claims, not every idea is mature enough for commercialization at this point, even if it has passed through the whole Funnel [9]. Therefore, it is crucial to make a well-considered decision regarding the next steps. The Design Thinking approach enables the evaluation of the idea's readiness, as it has also passed through all Design Thinking phases within the Innovation Funnel. Therefore, three possible scenarios emerge.

The idea is successfully validated. Users recognize its value and confirm that it addresses their problem. The organization then begins full-scale development to transform it into a real product [9].

The idea is rejected. The problem it addresses may not be important enough for users to be willing to pay for a solution. And this is an acceptable conclusion, as it is better to identify this early than to invest time and resources in marketing and selling a solution that users do not want [9].

The idea shows potential, but requires reshaping. And this is a common outcome in the innovation process, where Brown and Liedtka strongly recommend another iteration. By repeating the phases, the organization enters iterative incremental innovation: a process in which each iteration adds value for the user, until the solution fully addresses their problem [14,19].

5 Discussion

The integration of Design Thinking into the Innovation Funnel introduces a shift in how the organization approaches innovation, moving toward a more user - centered perspective [14]. In line with Proposition 2, this integration emphasizes a deeper understanding of users, enabling a more accurate problem definition in the early phases of the Funnel. Consequently, it reduces the risk of developing non-viable ideas, establishing a strong foundation for successful innovation [9,11]. In this context, Mihić et al. [21] emphasize the importance of effective risk management in innovation projects, particularly highlighting early recognition of risk factors, while Stošić et al. [22] highlight that systematic risk identification is a foundational step in improving project performance.

The findings also support Proposition 1, that ineffective innovation management originates from insufficient problem validation and premature commitment to ideas. Design Thinking mitigates these risks through iterative validation and continuous user feedback. Bjelica et al. [23] demonstrate that clearly defined success factors and evaluation criteria are strongly associated with project success, while the same authors [24] further highlight the importance of user appreciation success criteria, thereby suggesting that inadequate problem definition may negatively affect project outcomes.

Furthermore, the integration also enables more efficient resource management. It allows ideas to be evaluated with minimal resource use, without committing to full - scale development only to discover later that they are not viable. Through an iterative approach, organizations can refine ideas and improve their quality before entering later stages of the Innovation Funnel [9,11].

But the integration also brings certain limitations. It requires additional time and careful execution in each phase, which can be particularly challenging in startup environments that prioritize rapid growth and quick scaling [12]. Furthermore, its successful application depends on organizational culture, as teams need to adapt a more open mindset in order to fully integrate Design Thinking [14]. In addition, the proposed framework is conceptually grounded and has not yet been empirically validated. Its applicability may depend on organizational maturity and types of innovations.

Therefore, future research should focus on empirical testing of the framework in order to further evaluate its practical applicability in the innovation process.

Overall, the main contribution of this approach lies in improving innovation performance by enhancing product discovery and aligning solutions more closely with user needs. By addressing key causes of failure identified in the early stages of the Innovation Funnel, the proposed integration provides a structured approach to reducing innovation risk and increasing the likelihood of successful market adoption.

6 Conclusion

In a world where progress relies on change, innovation remains essential. However, innovation success depends not only on generating ideas, but on understanding user needs and managing the innovation process successfully.

The main contribution of this paper is the development of an innovation framework based on the integration of Design Thinking, the Innovation Funnel, and Diffusion of Innovations into a unified approach. The proposed framework theoretically supports organizations throughout the innovation process. It emphasizes that innovation does not need to start from grand ideas, but from small insights and empathy. In the end, it does not guarantee perfect solutions, but it may improve the chances of finding the right ones.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article. AI language assistance tools were used during the preparation of this manuscript for translation and grammar editing purposes. The authors take full responsibility for the content of the paper.

References

1. Milutinović, R., Stošić, B., Mihić, M.: Concepts and Importance of Strategic Innovation in SMEs: Evidence from Serbia. Management XX(77), 35–42 (2015). https://doi.org/10.7595/management.fon.2015.0030
2. Castellion, G., Markham, S.K.: Perspective: New product failure rates: Influence of argumentum ad populum and self-interest. Journal of Product Innovation Management 30(5), 976–979 (2013)
3. Trott, P.: Innovation Management and New Product Development. 3rd edn. Pearson Education Limited, Harlow (2005)
4. Tidd, J., Bessant, J.: Managing Innovation: Integrating Technological, Market and Organizational Change. 6th edn. Wiley (2018)
5. Birss, D.: How to Get to Great Ideas: A System for Smart, Extraordinary Thinking. Nicholas Brealey Publishing, London (2018)
6. Ulwick, A. W.: What Customers Want: Using Outcome-Driven Innovation to Create Breakthrough Products and Services. McGraw-Hill, New York (2005)
7. Thomke, S., von Hippel, E.: Customers as Innovators: A New Way to Create Value. Harvard Business Review 80(4), 74–81 (2002)
8. Ries, E.: The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses. Crown Business, New York (2011).
9. Cooper, R. G.: Winning at New Products: Creating Value Through Innovation. Basic Books, New York (2019)

10. McKinsey & Company: McKinsey on Finance: 20th Anniversary Edition. McKinsey Global Publications (2022). Available at: https://www.mckinsey.com
11. Rogers, E.M.: Diffusion of Innovations. 5th edn. Free Press, New York (2003)
12. Moore, G.A.: Crossing the Chasm: Marketing and Selling High-Tech Products to Mainstream Customers. Rev. edn. HarperBusiness, New York (2001)
13. Stanimirović, P., Mihić, M., Mitrović, Z.: Operationalizing Functional Analysis in Public Administration: A Diagnostic Framework for Enhancing System Efficiency in Eastern European Parliamentary Secretariats. Administrative Sciences 16(3), 142 (2026). https://doi.org/10.3390/admsci16030142
14. Brown, T.: Design thinking. Harvard Business Review 86(6), 84–92 (2008)
15. Curedale, R.A.: Design Thinking: Processes & Methods Manual. Design Community College Inc., Topanga, CA (2016)
16. Scrum.org: What are some differences among simple, complicated, and complex problems and work. Available at: https://www.scrum.org/resources/what-are-some-differences-among-simple-complicated-and-complex-problems-and-work. Accessed: Mar. 11, 2026
17. Carlgren, L., Rauth, I., Elmquist, M.: Framing Design Thinking: The concept in idea and enactment. Creativity and Innovation Management 25(1), 38–57 (2016)
18. Djordjevic, A., Mihić, M.: Strategic Importance and Sustainable Governance of High-Tech Business Incubators: Evidence from Serbia. In: Mitrović, N., Mladenović, G., Mitrović, A. (eds.) Current Problems in Experimental and Computational Engineering. CNNTech 2021. Lecture Notes in Networks and Systems 323, 199–221 (2022). https://doi.org/10.1007/978-3-030-86009-7_11
19. Liedtka, J.: Why Design Thinking Works. Harvard Business Review Press, Boston (2018)
20. Hasso Plattner Institute of Design at Stanford: An Introduction to Design Thinking Process Guide. Available at: https://web.stanford.edu/~mshanks/MichaelShanks/files/509554.pdf. Accessed: Apr. 3, 2026
21. Mihić, M., Dodevska, Z., Todorović, M., Obradović, V., Petrović, D.: Reducing risks in energy innovation projects: complexity theory perspective. Sustainability 10(9), 2968 (2018). https://doi.org/10.3390/su10092968
22. Stošić, B., Mihić, M., Milutinović, R., Isljamović, S.: Risk identification in product innovation projects: new perspectives and lessons learned. Technology Analysis & Strategic Management 29(2), 133–148 (2017). https://doi.org/10.1080/09537325.2016.1210121
23. Bjelica, D., Mihić, M., Kavčič, K., Gošnik, D.: Relationship between project success factors, project success criteria and project success in SME: evidence from selected European transitional economies. International Journal of Industrial Engineering and Management 14(4), 297–310 (2023). https://doi.org/10.24867/IJIEM-2023-4-340
24. Bjelica, D., Mihić, M., Gošnik, D.: Managing projects in large companies – project success factors in the crisis and post-crisis period: evidence from Serbia and Slovenia. International Journal of Industrial Engineering and Management 17(1), 79–92 (2026). https://doi.org/10.24867/IJIEM-404

Original Research Article *https://doi.org/10.5281/zenodo.23477004*

The Influence of Experience and Age on Perceptions of Managerial Factors in Theatre Projects

Dragana Pantić[1][0009-0000-8965-5254] and Dragan Bjelica[1][0000-0003-0203-2877]

[1]University of Belgrade – Faculty of Organizational Sciences, Jove Ilica 154, Belgrade, Serbia
dragana.pantic@fon.bg.ac.rs, dragan.bjelica@fon.bg.ac.rs

Abstract. Previous studies have investigated how theatre members perceive management issues, but less attention has been paid to project management-specific tools and techniques. The knowledge gap highlights the need to examine whether project management practices can contribute to overcoming the challenges of the performing arts sector. This study examined the perceptions of theatre project participants. The data were analyzed using the ANOVA test and t-test. The results show a statistically significant difference in attitudes toward team coordination and communication among different age groups, with older respondents (46 + years) expressing a higher level of agreement regarding the various aspects of team collaboration compared to younger participants (18–25 years). These findings indicate the potential of project-oriented approaches as a relevant framework for improving organizational and production processes in contemporary theatre.

Keywords: Theatre, Experience, Project Management.

1 Introduction

A theatre project is a temporary, project-based process for producing a theatrical performance that includes pre-production, formation of the production team, rehearsals, technical implementation, marketing activities, and the premiere performance, with a fixed final deadline (Lindgren and Packendorff, 2007). Due to the complexity and interdependence of components, theatre projects represent an inspiring and important topic of research (Schreyögg and Höpfl, 2004; Cwikla and Jalocha, 2015).

For example, the production of an opera performance, the organization of a theatre festival, or the international co-production can be considered typical examples of theatre projects, as they involve defined budgets, fixed timelines, multidisciplinary teams, and clearly established artistic and organizational objectives.

Existing research indicates that artistic activities are increasingly being integrated into project-based organizational models (Bauer, Viola, and Strauss, 2011), while several authors emphasize the growing managerial complexity of cultural production (Lehner, 2009; Lindgren and Packendorff, 2007). However, previous studies have primarily focused on institutional management, cultural policy, audience development, or the economic sustainability of the arts, while the perceptions of theatre practitioners

regarding the practical application of project management principles have remained insufficiently explored.

This research gap is particularly significant because theatre projects involve a unique combination of creative autonomy and managerial constraints. Theatre production requires not only artistic innovation, but also the coordination of financial, technological, temporal, and interpersonal resources under uncertain and often resource-constrained conditions. Consequently, balancing artistic objectives with economic and organizational demands becomes a central challenge in theatre project management, creating both tensions and opportunities for improving the efficiency of non-profit cultural institutions (Røyseng, 2008; Lee, Sejas-Portillo, and Fraser, 2023; Bennett, 2009; Donelli, Fanelli, and Zangrandi, 2021).

Therefore, this study aims to explore how participants in theatre projects perceive key dimensions of project-based management and whether these perceptions differ according to professional experience. By focusing on practitioners from the Serbian theatre context, this study contributes to the growing discussion on the projectification of the cultural sector and provides empirical insight into the applicability of project management principles in performing arts practice.

2 Literature Review

Although project management encompasses a broad range of managerial dimensions, this study focuses on three key dimensions considered particularly relevant in the context of theatre production: team coordination and communication, financial governance capability, and temporal resource management. These dimensions were selected because they reflect the structural characteristics of theatre projects as temporary, collaborative, and resource-constrained undertakings.

Team coordination and communication represent a central dimension due to the multidisciplinary and highly collaborative nature of theatre production, where successful implementation depends on the interaction of diverse artistic and managerial actors (Boerner and Gebert, 2012; Abfalter, 2013). Financial governance capability was included because theatre projects operate within fixed or limited budgets, often in uncertain funding environments that require continuous resource allocation and financial decision-making (Wu, Jiang, and Yuan, 2020; Zelenskaya and Shakina, 2023). Temporal resource management was selected because theatre projects are inherently driven by clearly defined deadlines, with premiere performances and production schedules creating strong temporal constraints that directly shape project execution (Lindgren and Packendorff, 2007; Gevers et al., 2009). Together, these dimensions capture key managerial challenges related to theatre project environments.

2.1 The Role of Team Coordination and Communication in Project Success

Based on previous studies (Lindgren and Packendorff 2007; Häti-Korkeila and Gröndahl, 2018), team roles, including those of directors, producers, performers,

dramaturges, and managers, as well as interactions between them, have been identified as essential for the implementation of theatre projects.

One fundamental challenge in collaboration relates to the handling of disagreements and differences in perspectives. In a creative environment such as the theatre, varying viewpoints can be both an asset and a catalyst for conflict. What is important for the team to recognize is that a well-developed culture of tolerance, transparent communication, determination, and empathy can contribute to constructive conflict resolution (Afzalur Rahim, 2002; Stingl et al., 2025).

Previous studies (Boerner and Gebert, 2012; Abfalter, 2013; Dunham and Freeman, 2000; Altaher et al., 2024) have found a link between the role of leadership, which was evaluated by participants as highly significant, and project success. In a study investigating project success factors in small and midsize enterprises, Bjelica et al. (2023) reported that project management skills, such as emotional intelligence and leadership skills, have a considerable impact on project success. In view of all that has been mentioned so far, one may suppose that this component functions as a common denominator in both artistic and business-related projects. When discussing leadership in performing arts projects, it is crucial to highlight differences from conventional business settings. Defining characteristics of art project leaders include precarious, project-based employment, short career spans, the central focus on networking and social capital, as well as exceptionally high levels of expertise and intrinsic motivation (Abfalter, 2013).

Teams whose cultures foster the capacity for creative problem-solving can better confront numerous obstacles and uncertainties (Slavich and Montanari, 2009). This ability, stemming from several indirect factors, such as relational coordination, experiential learning, and access to knowledge, has a direct impact on project outcomes (Carmeli, Levi, and Peccei, 2021).

2.2 Financial Governance Capability

The findings of the studies (Wu, Jiang, and Yuan, 2020; Shafranskaya, Potapov, and Ozhegov, 2021) indicate that different types of theatre (e.g., opera, ballet, and children's theatre) and the ability to recruit appropriate target groups make a difference in the efficiency of theatres. This results in the need for project budgeting to be aligned with the applicable mission of the theatre. Cost-effectiveness can be built around strategies related to size, as larger classical theatres might benefit from so-called economies of scale – the more often a production runs, the cheaper the average cost per production is. On the other hand, regional theatres need to centre their financial planning on strategies such as expanding their repertoire, establishing pricing models tailored to different audience segments, developing synergies with tourism, and enhancing their relationships with the community. Thus, managing budgets in theatre projects involves a process in which one needs to balance among the demands of institutional model types, artistic needs, geographical context, and audience profiles in a manner that accommodates cultural and economic objectives (Zelenskaya and Shakina, 2023). According to Melnichuk (2015), the quality of the budgeting has a significant effect on managing finances, ensuring an effective response to emerging trends and changes. Additionally, budgeting cannot be separated from overall organizational planning, as it involves the

estimation of necessary financial resources for project implementation and their allocation (Herath and Chong, 2021).

In studies investigating the financing of theatre sectors in Poland (Gałecka and Smolny, 2019), Belgium (Werck, Stultjes, and Heyndels, 2008), Italy (Fanelli et al., 2019; Castiglione, Infante, and Zieba, 2024; Slavich and Montanari, 2009), and Germany (Neligan, 2006), authors report that government subsidies, whether central or local, are among the most important sources of funding. Under certain circumstances, subsidizing can indeed be a major pillar of financial support, especially for non-commercial theatre projects. Much of the criticism that financing through direct budget subsidies has attracted relates to the optimal allocation of these donated funds, and the lack of coherence between legislation, principles of self-governance, and financial autonomy (Gałecka and Smolny, 2019). Distancing itself from the former all-prevalent state-supported funding model, China has experienced a significant transformation to more market-driven financial models since 2003. The latter is characterized by the corporatization of public performing arts institutions and the unprecedented growth of private enterprises. Consequently, ticket sales have become the main source of revenue, now contributing to around 60% of total income, with public subsidies remaining at only around 22% (Wu, Jiang, and Yuan, 2020).

2.3 Temporal Resource Management

The effective functioning of a theatre largely depends on the establishment of and consistent adherence to time schedules. Due to the unique nature of theatre projects, a systematic approach to managing their deadlines is necessary. A qualitative study by Lindgren and Packendorff (2007) described the presence of temporal inevitability, manifested through fixed project completion deadlines, most commonly the premiere performance. Audiences organize their free time and allocate funds to purchase tickets, which places the timely completion of a project within a wider context of professional behavior. Moreover, respondents recognized deadlines as a vital motivator, acknowledging that without definitive temporal limits, project work could potentially be prolonged unnecessarily and lead to disappointment among participants (Lindgren and Packendorff, 2007).

Although the specifics of defining timelines are important, the nature of the theatre production process requires a certain degree of flexibility due to unpredictable and frequent changes in both technical and creative matters. Teams whose members do not agree on time-related matters may find it more challenging to meet deadlines, as it becomes increasingly difficult to achieve a shared understanding of time constraints (Gevers et al., 2009). According to Ariely and Wertenbroch (2002), individuals who acknowledge their propensity to procrastinate often impose deadlines on themselves as a form of self-control. However, authors have also shown that self-imposed deadlines are generally less effective than evenly spaced, externally imposed deadlines (Ariely and Wertenbroch, 2002). This is why the best approach to deadlines in a team setting may be hybrid, combining structure and accountability via external imposition (such as producers or managers) and internal self-management by team members.

3 Method

The quantitative research method was used in order to gain insights into key components of effective management, such as enhancing teamwork, managing time effectively, optimizing financial planning, and applying contemporary project management tools and techniques in the theatre industry. Data was collected through a survey questionnaire. Table 1 presents an overview of the variables identified through the literature review. Each component was examined by asking respondents to assess the presented sub-components on a 5-point Likert scale ranging from 1 (strongly disagree/very low) to 5 (strongly agree/very high).

Table 1. Summary of research variables identified through literature review.

ID	Description	Literature
	Team Coordination and Communication Excellence (T)	
T1	Even and clear distribution of tasks	(Černevičiūtė and Strazdas, 2018; Erez, Lepine, and Elms, 2002)
T2	Team meetings, interaction dynamics, and feedback	(Gaunt and Treacy, 2020; Gibb, 2004; Kauffeld and Lehmann-Willenbrock, 2012)
T3	Transparent communication	(Salas, Sims, and Burke, 2005; De Dreu and Van Vianen, 2001)
T4	Utilization of digital platforms	(Anders, 2016; Pollack and Matous, 2019)
T5	Constructive conflict resolution	(De Dreu and Van Vianen, 2001; Järvinen, Ansio, and Houni, 2015)
	Financial Governance Capability (F)	
F1	Budget planning	(Werck, Stultjes, and Heyndels, 2008; Crawford, 2005)
F2	Budget monitoring	(Mokre, 1996)
F3	Financial reporting	(DeFond and Zhang, 2014)
F4	Differentiation of funding sources	(Amans, Mazars-Chapelon, and Villesèque-Dubus, 2015; Smith, 2007)

ID	Description	Literature
F5	Cost control mechanisms	(Amans, Mazars-Chapelon, and Villesèque-Dubus, 2015)
	Time and Priority Management (TP)	
TP1	Definition of precise timeframes	(Gevers, van Eerde, and Rutte, 2009; Crawford, 2005)
TP2	Completion of project phases on time	(Jalali Sohi, Bosch-Rekveldt, and Hertogh, 2019; Ariely and Wertenbroch, 2002)
TP3	Task prioritization	(Komander and König, 2024)
TP4	Flexibility and adaptability to changes	(Jalali Sohi, Bosch-Rekveldt, and Hertogh 2019; Lampel, Lant, and Shamsie 2000)
TP5	Meeting final deadlines	(Claessens et al. 2007; Ariely and Wertenbroch, 2002)

3.1 Sample Characteristics

The research was conducted in the context of theatres located in cities with active theatrical production in the Republic of Serbia, with a focus on professionals involved in theatre production processes. Although approximately 80 producers were invited to participate in the survey, the response rate was around 56.25%.

The final sample (Table 2) consisted of 45 respondents – theatre producers, of whom the majority identified as female (73.3%). The 18–25 years age group was the most represented (44.4%), while respondents aged 26–35 and 46+ were equally represented (24.4% each), and the least represented group was that of 36–45 years (6.7%). Most respondents were university graduates (64.4%), while 20% had a master's degree, 8.9% had a high school education, and 6.7% held a doctoral degree. Most respondents came from Belgrade (68.9%). The majority of respondents (60%) had worked on 1-9 theatre projects, 22.2% had worked on more than 30 projects, and a smaller proportion had worked on 10 to 29 projects.

Table 2. Demographic characteristics.

Characteristic	**N (%)**
Sex	
Female	33 (73.3)
Male	12 (26.7)
Age	
18-25	20 (44.4)
26-35	11 (24.4)
36-45	3 (6.7)

Characteristic	N (%)
46+	11 (24.4)
Education	
Secondary School	4 (8.9)
Bachelor	29 (64.4)
Post-graduate education (Master)	9 (20)
Post-graduate education (PhD)	3 (6.7)
Number of Projects	
1-9	27 (60.0)
10-19	4 (8.9)
20-29	4 (8.9)
30+	10 (22.2)

4 Results

Most respondents (42.2%) believe that team tasks are well distributed among team members. This result provides the highest percentage of agreement compared to other aspects of collaboration. Viewing the data in a more detailed way, a significant portion of respondents (22.2%) strongly agree with this statement, while a smaller percentage (6.7%) expresses strong disagreement. These differences in perception indicate that participants are often, but not always clear about particularities in the distribution of tasks, which may reflect different experiences and practices in team coordination. Regarding the regularity of team meetings, 51.1% of respondents agree or strongly agree with the statement that meetings are held regularly, while 24.4% are neutral. The result indicates that although the majority believe that meetings are frequent, there is a significant portion of respondents who do not feel completely confident about the regularity of meetings. When it comes to sharing information relevant to the project, 62.2% of respondents expressed a high degree of agreement with this statement. However, it is important to note that 24.4% of respondents expressed disagreement or total disagreement, which indicates potential problems in communication within teams. The percentage of respondents who believe that digital platforms are used effectively to coordinate activities is 55.5%, while 31.1% of respondents believe that this aspect is not implemented well enough, which indicates possible shortcomings in the use of digital tools or the need for additional education and training of team members to ensure that digital platforms are optimally integrated into work processes. The fact that constructive conflict resolution is important for maintaining productive team dynamics is indicated by the views of 28.9% of respondents who expressed agreement and 20% who completely agree with the statement that conflicts in their teams are resolved constructively. However, a significant percentage of respondents (26.7%) remained neutral. Overall, the findings reveal strengths in task distribution and information sharing but also identify areas for improvement, particularly in the standardization of meetings, the optimization of digital platform usage, and the enhancement of conflict resolution strategies.

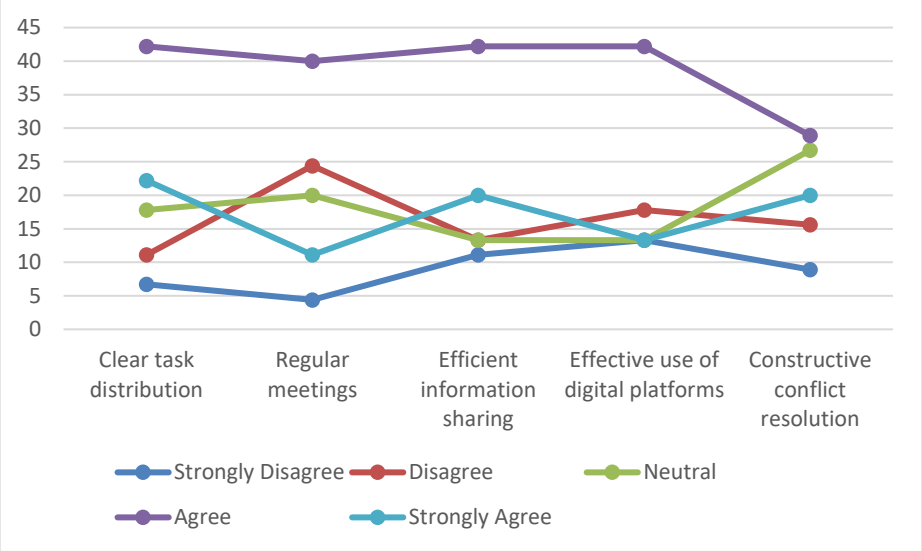


Fig. 1. Perceived importance of teamwork in theatre projects

A positive attitude towards detailed budget planning before the start of the project implementation was expressed by 40% of respondents (Figure 2), of which 20% strongly agreed with this attitude. On the other hand, 24.4% of respondents disagreed with this statement, while 24.4% remained neutral. This result indicates a mixed perception of budget planning, which may also suggest possible gaps in the preparation process. The adoption of modern budgeting tools could further improve resource management. According to our survey, 33.3% of respondents agreed, and 17.8% strongly agreed that costs are regularly monitored throughout the entire process. However, 42.2% of respondents remained neutral, suggesting some uncertainty about the importance or quality of existing financial reports. Awareness of the importance of preparing financial reports was expressed by 15.6% of respondents who agreed and 13.3% who strongly agreed that accurate reports are regularly prepared for all phases of theatre projects. However, 42.2% of respondents remained neutral. Exploring various funding sources, such as grants, donations, and ticket sales, is crucial for project sustainability. However, a high percentage of neutral responses (33.3%) indicates a lack of initiative or knowledge about these funding options.

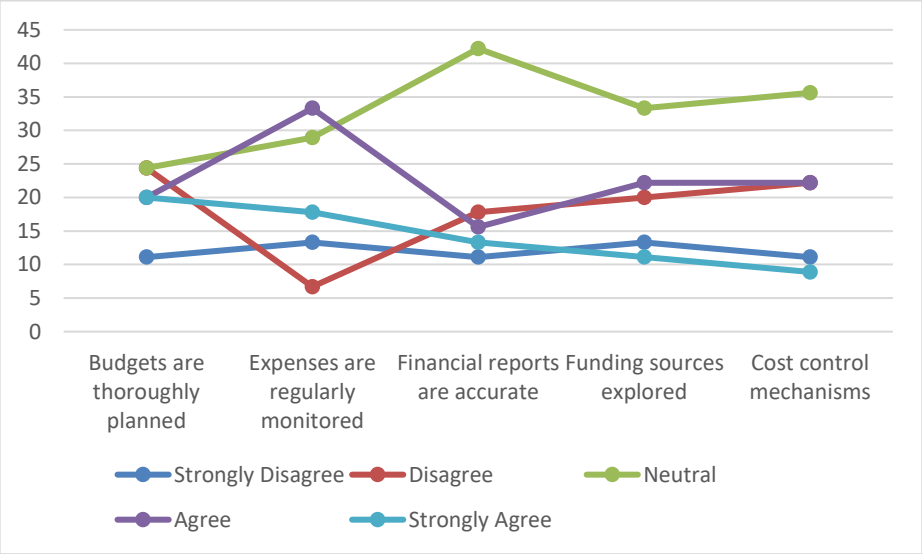


Fig. 2. Respondents' perspectives on financial management

The analysis of time management in theatre projects, based on the results of our research (Figure 3), reveals several important aspects. The majority, 37.8% of respondents, believe that time frames are carefully planned, which emphasizes the importance of detailed planning for successful project management. The fact that 22.2% of respondents expressed a neutral position on this issue may be worth considering, as it may invite stakeholders to improve the process of planning deadlines at each phase of the project. Given the importance of monitoring progress throughout projects, we wanted to find out the views of the respondents, of whom 33.3% expressed a neutral view, and 28.9% agreed that progress is monitored following the planned deadlines. Respondents (37.8%) feel that tasks in theatre projects are well prioritized, while 22.2% of respondents express a neutral view. Finally, the flexibility and adaptability of project teams to a dynamic environment are also present, as confirmed by the positive responses of the majority of participants, with 46.7% agreeing and 28.9% strongly agreeing with the given statement.

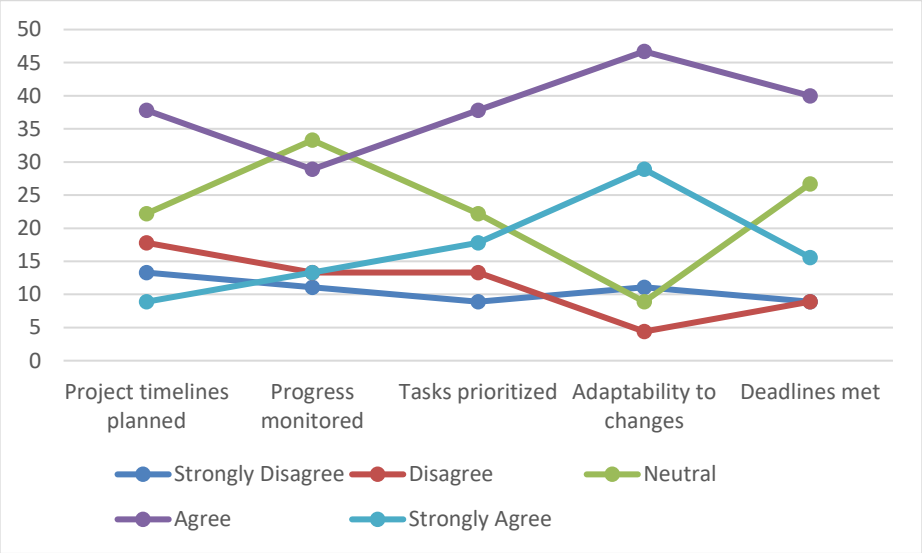


Fig. 3. Perceptions of respondents regarding time management

4.1 Age and Expertise in Theatre Project Perception

A one-way analysis of variance (ANOVA) was conducted to assess the differences in respondents' attitudes across the age categories, related to the various aspects of the theatre project management (Table 3). The aspects investigated included teamwork, financial management, and time management. Older respondents also reported participation in a greater number of theatre projects, suggesting a meaningful relationship between age and professional experience in this sample. Although age was used as the primary grouping variable, it can also be interpreted as a proxy for professional experience.

When it comes to attitudes regarding team collaboration across different age groups, the results show a statistically significant difference ($F = 3.185$, $p = .034$). Specifically, the youngest group of respondents (18–25 years) provides a mean value of 3.30 ($SD = 0.91$), while the oldest group (46+ years) shows the highest mean value of 4.00 ($SD = 0.83$). A possible interpretation is that older respondents consider teamwork more important within theatre projects, which comes from their longer professional experience and a stronger emphasis on efficiency through collaboration. On the other hand, respondents in the 36–45 age group have a significantly lower mean score ($M = 2.20$, $SD = 1.06$), which might suggest a certain reservation or different priorities regarding team collaboration.

Regarding financial management, no statistically significant differences were found between age categories ($F = 1.976$, $p = .133$). However, it is important to mention that the older cohort (46+ years) has the highest mean value ($M = 3.58$, $SD = 0.83$), implying that older participants give more importance to the need for financial management. This trend can be linked to greater awareness and experience in financial

responsibilities and resource management, which are crucial to project sustainability and success. As for time management, the results do not show statistically significant differences between age groups either (F = 2.015, p = .127). Once again, the oldest group (46+ years) demonstrates the highest mean value (M = 3.56, SD = 1.05), while the lowest mean is observed among respondents aged 36–45 (M = 2.00, SD = 1.40). Based on these results, we can assume that older respondents may be more effective in planning and time management, likely due to accumulated experience and awareness of the importance of time as a project resource.

Table 3. Differences in attitudes toward theatre project management among different age groups

		N	**M ± SD**	***F***	***Sig.***
Teamwork	18-25	20	3.30 ±.91	3.185	.034
	26-35	11	3.29±1.09		
	36-45	3	2.20±1.06		
	46+	11	4.00±0.83		
Financial Management	18-25	20	3.07±1.00	1.976	.133
	26-35	11	2.91±1.22		
	36-45	3	2.07±1.01		
	46+	11	3.58 ±.83		
Time Management	18-25	20	3.47 ±.82	2.015	.127
	26-35	11	3.45±1.23		
	36-45	3	2.00±1.40		
	46+	11	3.56±1.05		

5 Discussion

The findings of this study suggest that theatre practitioners perceive project management dimensions as relevant components of contemporary theatre production. Rather than viewing theatre production exclusively as an artistic process, respondents recognized the importance of organizational and managerial structures in ensuring the successful implementation of theatre projects. These findings support previous studies emphasizing the increasing projectification of cultural production and the growing managerial complexity of performing arts organizations (Lindgren and Packendorff, 2007; Bauer, Viola, and Strauss, 2011).

Among the examined dimensions, teamwork emerged as the most prominent and consistently valued aspect of project management in theatre. Positive perceptions regarding task distribution, communication, and collaborative dynamics indicate that respondents recognize coordination as a necessary condition for managing multidisciplinary creative processes. This finding is consistent with studies highlighting the importance of relational coordination, leadership, and communication quality in artistic teams (Boerner and Gebert, 2012; Abfalter, 2013). At the same time, the presence of neutral and negative responses regarding meetings, conflict resolution, and digital coordination tools suggests that collaboration within theatre projects remains uneven and highly dependent on organizational practices and interpersonal dynamics. In this sense,

the study indicates that teamwork in theatre cannot be understood solely as spontaneous artistic collaboration, but also as a structured managerial process requiring clear communication mechanisms and coordination strategies.

One of the central contributions of this study relates to the role of professional experience and age in shaping managerial perceptions. The results demonstrated statistically significant differences in perceptions of teamwork across age groups, with older respondents assigning greater importance to collaboration and coordination within theatre projects. Since older participants also reported involvement in a larger number of productions, the findings suggest that accumulated professional experience may increase awareness of the importance of collective work processes in theatre environments. These findings partially align with Ojiako et al. (2014), who argue that individual and professional characteristics can influence perceptions of project management practices.

From a theoretical perspective, this study contributes to the limited body of research connecting project management theory and theatre production practice. By examining managerial perceptions within performing arts environments, the study extends discussions on projectification beyond traditional business and organizational contexts. From a practical perspective, the findings suggest that theatre organizations may benefit from strengthening communication structures, improving access to financial knowledge, and developing more consistent coordination mechanisms adapted to collaborative creative work.

6 Conclusion

This study aimed to examine how participants in theatre projects perceive key dimensions of project-based management and whether these perceptions differ according to professional experience. By focusing on teamwork, financial management, and time management within the Serbian theatre context, the research sought to explore the relevance of project management principles in contemporary theatre production.

The findings indicate that respondents perceive all analyzed dimensions as important components of successful theatre project implementation. Among these dimensions, teamwork emerged as the most prominent aspect, particularly among older and more experienced respondents, who assigned greater importance to collaboration and coordination within theatre projects. These results suggest that professional experience influences perceptions of managerial factors, especially those related to collective work processes and team dynamics. In contrast, perceptions regarding financial and time management appeared to be more consistent across different age and experience groups.

The study contributes to the limited body of research connecting project management and theatre practice by providing empirical insight into how theatre practitioners perceive managerial processes in project-based cultural production. The findings further support the view that theatre production increasingly requires not only artistic competencies, but also organizational and managerial capacities adapted to collaborative and resource-constrained working environments.

Finally, several limitations of this study should be acknowledged. The present study focuses on managing finances, time, and teamwork in theatre projects. It is important to emphasize that, to understand theatre projects well, one must investigate the problem of management with regard to multiple contexts. Performance quality, artistic excellence, creativity, and other factors such as audience interaction may be some of the areas requiring detailed research if the broader goal is to provide a comprehensive overview of the components that contribute to successful processes in theatre.

The limitations of this study include running the study in a specific geographic region, which may limit the generalizability of the findings. A longitudinal study, involving a more diverse sample with wider geographic coverage, is needed in order to gain a comprehensive understanding of this problem. Furthermore, there is no empirical evidence that participants actively implement project management tools and techniques in the theatre industry. To support such claims, further research needs to include active engagement or case studies of specific theatre productions or festivals, examined as potential representative forms of project-based initiatives.

Acknowledgments. The authors acknowledge the use of AI-assisted language tools for stylistic editing and language refinement, given that English is not the authors' native language.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Abfalter, D.: Authenticity and respect: Leading creative teams in the performing arts. Creativity and Innovation Management 22(3), 295–306 (2013). https://doi.org/10.1111/caim.12004
2. Altaher, A.M., AL-Obaidly, G.A., Al-badaineh, G., Kashoob, S.S.M., Hammouri, Q., Khasawneh, M.A.S.: The impact of leadership styles on project success: The mediating role of team collaboration. Journal of Project Management 9(4), 485–492 (2024). https://doi.org/10.5267/j.jpm.2024.7.001
3. Amans, P., Mazars-Chapelon, A., Villesèque-Dubus, F.: Budgeting in institutional complexity: The case of performing arts organizations. Management Accounting Research 27, 47–66 (2015). https://doi.org/10.1016/j.mar.2015.03.001
4. Anders, A.: Team communication platforms and emergent social collaboration practices. International Journal of Business Communication 53(2), 224–261 (2016). https://doi.org/10.1177/2329488415627273
5. Ariely, D., Wertenbroch, K.: Procrastination, deadlines, and performance: Self-control by precommitment. Psychological Science 13(3), 219–224 (2002). https://doi.org/10.1111/1467-9280.00441
6. Bauer, C., Viola, K., Strauss, C.: Management skills for artists: Learning by doing? International Journal of Cultural Policy 17(5), 626–644 (2011). https://doi.org/10.1080/10286632.2010.531716
7. Bennett, R.: Perceived importance of performance management metrics among UK theatre companies. International Journal of Productivity and Performance Management 58(7), 670–693 (2009). https://doi.org/10.1108/17410400910989476

8. Bjelica, D. LJ., Mihić, M., Kavčič, K., Gošnik, D.: Relationship between project success factors, project success criteria and project success in SMEs: Evidence from selected European transitional economies. International Journal of Industrial Engineering and Management 14(4), 297–310 (2023). https://doi.org/10.24867/IJIEM-2023-4-340
9. Bilton, C., Leary, R.: What can managers do for creativity? Brokering creativity in the creative industries. International Journal of Cultural Policy 8(1), 49–64 (2002). https://doi.org/10.1080/10286630290032431
10. Boerner, S., Gebert, D.: Fostering artistic ensemble performance: Exploring the role of transformational leadership. Nonprofit Management and Leadership 22(3), 347–365 (2012). https://doi.org/10.1002/nml.20058
11. Carmeli, A., Levi, A., Peccei, R.: Resilience and creative problem-solving capacities in project teams: A relational view. International Journal of Project Management 39(5), 546–556 (2021). https://doi.org/10.1016/j.ijproman.2021.03.007
12. Castiglione, C., Infante, D., Zieba, M.: Is it worth subsidising the cultural sector? New insights from Italian theatre companies. Journal of Policy Modeling 46(1), 20–38 (2024). https://doi.org/10.1016/j.jpolmod.2023.11.006
13. Černevičiūtė, J., Strazdas, R.: Teamwork management in creative industries: Factors influencing productivity. Entrepreneurship and Sustainability Issues 6(2), 503–516 (2018). https://doi.org/10.9770/jesi.2018.6.2(3)
14. Claessens, B.J.C., van Eerde, W., Rutte, C.G., Roe, R.A.: A review of the time management literature. Personnel Review 36(2), 255–276 (2007). https://doi.org/10.1108/00483480710726136
15. Crawford, L.: Senior management perceptions of project management competence. International Journal of Project Management 23(1), 7–16 (2005). https://doi.org/10.1016/j.ijproman.2004.06.005
16. Cwikla, M., Jalocha, B.: Unspread wings. International Journal of Managing Projects in Business 8(4), 626–648 (2015). https://doi.org/10.1108/IJMPB-11-2014-0078
17. DeFond, M., Zhang, J.: A review of archival auditing research. Journal of Accounting and Economics 58(2–3), 275–326 (2014). https://doi.org/10.1016/j.jacceco.2014.09.002
18. Donelli, C.C., Fanelli, S., Zangrandi, A.: Inside and outside the boardroom: Collaborative practices in the performing arts sector. International Journal of Arts Management 24(1), 48–62 (2021).
19. De Dreu, C.K.W., Van Vianen, A.E.M.: Managing relationship conflict and the effectiveness of organizational teams. Journal of Organizational Behavior 22(3), 309–328 (2001). https://doi.org/10.1002/job.71
20. Dunham, L., Freeman, R.E.: There is business like show business. Organizational Dynamics 29(2), 108–122 (2000). https://doi.org/10.1016/S0090-2616(00)00018-8
21. Edelman, J., Šorli, M.: Measuring the value of theatre for Tyneside audiences. Cultural Trends 24(3), 232–244 (2015). https://doi.org/10.1080/09548963.2015.1066074
22. Erez, A., Lepine, J.A., Elms, H.: Effects of rotated leadership and peer evaluation on the functioning and effectiveness of self-managed teams: A quasi-experiment. Personnel Psychology 55(4), 929–948 (2002). https://doi.org/10.1111/j.1744-6570.2002.tb00135.x
23. Fanelli, S., Donelli, C.C., Zangrandi, A., Mozzoni, I.: Balancing artistic and financial performance: Is collaborative governance the answer? International Journal of Public Sector Management 33(1), 78–93 (2019). https://doi.org/10.1108/IJPSM-05-2019-0138
24. Gałecka, M., Smolny, K.: Criteria for the optimal financing model of public theatres. Review of Economic Perspectives 19(2), 119–136 (2019). https://doi.org/10.2478/revecp-2019-0007

25. Gaunt, H., Treacy, D.S.: Ensemble practices in the arts: A reflective matrix to enhance team work and collaborative learning in higher education. Arts and Humanities in Higher Education 19(4), 419–444 (2020). https://doi.org/10.1177/1474022219885791
26. Gevers, J.M.P., van Eerde, W., Rutte, C.G.: Team self-regulation and meeting deadlines in project teams: Antecedents and effects of temporal consensus. European Journal of Work and Organizational Psychology 18(3), 295–321 (2009). https://doi.org/10.1080/13594320701693217
27. Gibb, S.: Arts-based training in management development: The use of improvisational theatre. Journal of Management Development 23(8), 741–750 (2004). https://doi.org/10.1108/02621710410549594
28. Hamilton, J.R.: The art of theater—A précis. The Journal of Aesthetic Education 43(3), 4–14 (2009). https://doi.org/10.1353/jae.0.0053
29. Häti-Korkeila, M., Gröndahl, L.: From permanent positions to visiting jobs. Nordic Theatre Studies 30(1), 89–114 (2018). https://doi.org/10.7146/nts.v30i1.106925
30. Herath, S., Chong, S.C.: Key components and critical success factors for project management success: A literature review. Operations and Supply Chain Management 14(4), 431–443 (2021). https://doi.org/10.31387/oscm0470314
31. Jalali Sohi, A., Bosch-Rekveldt, M., Hertogh, M.: Does flexibility in project management in early project phases contribute positively to end-project performance? International Journal of Managing Projects in Business 13(4) (2020). https://doi.org/10.1108/IJMPB-07-2019-0173
32. Järvinen, M., Ansio, H., Houni, P.: New variations of dual leadership: Insights from Finnish theatre. International Journal of Arts Management 17(3), 16–27 (2015).
33. Kauffeld, S., Lehmann-Willenbrock, N.: Meetings matter. Small Group Research 43(2), 130–158 (2012). https://doi.org/10.1177/1046496411429599
34. Komander, V., König, A.: Organizations on stage: Organizational research and the performing arts. Management Review Quarterly 74(1), 303–352 (2024). https://doi.org/10.1007/s11301-022-00301-9
35. Lampel, J., Lant, T., Shamsie, J.: Balancing act: Learning from organizing practices in cultural industries. Organization Science 11(3), 263–269 (2000). https://doi.org/10.1287/orsc.11.3.263.12503
36. Lee, B., Sejas-Portillo, R., Fraser, I.: Balancing perspectives on performance: Measurement from the inside and measurement from the outside. Nonprofit Management and Leadership 34(1), 13–34 (2023). https://doi.org/10.1002/nml.21561
37. Lehner, J.M.: The staging model: The contribution of classical theatre directors to project management in development contexts. International Journal of Project Management 27(3), 195–205 (2009). https://doi.org/10.1016/j.ijproman.2008.02.010
38. Lindgren, M., Packendorff, J.: Performing arts and the art of performing: On co-construction of project work and professional identities in theatres. International Journal of Project Management 25(4), 354–364 (2007). https://doi.org/10.1016/j.ijproman.2007.01.005
39. Locatelli, G., Konstantinou, E., Geraldi, J., Sainati, T.: The dark side of projects: Dimensionality, research methods, and agenda. Project Management Journal 53(4), 367–381 (2022). https://doi.org/10.1177/87569728221103911
40. Melnichuk, N.: Budget planning in financial management. Baltic Journal of Economic Studies 1(2), 95–100 (2015). https://doi.org/10.30525/2256-0742/2015-1-2-95-100
41. Mokre, M.: Austrian theatres cost too much! A summary of a research project in Vienna. The European Journal of Cultural Policy 2(2), 289–302 (1996). https://doi.org/10.1080/10286639609358021

42. Neligan, A.: Public funding and repertoire conventionality in the German public theatre sector: An econometric analysis. Applied Economics 38(10), 1111–1121 (2006). https://doi.org/10.1080/00036840500405961
43. Ojiako, U., Chipulu, M., Gardiner, P., Williams, T., Mota, C., Maguire, S., Shou, Y., Stamati, T.: Effect of project role, age and gender differences on the formation and revision of project decision judgements. International Journal of Project Management 32(4), 556–567 (2014). https://doi.org/10.1016/j.ijproman.2013.09.001
44. Pollack, J., Matous, P.: Testing the impact of targeted team building on project team communication using social network analysis. International Journal of Project Management 37(3), 473–484 (2019). https://doi.org/10.1016/j.ijproman.2019.02.005
45. Rahim, M.A.: Toward a theory of managing organizational conflict. International Journal of Conflict Management 13(3), 206–235 (2002). https://doi.org/10.1108/eb022874
46. Røyseng, S.: Arts management and the autonomy of art. International Journal of Cultural Policy 14(1), 37–48 (2008). https://doi.org/10.1080/10286630701856484
47. Salas, E., Sims, D.E., Burke, C.S.: Is there a "Big Five" in teamwork? Small Group Research 36(5), 555–599 (2005). https://doi.org/10.1177/1046496405277134
48. Schreyögg, G., Höpfl, H.: Theatre and organization: Editorial introduction. Organization Studies 25(5), 691–704 (2004). https://doi.org/10.1177/0170840604042410
49. Shafranskaya, I., Potapov, D., Ozhegov, E.: Individual versus companion visitors' willingness-to-pay for theater tickets: The case of Perm Opera. Journal of Philanthropy and Marketing 26(3), e1698 (2021). https://doi.org/10.1002/nvsm.1698
50. Slavich, B., Montanari, F.: New trends of managerial roles in performing arts: Empirical evidence from the Italian context. Cultural Trends 18(3), 227–237 (2009). https://doi.org/10.1080/09548960903065261
51. Smith, T.M.: The impact of government funding on private contributions to nonprofit performing arts organizations. Annals of Public and Cooperative Economics 78(1), 137–160 (2007). https://doi.org/10.1111/j.1467-8292.2007.00329.x
52. Stingl, V., Gilchrist, A., Lawani, A., Flin, R., Zwikael, O.: Thinking is for doing: Project cognition as the foundation of project behaviour. International Journal of Project Management 43(1), 102678 (2025). https://doi.org/10.1016/j.ijproman.2025.102678
53. Walmsley, B.: Co-creating theatre: Authentic engagement or inter-legitimation? Cultural Trends 22(2), 108–118 (2013). https://doi.org/10.1080/09548963.2013.783176
54. Werck, K., Grinwis Plaat Stultjes, M., Heyndels, B.: Budgetary constraints and programmatic choices by Flemish subsidized theatres. Applied Economics 40(18), 2369–2379 (2008). https://doi.org/10.1080/00036840600970153
55. Wu, J., Jiang, K., Yuan, C.: Market-oriented reforms, economic efficiency and commercial performing arts in China. The Journal of Arts Management, Law, and Society 50(1), 33–51 (2020). https://doi.org/10.1080/10632921.2019.1664692
56. Zelenskaya, E., Shakina, E.: The efficiency of the performing arts: Evidence from data envelopment analysis for Russian theaters. The Journal of Arts Management, Law, and Society 53(5), 273–290 (2023). https://doi.org/10.1080/10632921.2023.2273442

Original Research Article *https://doi.org/10.5281/zenodo.23477031*

Toward Integrated IT Resilience: Synergizing Risk Management, Maturity Assessments, and IT Audit

Kristijan Lazić[1] [0009-0007-6243-3310], Dragan Jovičić[2] [0009-0005-1868-9584], Vladan Pantović[3] [0009-0001-6319-8822], Zoran Keković[4] [0000-0003-0971-5371]

[1] Fakultet za projektni i inovacioni menadžment dr Petar Jovanović, Belgrade, Serbia
kristijan.lazic@gmail.com
[2] 2 PHOENIX Pharma SE, Mannheim, Germany
e-mail: dragan.r.jovicic@gmail.com
[3] Fakultet za projektni i inovacioni menadžment dr Petar Jovanović, Belgrade, Serbia
e-mail: vladan@pantovic.rs
[4] Faculty of Security Studies, Gospodara Vučića 50, 11000 Belgrade Serbia
e-mail: zorankekovic@yahoo.com

Abstract. Modern organizations increasingly rely on complex, hybrid IT environments combining on-premises infrastructure, cloud computing, and third-party services. Recent geopolitical developments have reintroduced availability and resilience risks previously considered mitigated by cloud adoption. This paper argues that an integrated approach, combining risk management, maturity and capability assessments, and IT audit functions, provides a more comprehensive and cost-effective basis for evaluating organizational IT resilience than any of these methods applied in isolation. The paper maps resilience-relevant controls from two widely adopted cybersecurity frameworks, NIST Cybersecurity Framework 2.0 and ISO/IEC 27001:2022, to a generic Capability Maturity Model scale, and outlines a practical self-assessment methodology applicable to organizations of varying size and sector. IT audit is examined as a mechanism for providing independent assurance over assessment results. The authors conclude that while each approach individually contributes to resilience, their structured combination offers synergistic benefits not yet systematically addressed in the literature and propose this integration as a direction for future research.

Keywords: Resilience, Maturity model, IT audit, NIST CSF, ISO 27001, Risk management.

1 Introduction

Most modern organizations are inherently dependent on their information technology (IT) systems. Having evolved beyond its role as a support function, today's modern IT has evolved from a support function into a strategic asset and foundational driver of organizational value creation [1]. This includes data, processes (both technical processes as well ones supporting business functions) and technical systems (software, networking and hardware). However, managing and securing modern IT environments

has become increasingly complex owing to layered architectures and a mix of technologies: on-premises systems, combined with cloud computing, internet of things (IoT), third-party dependencies and artificial intelligence (AI). Adding people factors to this list, primarily skills and expertise, geographical distribution of various (virtual) teams and high staff turnovers, create additional complexity which is in some cases even challenging to navigate through.

Many enterprises have fully or partially migrated their IT environments to cloud infrastructure, with recent industry surveys indicating that over 90% of organizations will adopt hybrid cloud through 2027 [2]. Trend of moving key production systems off the organization premises, clearly brings many advantages, as well as shifted focus to different types / classes of risks. One of the key ones addresses lack of availability – limited or broken access to any critical part of organizations IT infrastructure and services may lead to high disruption or unavailability of key business processes. Also, there are many obvious solutions to this issue which may minimize the risk: choosing reliable cloud vendor(s) with guaranteed minimum downtime, implementing synchronization between two different cloud providers, even storing some of the IT elements on premise, as an "old-fashioned" way to maximize the continuity of operations [3].

Other types / classes of risks are related to compliance (e.g. privacy-related risks), information and cybersecurity risks (e.g. advanced persistent threats, data leakage etc.) and various technology and platform-specific risks and limitations, making risk ecosystem more complex and challenging to manage. Although cloud computing has matured in the past decade in all above-mentioned aspects, primary driver for migrating into cloud is still availability and reliability [3] [4]. Responsibilities for the up-time, better performance and scalability were transferred to the could vendor. However, latest geopolitical developments initiate reconsideration of decisions to move into cloud. News that Amazon Web Services (AWS) datacenters has been physically attacked and damaged in United Arab Emirates (UAE) and Bahrain in the March 2026 [5] [6], brings availability and reliability risks, which were previously thought to be practically remediated with the cloud computing, again to the C-level executives agenda.

Systematic and comprehensive approach toward organizations IT readiness in a complex geopolitical period, become increasingly important to establish robust resilience and maximize key IT and dependent business processes survivability [7]. A maturity level and process capability assessments, particularly in the area of incident management and business continuity areas, may help address increased risk, and be both cost-effective and goal-effective method to optimize limited resources toward achieving desired state of IT processes resilience.

1.1 The Complementary Roles of Risk Management, Maturity Assessment, and IT Audit in Organizational Resilience

Modern organizations regularly measure and oversight variety of "common" types of risks: financial risks, operations risks, legal and compliance risks, different reputational risks (e.g. brand risks), IT and cybersecurity risks (often as a sub-category of operational risks), market risks, strategic risks etc. Transparency of organization's weak

point areas and how they are assessed / measured is not only a good business practice but a mandatory compliance obligation as well [8]. In financial institutions, for example, risk departments are often perceived (and consequently positioned) as an important business units, with representatives within highest governing bodies (Chief Risk Officers – CRO, is a common high role responsible for managing overall organizations risks). Risks may be identified and measured on a various way, where most common approaches are by self-assessment exercises, relying on a result of different audits (internal, external), and in IT area, commonly by utilizing internal automation tools, post-incident reviews and regular external checks (e.g. penetration testing).

Improving overall process capabilities and maturity levels of continuity-related process groups, may provide positive synergic effect with risk management activities, resulting in more effective organization resilience to different crisis. Finally, engaging IT audit function to provide additional level of assurance, e.g. by confirming maturity assessment levels and identifying opportunities for improving assessment process, may further strengthen organizational resilience and add additional value. This set of complementary activities – risk management, maturity / capability assessments and IT audits – can jointly provide assurance and guidance to executive leadership on the state of key business processes reliability and may help set measured and informed expectations toward further improvements.

1.2 Research Methodology

This paper presents a conceptual framework for integrated IT resilience assessment, developed through mapping of established international standards in risk management, cybersecurity and IT Audit. The research adopts an approach focused on artifact construction (framework and self-assessment method).

Initially, two well-known cybersecurity standards were selected: NIST CSF 2.0 and ISO/IEC 27001:2022. These two were chosen primarily because they represent the most widely adopted, internationally recognized, and complementary standards in information security management. NIST CSF 2.0 is free to use and emphasizes outcomes and flexibility across sectors, while ISO 27001:2022 provides certifiable, auditable requirements organization must comply with. For the maturity model, instead of NIST CSF Tier model, a well-known and (also) widely adopted Capability Maturity Model (CMM) was selected. Generic, five-level maturity scale (Initial to Optimized) provides simplicity in maturity evaluation.

Synergy of NIST CSF 2.0, ISO 27001:2022 and CMM is referred as a “Unified Framework for Cybersecurity Maturity Assessment”, where flexibility of joint approach has been proven through practical applications in different industries resulting in improved and measurable information security function [9].

Risk management activities, primarily risk identification and evaluation in organizations are carried out regularly in line with established risk management methodology. In this paper, authors assume that identified risks are results from the mature risk management processes, that are properly quantified and that can be treated as a reliable.

Role of IT audit in integrated IT Resilience is twofold: as an (independent) assurance layer and as a consultant function in activities where there is no conflict of interest.

No primary empirical data was collected for this conceptual paper; instead, the framework synthesizes existing best practices to propose synergistic integration and prepares a theoretical / normative baseline for future case studies.

2 Aligning the IT Control Environment with Established Cybersecurity Frameworks

Organizing cybersecurity functions, protecting IT systems (infrastructure, applications, services etc.) managing cybersecurity incidents and recovering from the potential damage, can be done on an *ad-hoc* basis or systematically, by following recommendations (best practices) organized in one of the internationally well-known cybersecurity frameworks. Cybersecurity frameworks are commonly defined as a structured set of security controls, criteria, requirements, objectives, or desired states of IT systems that organizations use to manage and mitigate risk, primarily cybersecurity risks. Frameworks vary in their approach to governance and organizational processes, but most often contain a set of governance and security controls that are considered a basic standard or are selected using a risk-based approach [8].

By aligning with a chosen and internationally recognized cybersecurity framework, organizations can improve their cyber defenses and reduce exposure to cyber-attacks. Also, other, indirect benefits can be achieved, such as:

- More efficient management of the IT security organization, e.g. establishing and managing an adequate structure of the department for information / cyber security.
- More effective management of cyber security risks and a more adequate response to them.
- Improved system of internal controls in the field of cyber security and general IT controls.
- More effective protection of digital assets, primarily critical information.
- The possibility of certification, that is, independent confirmation of compliance with controls.
- Cyber security frameworks can also help organizations meet regulatory requirements and facilitate communication about cyber security issues among internal and external stakeholders (e.g. establishing an effective exchange of information about cyber security threats in the financial sector).
- Many cybersecurity frameworks, such as the NIST Cybersecurity Framework (CSF), currently in a version 2.0, encourage a continuous cycle of assessment, implementation, monitoring, and improvement.

Successful examples of improving cyber security, which are the result of alignment with the most well-known frameworks [10], can be found both in private [11] and in government organizations [8].

Organization resilience readiness assessment by using NIST CSF or ISO 27001:2022 frameworks is commonly measured by comparing percentage of compliance with these two (or any other) frameworks. Executive level leadership may take the information of cybersecurity status through a "black or white lenses" – if compliance levels are above required threshold, there is no need for further (detailed) examination and often, no further investments in cybersecurity. However, more operational IT and IT security functions benefit from deeper technical analysis, where penetration tests are often treated as an addition or compensation mechanisms for "shallowness" of the compliance audits.

Several documented cases demonstrate that organizations holding valid ISO 27001 certifications have nonetheless experienced significant security incidents and data breaches, suggesting that certification alone does not guarantee operational security maturity [10]. Although this may not be logical at first, these occurrences may be understandable. Certificate itself assures that minimum set controls covering all defined areas exists, that they are (in most cases) effective, at least on the design level, and that evidence demonstrates that organization is (also at least) aware what cybersecurity aspects it needs to address. On the other side, penetration test reveals lack of deep technical controls, addresses sophisticated cybersecurity risks in implementation of controls and may require and initiate immediate mitigation actions. Business continuity test, which are in most industries mandatory (by validating business continuity plans), often are performed only for compliance purposes, end even then, the least invasive tests are executed (e.g. walk through, partial simulations etc.), purely to fulfill the form.

2.1 Resilience-Relevant Controls in NIST CSF 2.0 and ISO/IEC 27001:2022

NIST Cybersecurity Framework (NIST CSF) 2.0 provides guidance to industry, government agencies, and other organizations to manage cybersecurity risks. It offers a taxonomy of high-level cybersecurity outcomes that can be used by any organization — regardless of its size, sector, or maturity — to better understand, assess, prioritize, and communicate its cybersecurity efforts [12]. NIST CSF v2 contains 106 subcategories (called outcomes) organized into 22 categories and grouped into 6 functions: GOVERN (GV), IDENTIFY (ID), PROTECT (PR), DETECT (DE), RESPOND (RS) and RECOVER (RC). Latest three are common phases of incident management process, starting with (adverse) event detection and analysis, through incident management process and finalizing with recovery activities. In total, 8 categories and 32 subcategories are related to the operational resilience activities (excluding governance part), as shown in the table 1:

Table 1. NIST CSF v2.0 resilience controls

Category ID	Category Title	# of sub-categories
DE.CM	Continuous Monitoring	5
DE.AE	Adverse Event Analysis	6
RS.MA	Incident Management	5
RS.AN	Incident Analysis	4
RS.CO	Incident Response Reporting and Communication	2
RS.MI	Incident Mitigation	2
RC.RP	Incident Recovery Plan Execution	6
RC.CO	Incident Recovery Communication	2

Although common and logical focus organization may be to organize cybersecurity function and to strengthen prevention controls to minimize likelihood of cyber incidents occurrence, lack of mature incident management processes may result in high level of overall damage (e.g. financial, operational, reputational impact) and organization may experience. And one of the key performance indicators within incident management process is availability restoration to defined levels (ideally to pre-incident levels).

The ISO/IEC 27001 standard provides companies of any size and from all sectors of activity with guidance for establishing, implementing, maintaining and continually improving an information security management system [13]. The ISO 27001:2022 Annex A resilience activities are mostly organized in a 7 annex A controls, as shown in the table 2:

Table 2. ISO 27001:2022 resilience controls

Control ID	Control Title
A 5.24	Information Security Incident Management Planning and Preparation
A 5.25	Assessment and Decision on Information Security Events
A 5.26	Response to Information Security Incidents
A 5.27	Learning From Information Security Incidents
A 5.28	Collection of Evidence
A 5.29	Information Security During Disruption
A 5.30	ICT Readiness for Business Continuity

2.2 Selecting an Appropriate Maturity Model Framework

Conducting maturity or capability assessment jointly with compliance activities (e.g. ISO audits), may improve risk assessment efforts and provide to the higher leadership,

middle and operational management and cybersecurity practitioners more comprehensive, interconnected insight of overall IT and cybersecurity maturity state of the entire organization. Process capability is defined as characterization of the ability of a process to meet current or projected business goals [14].

Selecting proper maturity / capability model itself may provide a challenging task. Comprehensive literature reviews in this area [15] suggested that there are hundreds of models available, starting from general, well-known and highly adaptable frameworks for most use-cases (e.g. CMM-I, ISO 33000series, primarily ISO 33020:2021 [16] standard etc.) to the industry-specific, and even capability-specific maturity model frameworks [17] [18]. In a cybersecurity field of expertise, most well-known maturity and capability level frameworks are almost readily available "out-of-the-box"; user base is international, many use-cases are well described and both scientific literature as well as many practice implementation guidelines are freely accessible on the Internet.

Maturity assessment of complete NIST CSF subcategories or ISO 27001:2022 controls may represent a cost-efficient, but at the same time time-consuming activity, particularly if done internally, relying on skills and effort of the employees. Initial assessment may use a generic Capability Maturity Model (CMM) to evaluate general state of the NIST CSF sub-category or ISO 27001:2022 control. Generic CMM contains five levels, as shown in the fig 1:

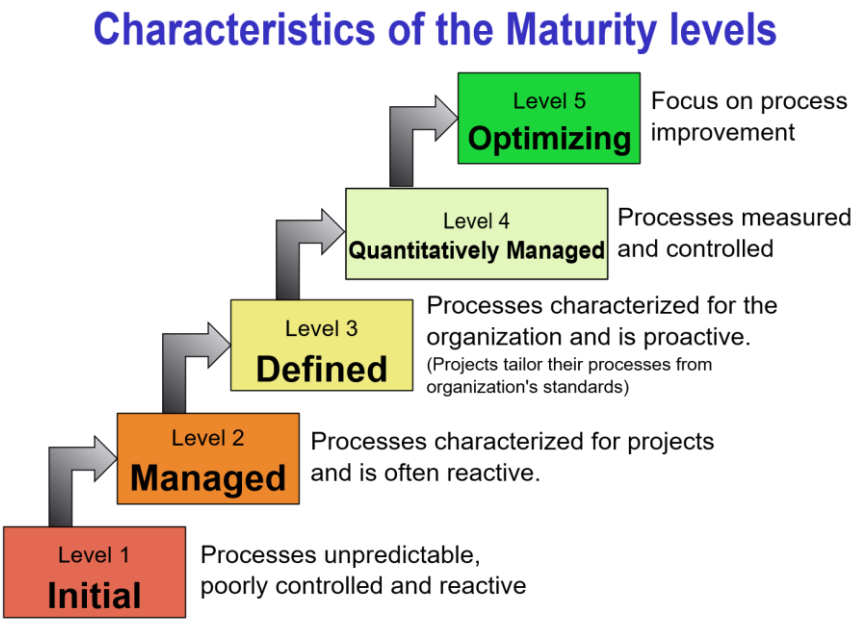


Fig. 1. Generic CMM scale / levels.

In practice, most ISO 27001 Annex A controls are expected to demonstrate at least a basic, documented level of implementation, while NIST CSF sub-categories are defined as outcomes, and achieving level of alignment is voluntarily, providing organization a flexibility in setting goals.

2.3 Implementing a Cost-Effective Maturity Self-Assessment

Studies have confirmed positive relationship between organizations resilience and higher maturity levels [19]. Maturity levels 3 (defined process) and 4 (Quantitatively managed) transforms organization from reactive to proactive culture. This is particularly important in IT, where certain IT unavailability (ITU) risks / incidents may lead to substantial strategic consequences [20].

Maturity assessment process is often internally implemented by utilizing self-assessment techniques, and verified by another internal team or external entity, to avoid conflict of interest and to improve the quality of results. For the cost-effective exercises, a simple questionnaire shall be created with the following guidelines:

- A list of desired set of NIST CSF sub-categories or ISO 27001:2022 controls, in a form of questions is created. Questions were derived directly from control/subcategory descriptions, rephrased for practitioner usability while preserving original intent.
  - Ex: an ISO 27001 control "A 5.27 Learning From Information Security Incidents" can be rephrased as follows: "A 5.27 How would you evaluate Learning From Information Security Incidents process?"
- Response scale (1–5) and supporting columns (evidence/comments) follow established self-assessment practices in maturity modeling:
  - A column where answer (maturity level) can be provided shall follow, preferably by limiting possible values from 1 to 5.
  - Short description of maturity level categories shall be visible permanently to the assessor, or be easily available for the reference
  - A comment column for additional information shall be available, where assessors may provide useful info which may be used to further improve the quality of specific level value or some other part of the questionnaire.

The questionnaires shall be then distributed to selected group of employees (assessors). All assessors shall be motivated to respond truthfully and to be ready to support their selected value with adequate rationale, if necessary. One of the examples how a successful self-assessment may be created is provided in the literature [21].

2.4 The Role of IT Audit in Assuring Organizational Resilience

Role of the IT audit in supporting organization resilience has been increasingly discussed in last period, with general conclusions that added value requires a fundamental

shift in audit mandate, skills, and positioning consistent with broader calls for internal audit reform [22]. Furthermore, several reference models for auditing organizational resilience have been proposed [23], and can be easily adapted to the specific IT audit methodologies and audit programs.

3 Conclusion

Risk management, maturity/capability assessments, and IT audit offer complementary perspectives through which organizational IT resilience can be evaluated. While each approach individually provides valuable insights, their structured integration offers synergistic benefits that have not yet been systematically addressed in the literature.

To validate and operationalize the proposed framework, authors emphasizes that practical steps toward combining results from three different perspectives shall be developed, with the following proposed future research agenda:

- Development of a formalized methodology / processes for synergizing risk management, maturity assessment and IT audit.
- Empirical validation through multiple case-studies in different sectors (e.g. finance / banking, pharmaceutical, critical infrastructure etc.) and organization sizes.
- Longitudinal studies measuring the impact of the integrated approach on actual resilience metrics (e.g., Mean Time to Recovery, business downtime costs, audit findings reduction etc.)

Such research would provide practitioners with actionable tools and contribute to the evolving literature on organizational resilience in hybrid IT environments.

4 Declarations

Declaration of Interests: The authors declare no competing interests.

Declaration of Generative AI and AI-assisted Technologies: During the preparation of this manuscript, the authors used Google Gemini to improve language clarity, grammar, and style. All content, ideas, structure, and scientific arguments were developed and reviewed by the authors, who take full responsibility for the final version.

References

1. J. C. Henderson and N. Venkatraman, "Strategic alignment: Leveraging information technology for transforming organizations," *IBM Systems Journal,* vol. 38, p. 472–484, 1999.
2. "Gartner Forecasts Worldwide Public Cloud End-User Spending to Total $723 Billion in 2025," gartner, 19 11 2024. [Online]. Available: https://www.gartner.com/en/newsroom/press-releases/2024-11-19-gartner-forecasts-worldwide-public-cloud-end-user-spending-to-total-723-billion-dollars-in-2025. [Accessed 26 4 2026].
3. K. C. Thatikonda, "Comprehensive Cloud Migration Strategies: Optimizing for Reliability, Security, and Performance," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology,* vol. 11, no. 2, pp. 72-82, 3 3 2025.
4. J. Vupputuri, "Data Migration to Cloud - Strategies, Challenges, and Best Practices," *International journal of computer engineering & technology,* vol. 15, no. 5, pp. 1084-1092, 2024.
5. "Iranian Drones Damage AWS Data Centres in the UAE and Bahrain," Wired Middle East, 3 3 2026. [Online]. Available: https://www.wired.me/story/when-iranian-drones-hit-the-cloud-aws-data-centres-damaged-in-the-gulf. [Accessed 24 4 2026].
6. S. Humphreys, "Reports: Iran Missiles Hit AWS Cloud Site in Bahrain," DataCentre Magazine, 2 4 2026. [Online]. Available: https://datacentremagazine.com/news/iran-missiles-hit-aws-cloud-infrastructure-in-bahrain. [Accessed 24 4 2026].
7. Z. Keković, "Efekti neizvesnosti u odlučivanju o sistemskim rizicima," *faculty of security studies annual,* pp. 15-37, 2022.
8. A. Vestad and B. Yang, "From Security Frameworks to Sustainable Municipal Cybersecurity Capabilities," *Journal of Cybersecurity and Privacy,* vol. 5, no. 2, p. 297–313, 2025.
9. T. Abdiukov, "A Unified Framework for Cybersecurity Maturity Assessment Across CISSP Domains: Integrating NIST, ISO 27001, and CMMI for Holistic Security Posture Evaluation," *Journal of Emerging Technologies and Innovative Research (JETIR),* vol. 10, no. 11, pp. 743-753, 2023.
10. S. Todström, *The effects of ISO 27001 certification,* Umeå University, Department of Informatics, 2024.
11. A.-M. Udroiu, M. Dumitrache and I. Sandu, "Improving the cybersecurity of medical systems by applying the NIST framework," in *14th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, 2022.

12. National Institute of Standards and Technology (NIST), "The NIST Cybersecurity Framework (CSF) 2.0," [Online]. Available: https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.29.pdf. [Accessed 30 5 2025].
13. " ISO/IEC 27001:2022 Information security, cybersecurity and privacy protection — Information security management systems — Requirements," International Organization for Standardization, 2022. [Online]. Available: https://www.iso.org/standard/27001. [Accessed 26 4 2026].
14. "ISO/IEC 33020:2019(en) Information technology — Process assessment — Process measurement framework for assessment of process capability," International Organization for Standardization, 2019. [Online]. Available: https://www.iso.org/obp/ui/es/#iso:std:iso-iec:33020:ed-2:v1:en. [Accessed 26 4 2026].
15. R. Wendler, "The maturity of maturity model research: A systematic mapping study," *Information and Software Technology,* vol. 54, pp. 1317-1339, 2012.
16. "SRPS ISO/IEC 33020:2021 Информационе технологије – Оцењивање процеса – Оквир мерења процеса за оцењивање способности процеса," [Online]. Available: https://iss.rs/sr_Cyrl/project/show/iss:proj:101532. [Accessed 31 01 2026].
17. J. Stocker, N. Herda, M. Wolf and S. Ruf, "A Maturity Model to Assess and Foster the Resilience of Organizations," *Applied Sciences,* 2022.
18. S. Grobbelaar, "Development of a Resilience Maturity Framework for Project-Based Organisations," in *33rd IPMA World Congress 2024*, DHL Stadium, Cape Town, South Africa, 2024.
19. L. Alves Sincorá, M. Paulo Valadares de Oliveira, H. Filho and M. Zamboni Alvarenga, "Developing organizational resilience from business process management maturity," *Innovation & Management Review,* vol. 20, no. 2, 2023.
20. A. K. Amiri, H. Cavusoglu, I. Benbasat and V. Grover, "A Theory of Strategic Information Technology Unavailability," *Information Systems Research,* 2025.
21. K. Lazić and V. Pantović, "NIST cybersecurity framework: preparation steps for successful assessment," *International Conferences INFOTECH 2020 – 2024 Selected Papers,* pp. 71-75, 2025.
22. M. J. Caldwell and D. Olasehinde, *Internal Audit's Contribution to Organizational Resilience and Crisis Management,* Lake Mary, FL, USA: The Institute of Internal Auditors (IIA) Research Foundation, 2022.
23. I. van Maaren, "A reference model for auditing organisational resilience," *Maandblad voor Accountancy en bedrijfseconomie,* vol. 96, no. 7/8, pp. 201-211, 9 2022.

THEMATIC SECTION

Quality Management and Standardization

Resilient and Intelligent Organizations in Times of Permanent Disruption

SYMORG 2026 • CONFERENCE PROCEEDINGS

Original Research Article *https://doi.org/10.5281/zenodo.23477089*

The Impact of the Application of Quality Management System Standards in Automotive Supply Chains

Simo Denić[1][0009-0000-0882-6219] and Jelena Ruso[2][0000-0003-4120-8759]

[1]Brose Pančevo, Pančevo, Serbia
[2]University of Belgrade – Faculty of Organizational Sciences, Jove Ilića 154, Belgrade, Serbia

Abstract. After Henry Ford revolutionized the automotive industry by creating mass production on the assembly line, automakers worked ways to create a high-quality product. In the beginning, the achievement of quality was achieved through mere quality control, but over time the idea of incorporating quality into the product was developed. The idea of shifting responsibility from quality controllers to the entire organization with an emphasis on management was popularized in the late 1970s, when the Japanese automotive industry developed at an incredible speed, which, along with Ford, undeniably brought the second biggest revolution in this industry. Driving in a car, in itself, poses a risk to the life and health of those inside and outside the car, and consequently achieving product quality is of primary importance to car manufacturers. In order to create high-quality products, it was first necessary to adequately arrange the system of organization with an emphasis on quality – the quality management system. By merging organizations operating in the automotive industry, but also individually, mainly by geographical affiliation, standards for the quality management system specific to this branch of industry were developed, and it was not until the end of the 1990s that there was a big shift and the creation of the IATF 16949 standard, which is the basis of the automotive industry. This paper will present the role of the mentioned standard, developed by an informal sectoral organization for standardization, in quality assurance, and then through the presentation of the results of the conducted research, the level of understanding of this standard in an organization operating in the automotive industry will be presented.

Keywords: Quality, Standard, Quality Management System, Automotive, Suppliers, IATF 16949, VDA, AIAG.

1 Introduction

Although we are often not aware of it, modern life as we know it is conditioned by the existence and respect of a large number of standards, and as Professor Mijatović adequately described: "Modern life is not possible without standards and they are all around us" (Mijatović, 2019, p. 14). In the automotive industry, of course, standards occupy a central place – for example, the production of tires according to dimensions that are standardized allows the interoperability of rims and tires regardless of the geographical position and manufacturer of these two parts of the car. The development of

standards that define the requirements for the characteristics of fuel for cars has made it possible to refuel at any gas station on the way from Siberia to Portugal. The importance of the standard is also distinguished by the labelling used in the automotive industry – every car has the same label on the dashboard – lack of oil, lights on, brake problem, speed expressed in the appropriate unit. These seemingly simple and less important things can directly affect the safety of road users, which is why they are often found in the laws of countries (UNECE, 2025). On the other hand, the development of the NCAP series of car safety rating standards has made it possible to achieve a realistic comparison of the safety of different car models, since they are all tested to the same standard. Some organizations, such as the German National Institute for Standardization (DIN), also develop standards for the validation of components. In order to establish the same criteria for assessing the quality of components and compliance with the requirements throughout the supply chain (DIN, 2025).

However, the standards do not only define the requirements for product characteristics, they also define the most basic concepts such as units of measurement systems and their designations, they also define terms for easier communication, but also requirements for management systems. In the late 1980s and early 1990s, the International Organization for Standardization developed a series of standards for management systems in order to create a set of requirements and guidelines for their implementation in various spheres from quality to environment. By creating standards for management systems and their implementation, it is ensured, referring only to the production sector, improved business in the aspects of implemented standards, which is especially important for the aspect of quality. The development of specific standards for the quality management system of companies operating in the automotive industry has increased the degree of certainty that production will result in compliant products.

2 Automobile and Automotive Industry

According to the *European Automobile Manufacturers' Association (ACEA)*, a car is defined as: "vehicles designed and manufactured for the carriage of passengers, which have a maximum of eight seats, including the driver's seat, and have a maximum gross vehicle mass without passengers not exceeding 3.5 tonnes" (ACEA, 2025). Modern cars are very complex technical systems and consist of thousands of components, some of which have been present from the very beginning of the car's development, such as wheels and steering wheels, all the way to modern systems such as radar sensors for keeping distances and reversing cameras (Cromer & Purry, 2025). All companies whose activities relate to the development and production of cars, car parts, their distribution as well as the necessary subsequent maintenance of cars belong to the automotive industry sector. It is important to note that the automotive industry also includes companies that have focused their activities on commercial vehicles and trucks (Cromer & Purry, 2025).

In fact, the beginning of the automotive industry as we know it today is considered to be back in 1886, when engineer Karl Benz, whose surname would later become synonymous with the production of high-quality vehicles, created the first car that was

powered by an internal combustion engine (Warmuz, 2025). The center of the automotive industry moved from Europe to the United States in the early twentieth century, when Henry Ford founded the Ford Motor Company, which is still successful today. Ford succeeded in using mass production, assembly line, interchangeability of parts and the use of specialized machinery to create a car that was affordable for everyone. Such has been the success of the company that in just two decades the company manages to produce as much as half of all cars globally (Jaško et al., 2014; Ford, 2024). Henry Ford emphasizes that mass production as a term does not refer to the quantity of production, standardization, or mechanization of labor processes (*because these techniques can be applied independently, without mass production*), but is a modern method of production. Mass production integrates a large number of production factors such as cost-effectiveness, efficiency, quality, standardization, speed and continuity in order to create a product at the lowest possible cost. Also, Ford states that the condition for establishing mass production is the existence of a high demand for a certain product, but states that customers do not necessarily have to be aware that they can meet the need with that product (*demand can be created*). He also adds that mass production makes it possible to create mass demand, that is, one is not possible without the other (Ford, 2024). After the Second World War, there was a revival and great technological development of the automotive industry in Europe and Asia, especially in Japan, Germany and France, where the largest number of car manufacturers and manufacturers of car components are still located, but until 1980 the largest car manufacturer was the United States, until it was replaced by Japan thanks to the incredible success of Toyota (Binder & Rae, 2024).

3 Automotive Supply Chains

Car manufacturers or so-called OEMs (Original Equipment Developers) are not the only members of the automotive industry. One group of companies that certainly plays the largest role in the automotive industry, apart from OEMs themselves, are manufacturers of automotive components (ZF, Bosch, Continental, Brose, Lear, Leoni). By entrusting the production and development of components, car manufacturers significantly reduce costs and increase product quality (Binder & Rae, 2024). According to research by the European Automotive Suppliers Association – CLEPA, 75% of the value of cars is made up of parts that are not produced by OEMs but by suppliers, and these companies employ over 5 million people in the EU. In addition, various other companies such as dealerships, car repair and maintenance workshops, are part of this industry (CLEPA, 2025).

As the final product – a car can have from 15 to 25 thousand components, supply chain management is of great importance in this industry (Leggett, 2024). Different authors define the supply chain in different ways, and it is possible to create an entire overview of the proposed definitions, but in general, the supply chain is "a network of organizations involved, through upstream and downstream links, in various processes and activities that produce value in the form of products and services delivered to the end consumer" (Mentzer et al., 2001, p. 3). Supply chains need to be managed. Supply

chain management (SCM) must be in place in order to provide the most efficient and effective way to the user. The buyer of the goods is delivered to the same and thus reduces logistics costs and maximizes the value of the product, and consequently achieves a competitive advantage for the supplier (Christopher, 1994). The supply chain in the automotive industry is usually divided according to levels (*eng. Tier*) starting with the manufacturer of raw materials (*aluminum, copper, etc.*) or in some cases companies that carry out technological processing of materials and thus get to the basic parts (*screws, wire, castings, etc.*). These companies are Tier 2 or Tier 3 suppliers who send their parts as a gateway to the next process to the sub-assembly companies and send their parts directly to the automakers, Tier 1 suppliers, i.e. Tier 1 suppliers (ZF, Bosch, Continental, Brose). Of course, this distribution can be adapted to the specific situation in the supply chain, but the division by levels is always present in the automotive industry (Leggett, 2024).

4 Development and Implementation of QMS Standards in the Automotive Industry

The reasons for the implementation of standards for quality management systems in the automotive industry, but also in general, can be categorized into two groups of factors, internal and external implementation factors (Bravi et al., 2019): Internal factors: Striving to achieve a more orderly quality management system, increasing efficiency, reducing entropy, increasing quality awareness; External factors: Access to new markets, customer and state pressures, increased competitiveness. In any case, regardless of the primary motivation factor, the implementation of standards for a quality management system is a strategically good decision, but often also an obligation imposed by contracts between supplier and customer (Bravi et al., 2019; Hanzličková, 2019).

The full implementation of standards for quality management systems requires a review of all processes, starting from the analysis of customer requirements to the creation of work instructions for workers and their marking, but also the setting of quality goals at all levels (Hanzličková, 2019). In fact, due to increased competition and stringent requirements in the automotive industry, automakers and their suppliers have strived to increase quality and efficiency with a constant need to reduce costs. The adoption of quality as the basic philosophy of this branch of industry has imposed the need to formalize quality assurance through quality requirements and numerous standards governing the quality management system specific to the automotive industry (VDA, AIAG, IATF, TQM) (Goicoechea & Fenollera, 2012).

Modern manufacturing, as already described, is characterized by the specialization of companies in what they are good at, the business model of companies such as the one devised by Henry Ford – mastering the entire production and distribution chain – is almost extinct. Modern organizations reduce the range of products and services they produce and provide and use outsourcing to compensate for the processes they lack (Jaško et al., 2014). Nowhere is this business model more noticeable than in the automotive industry, where automakers have turned their industrial plants into assembly halls, producing only one component of the final product on their own. Responsibility

for the production of the necessary components for the car, and often the design and validation of the components, is left to a nominated supplier specializing in a particular system (suspension, steering, cooling, interior, software) (ACEA, 2025).

By delegating responsibility for design, manufacturing, and other responsibilities, automakers such as Ford, Toyota, and others will significantly reduce the cost of producing a car and increase its quality. This way of nominating suppliers based on set requirements is present throughout the supply chain, and achieving quality and efficiency is the main goal of all of them (Goicoechea & Fenollera, 2012). Given this, the application of quality management system standards is a very important factor in the automotive industry, since its implementation will increase the likelihood that suppliers will achieve a satisfactory level of product quality. By regulating processes, as previously mentioned, the standard for a quality management system reduces the entropy of the system and encourages the use of quality tools that, in addition to increasing the level of quality, reduce costs and prevent the occurrence / recurrence of problems in production. During product development, buyers of parts (at all levels of supply) will check the supplier's compliance with the relevant requirements of the Quality Management System (QMS) standard through on-site inspections. And then there are the specific requirements of the customer (Goicoechea & Fenollera, 2012).

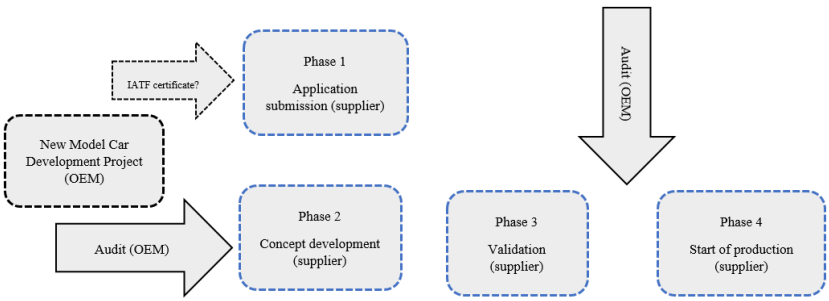


Fig. 1. Abstracted process of component development (Goicoechea & Fenollera, 2012)

5 Quality – Definition and Meaning

"Quality is a timeless concept, which is why the beginnings of quality management are hidden somewhere in the mists of the past" (Filipović & Đurić, 2009, p. 17). The root of the word quality, which is used today on a daily basis, originates from the Latin word "qualitas", which denotes a property or feature and can be associated with a product, process, service, life and other entities (Živković & Glogovac, 2015). From humble beginnings in the distant past and the concept of *caveat emptor* where quality was reduced to mere quality control by the customer, through the first artisanal guilds and quality guarantees by means of trademarks, and then the first production lines and automation created in the company of Henry Ford. During the Industrial Revolution, which enabled the establishment of mass production, quality was assigned to individuals, i.e. responsibility for quality was assigned exclusively to quality controllers. With the subsequent development of the science of quality, scientists realized that quality

must be built into the product and that responsibility for quality must be divided among all sectors of the organization. In fact, quality as a scientific and technical discipline experienced its significant expansion in Japan only after World War II, and on the basis laid by American engineers and scientists (Filipović & Đurić, 2009; Hoyle, 2007). It is also important to note that the term quality has three meanings (Filipović & Đurić, 2009):

- Quality as a property of the entity: Quality of education, quality of product workmanship;
- Quality as a function in the organization: Quality planning, quality engineering;
- Quality as a scientific-professional, engineering-managerial discipline.

ISO (International Organization for Standardization) in the standard SRPS ISO 9000:2015 (Fundamentals and Vocabulary) defines quality as: "The degree to which a set of inherent characteristics meets the requirements". Of course, in addition to the definition of ISO, there are many other definitions that can differ according to the angle of looking at the concept of quality – that is, quality can have a different meaning for different people and even a different meaning for the same person in a different situation. In view of this, authors Živković and Glogovac propose the following criteria for categorizing definitions of quality (Živković & Glogovac, 2015):

- User-oriented criterion – represents the purposefulness or degree of goodness of performing the basic function of the entity;
- Process-oriented criterion – compliance with the specification;
- Product or service-oriented criterion – the difference in quality is based on the difference in quantity of a certain quality characteristic (the more the better, i.e. the less the better);
- Value-oriented criterion – "sufficiently meet the needs of users and be cheaper than the competition";
- Subjective criterion – the idea of quality as a synonym for superiority where it is not entirely clear what quality really represents.

For a complete understanding of quality, it is necessary to refer to the concept of Juran's quality trilogy, which will later be used when analyzing the role of standards for quality management systems. The Quality Trilogy is the basic concept of quality, which involves a combination of global processes of planning, management and quality improvement (Filipović & Đurić, 2009). "Process-based organizations perform quality management through a trilogy of quality throughout the product life cycle, which achieves the delivery of products of a certain quality" (Filipović & Đurić, 2009, p. 91).

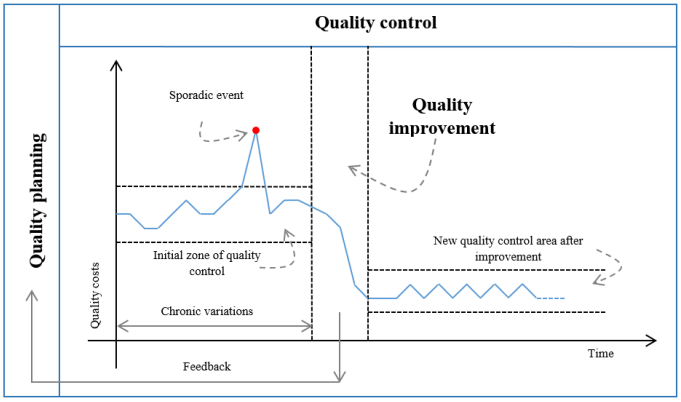


Fig. 2. Graphic representation of the quality trilogy (Juran & Godfrey, 1998)

6 Standard and Standardization

The generally accepted and most frequently cited definition of a standard, which does not reflect strategically and economically on the goals of the standardization organization, is the de Vries definition: "A standard is an approved specification of a limited number of solutions to existing and potential common and related problems, designed for the benefit of the party or parties involved, balancing their needs, whereby the solutions are intended and expected to be used by the parties repeatedly or continuously over a specified period of time" (Mijatović, 2019, p. 35). The United Nations defines a standard as "a document that defines, among other things, requirements, rules and guidelines for processes, products and services" (United Nations Industrial Development Organization (UNIDO), 2006). Similar to the concept of quality, the word standard is used in everyday speech, acquiring very different meanings depending on the context in which it is used. Professor Mijatović lists three meanings of the word standard (Mijatović, 2019):

- Standard in terms of acceptable and consistent quality: In this context, a standard is considered to be an acceptable quality to which customers are accustomed. For example, a 22-inch monitor with a resolution of 1080p is considered acceptable/sufficient;
- A standard in terms of recognizability for its excellence: In this sense, a standard is equated with a best-in-class solution, that is, a standard is something with which everything else compares. For example, Toyota car reliability – Toyota has set a standard for car reliability that serves as a benchmark for comparison;
- A standard in terms of something that is generally accepted / in terms of a model: In this context, a standard is established by officials such as national governments or international organizations. For example, A4 paper size, dimensions of the shipping container, flammability of the fuel.

On the other hand, standardization is defined as: "the activity of establishing and recording a limited set of solutions, existing or potential common and related problems, where the solutions benefit the party or parties involved; based on a balance of the needs of the parties involved; intended for a significant number of parties with the expectation that the solutions will be applied repeatedly or continuously during a certain period" (Mijatović, 2019, p. 35). The aforementioned scientist also lists several goals of the development of standards and says that standards are developed in order to (De Vries, 1999): Enable interoperability of products; Reduce the diversity of solutions; Facilitate communication and organization; Facilitate the establishment of global trade and compatibility solutions; Protect the interests of users of products/services through the reduction of information. Standardization can be divided, from the point of view of the organization developing the standard (Mijatović, 2019):

- Formal standardization / de jure: This type of standardization – the development of standards – takes place within the framework of formal standardization organizations, which are divided into international (ISO, IEC and ITU), regional (CEN, CENELEC, ETSI, PASC) and finally national standardization organizations (DIN, AFNOR, ISS, ANSI). Although they are referred to as formal standards organizations, there is no official organization (such as the UN or WTO) that recognizes these standards organizations as formal or selected, in fact they are all private organizations established by industry representatives. On the other hand, many countries have recognized the work of national standardization organizations in their standardization laws, which participate in the work of international and regional organizations (Mijatović, 2019);
- Informal standardization / de facto: This group of standardization organizations includes industry/professional associations, business associations, consortia and forums. Professional associations are organizations of professionals (engineers, doctors, etc.) that, among other activities, are engaged in the development of standards and have a wider field of activity than other informal organizations. Business associations, consortia, and forums are associations that are formed to develop standards in a particular area of interest and thereby generate economic or strategic gain. Examples of these organizations are VDA, AIAG, IATF, DVD, W3C, USB (Mijatović, 2019).

7 The Most Important Actors of Standardization in the Automotive Industry

Of course, the largest number of standards applicable and intended for the automotive industry have been developed by the formal organization **ISO** – International Organization for Standardization. ISO is made up of national standardization organizations – one country one representative – which represent intermediaries between national needs and global standardization, but also represent ISO in their country (Institute for Standardization of Serbia – ISS) (Mijatović, 2019). Technical committees, based in one of the national standardization organizations, are predominantly composed of national

committees and working groups for standards, in which experts in certain fields and representatives of interested companies work. Essentially, technical committees work on the development of standards by consolidating the work of several national commissions and working groups for standards, which are formed by national standardization organizations in areas of strategic and economic interest to the state (Mijatović, 2019). An overview of the "International Classification of Standards" available on the ISO website shows Section 43, which is dedicated to standards for road vehicles.

One of the most important actors in the development of standards in the automotive industry (according to section 43) is the ISO Technical Committee ISO/TC 22, whose secretariat is located within the French Association for Standardization – AFNOR (ISO, 2025). The area of activity of this technical committee is the development of standards for all types of road vehicles used on public roads, which are focused on safety, sustainability, compatibility and interchangeability of parts, maintenance, performance determination and, of course, the quality of vehicles. In addition, this technical committee is also active in the field of standard development for vehicle hardware and software, autonomous driving, test equipment and other relevant areas (ISO, 2025). Of course, in the automotive industry, ISO standards for management systems such as ISO 9001, ISO 45001, ISO 14001 and others are also applicable.

The IATF is a group (association) of car manufacturers and their national sectoral standardization organizations (AIAG, ANFIA, FIEV, SMMT and VDA), with the aim of reaching a consensus on the basic requirements for the quality management system formalized through the IATF 16949 standard, which will be distributed throughout the supply chain all the way to the supplier of raw materials. By creating representative offices in the USA, Germany, France, Italy, the United Kingdom and the People's Republic of China, the IATF exerts global influence, accredits certification bodies and oversees their work (IATF, 2025). It is important to note that in addition to the implementation of the IATF 16949 standard, very important Customer Specific Requirements (CSR) must be analyzed and met in practice, which are officially communicated to suppliers through the IATF website. In addition to participating in the development of universal standards for the automotive industry, each OEM creates a list of additional requirements and thus contradicts the idea of synchronization and cost reduction.

VDA – The *German Association of the Automotive Industry* is one of the most influential informal standardization organizations (business associations) in the automotive industry and Germany's representative in OICA, CLEPA and ACEA. The VDA unites over 620 organizations representing manufacturers of automobiles, trucks and commercial vehicles as well as automotive parts and equipment, with the aim of representing the interests of the German automotive industry at European and global level (VDA, 2025). The members of the VDA are divided into the following groups: Group I consists of German car manufacturers (VW, BMW, Mercedes); Group II consists of German manufacturers of buses and trailers (MAN, Kress, Volvo); Group III consists of German manufacturers of car parts (Bosch, ZF, Brose); Associate members deal with infrastructure, raw material production, software development, and other areas of relevance to the automotive industry (VDA, 2025). VDA is divided into sectors, and within the VDA QMC sector, standards for the automotive industry have been developed since 1997. The most well-known standards, in addition to VDA 2 – PPAP and VDA 5 –

measurement systems and quality control, developed by the VDA QMC are shown in Fig. 3 (Đošić et al., 2019; VDA, 2025).

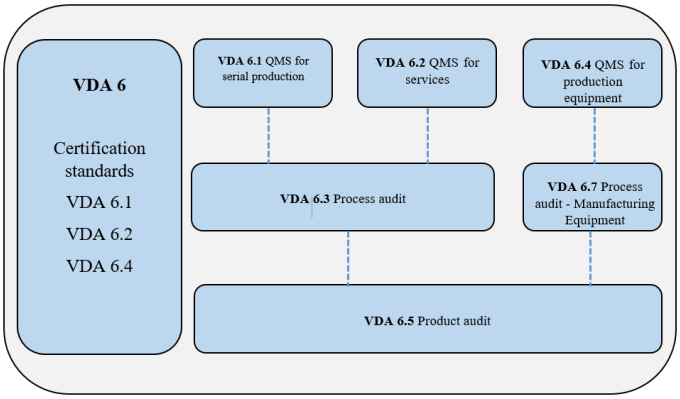


Fig. 3. Structure of the VDA 6.x series of standards (VDA, 2025)

Just as the VDA was founded to unite car manufacturers from Germany, on the other side of the planet in 1982 the **AIAG** – Automotive Industry Action Group was founded in the USA to unite American motor vehicle manufacturers (GM, Ford, CAT, Humvee, Jeep) and is the US representative in the IATF. Today, AIAG has over 5,000 members and unites other motor vehicle companies with headquarters around the world: Honda, Nissan, Toyota, Geely, Stellantis, JLR Group, Rivian, Lucid (AIAG, 2025). In addition, AIAG members include component and software manufacturers, dealers and vendors, the scientific community, and others. The goal of this organization is to develop standards and solutions that aim to reduce the cost and complexity of the automotive supply chain as well as increase quality. The standards developed by AIAG are focused on the entire life cycle of products in the automotive industry – development and serial production (*components or cars*) and are grouped into: Standards for 3D models, standards for production and troubleshooting, traceability, complaint management and warranty), standards for basic quality tools (PPAP, MSA, SPC, APQP, FMEA, CP, self-checks) and standards for specific processes known as CQI-xx (welding, heat treatment, pouring, injection molding) (AIAG, 2025).

In addition to the above-mentioned organizations, standards, manuals, instructions and other material unchanged in the automotive industry are developed by other organizations such as: IEC, ACEA, CLEPA, OICA, but also car manufacturers themselves develop a large number of standards and requirements often imposed on suppliers through regulations or CSR documents.

8 Standard for Quality Management System in the Automotive Industry

Before considering who and what standards are developing for the purpose of regulating the automotive industry, it is important to briefly explain what a quality management system actually means, and then a standard for a quality management system. By reviewing the various available definitions of systems, management systems, quality systems and in short quality management systems, it can be concluded that a quality management system is a set of interrelated or interacting elements aimed at coordinating activities to determine the direction and management of an organization in relation to quality. With further analysis, it can finally be understood that the QMS is a set of interrelated or interacting elements for planning, managing, improving and ensuring quality (Filipović & Đurić, 2009). Standards developed for a quality management system are specifications that contain a limited number of solutions for a quality management system, i.e. they are specifications for regulating the processes of planning, management, improvement and quality assurance that are intended for all interested parties. It is also important to note that the management system did exist before the development of standards for management systems, in fact, its existence initiated the need for the development of both standards (Živković & Glogovac, 2022).

As the application of the standard, in principle, is not mandatory, buyers of goods (throughout the supply chain) required in their commercial contracts to have a certificate confirming the successful implementation of the quality management system, in order to increase the likelihood of achieving high and constant quality of the supplier's products. In the early 1990s, most automotive suppliers were ISO 9001 certified, which essentially constituted a license to operate in the automotive industry (Laskurain-Iturbe et al., 2021; Fahmi et al., 2021). Professor Hoyle states that the reason for this is the lack of a standard that is intended for a specific industry, since ISO 9001 is applicable in every organization, regardless of size, ownership structure / organization and activity of the same, as well as the different interpretation of the requirements of the standard by auditors (Hoyle, 2000; Janjić et al., 2022).

At the end of the 1990s, there was a need to unite different standards for the quality management system developed by different organizations/associations in the automotive industry. The idea was clear to everyone, different suppliers (for different components such as doors, windows, seats, but also rubber, capacitors, aluminum castings) in a large number of cases produce components for a large number of customers (regardless of what level they are at – Tier 1 / Tier 2 or OEM). In addition, there was pressure from suppliers to create a single standard for the quality management system, given that the certification of the organization according to different standards for the quality management system was an expensive and inefficient process, but without additional benefits for the entire supply chain (Yılmaz & Acaroglu, 2024; Hoyle, 2000).

The road to the creation of a common standard (in foreign literature there is also a name – specification) was long and filled with disagreements between interested parties. Initially, due to the need to create additional quality interventions in the automotive industry, different companies joined together in different organizations, often according to geographical criteria, or individually developed their own standards for the quality

management system as well as other types of standards such as the VDA 6 series (Guidelines for Conducting Audits) (Hoyle, 2000; Petkova et al., 2023). A revolutionary step took place in 1999 when, in cooperation with ISO and IATF, the technical specification ISO/TS 16949 was developed, and in 2016 it evolved into the IATF 16949 standard. As already explained, the IATF (*eng.* The International Automotive Task Force) is a group (association) of automobile manufacturers and their national sectoral standardization organizations (AIAG, ANFIA, FIEV, SMMT and VDA), with the aim of reaching a consensus on the basic requirements for the quality management system formalized through the IATF 16949 standard, which will be distributed throughout the supply chain all the way to the supplier of raw materials (Hoyle, 2000; IATF, 2025).

The development of the IATF 16949 standard did not start from a "blank sheet of paper", its basis is the ISO 9001 standard together with other standards for the quality management system for the automotive industry developed long before the IATF 16949 standard. Table 1 shows the contribution to the development of this IATF 16949 standard through a brief retrospective of the standards developed prior to IATF 16949, which Professor Hoyle cites as the basis for the creation of the common standard IATF 16949 (Hoyle, 2000). In the table you can see which organizations known to end customers (car manufacturers) have joined together in order to develop standards for quality management systems in the automotive industry, as well as which names of standards have arisen as a result of their work. In addition, the table contains a brief description of the standard and the year of its adoption. All of the standards listed below are now out of use, with the exception of the standards developed by the German organization VDA.

Table 1. Contribution to the development of ISO/TS 16949 (modified from Hoyle, 2000)

Organizations	Stand-ard	The basis for	Comments
Fiat, IVECO *Italian Association of the Automotive Industry [ANFIA]*	AVSQ	**IATF 16949**	First published in 1990. It is a manual for the auditor in the automotive sector together with a list of questions for the verification of the system that accompanies the chapters of the ISO 9001 standard with additional requirements, but does not supplement it.
PSA, Renault *French trade association for the automotive equipment and supplier industry [FIEV]*	EAQF		First published in 1994. It is a manual for auditors in the automotive sector together with a list of system audit questions that follow the chapters of ISO 9001, but does not constitute a supplement to it.
Opel, VW group, BMW, Mercedes *VDA*	VDA 6.x		First published in 1991. It is a manual for auditors in the automotive sector and is not an appendix to ISO 9001, but seeks to unify the interpretation of the requirements of that standard.
Chrysler, Ford, GM *AIAG*	QS-9000		First published in 1994. It integrates the ISO 9001 standard with specific requirements for the automotive industry.

As can be concluded from Table 1, different standards have been developed with the same idea as the basis – to improve, i.e. to improve. Adapting the ISO 9001 standard for the automotive industry. In addition to this review, Professor Hoyle also provides a chronological retrospective of the creation of the ISO/TS 16949 technical specification, starting with the development of the BS 5750 standard back in 1979, which was based on a military standard developed for the needs of the NATO alliance. The British National Institute for Standardization proposes to the International Organization for Standardization to develop and adopt an international standard for a quality management system on the basis of the BS 5750 standard, which is done in the form of the ISO 9001 standard. In fact, the British Institute for Standardization and the BS 5750 standard represent, at the highest level of observation, the basis of the IATF 16949 standard. A more detailed chronological overview (Fig. 4) also defines the relationship between the standards mentioned in Table 1 in the development of the technical specification ISO/TS 16949, which later evolves into the standard for the quality management system in the automotive industry IATF 16949 (Hoyle, 2000).

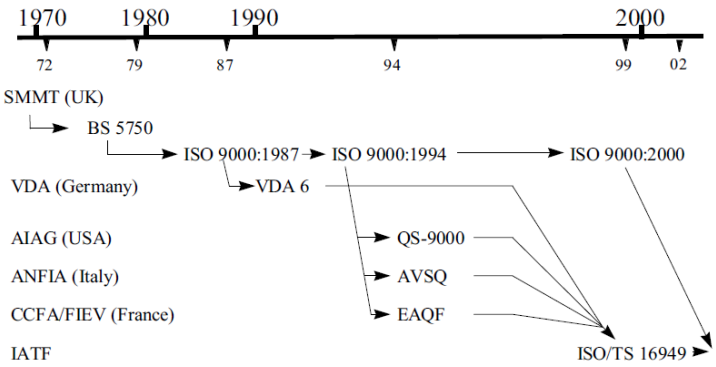


Fig. 4. Chronology of the development of ISO/TS 16949 (Hoyle, 2000)

Fig. 4 could be supplemented today by the evolution of the technical specification ISO/TS 16949 into a standard that was adopted in 2016 by the IATF and designated as **IATF 16949**. It is interesting to add that even today (in 2026) it is possible to find ISO/TS 16949 on the ISO website, but with the status "inactive" and the comment that it is necessary to perform a search on the IATF website (ISO, 2025). One of the reasons for the evolution of the technical specification into a standard is the development of a new version of ISO 9001 in 2015, when the requirements related to risk analysis appeared within this standard for the first time, which had to be integrated into ISO/TS 16949 also (Bozola et al., 2023; Cagnin et al., 2021).

In addition to establishing compliance with quality requirements through the IATF 16949 standard, the association of car manufacturers in the IATF organization established an equally important system for conducting audits and interpreting the requirements of the IATF 16949 standard, which consequently harmonized the assessment criteria throughout the supply chain (Hoyle, 2000). Harmonization of assessment methods has been achieved by a limited number of accreditation bodies that can issue a certificate of competence of the certification body to perform conformity checks of the test object according to the requirements of the IATF 16949 standard. Accreditation can only be carried out by IATF offices located within the sectoral organizations for standardization, namely: VDA, Germany; IATF, France; ANFIA, Italy; IAOB, USA; SMMT, UK (IATF, 2025). In this way, the responsibility of supplier certification according to the IATF 16949 standard has been shifted from the car manufacturer to the certification body, but with the full confidence of the end buyers of the components – the car manufacturer with regard to this accreditation scheme that reduces the possibility of abuse and corruption of the certification system that is under the control of the IATF itself.

In addition to a large number of requirements of the IATF 16949 standard related to the implementation and support, it is important to note that significant attention is paid to the design and development phase of the product, as a basis for adequate implementation of the product in further phases of the life cycle (Hoyle, 2000). Just as manufacturers of the final product require the possession of a certificate for the IATF 16949 standard, so do manufacturers of components require the possession of this certificate

from their suppliers, since a bad product at the lowest level of supply can lead to a defect in the end product and thus, in the worst case, jeopardize the safety of end customers – drivers and passengers. It is also in the best interest of suppliers to implement IATF 16949, given that product quality is the basis for long-term cooperation, good reputation and winning new business in the future (Hoyle, 2000; Waseem & Yusoff, 2025).

In addition to maintaining quality at a high level, the IATF 16949 standard also defines a significant number of requirements for reducing variations in the process that have arisen from the assumption that variations lead to costs and waste, which consequently leads to a waste of time and money, which is not acceptable in a very competitive industry such as automotive – time is money. Various tools such as SPC, MSA, and FMEA have been created with variation reduction as the ultimate goal and their use is required by IATF 16949 (Hoyle, 2000). The IATF states on its official website that as many as 90% of companies that have IATF 16949 certification meet the requirements of their customers (IATF, 2025). There are five key pillars of IATF 16949 (Hoyle, 2007; Sekulović & Čabarkapa, 2019):

- Advanced product quality planning (APQP);
- Production part approval process (PPAP);
- Failure modes and effect analysis (FMEA);
- Statistical process control (SPC);
- Measurement system analysis (MSA).

Despite the development of IATF 16949, the automotive industry has remained broadly divided into three independent spheres. The IATF 16949 standard defines the requirements for a quality management system in the automotive industry, but not the ways to implement them. Regional organizations have therefore developed a series of manuals and standards for various aspects such as risk analysis, the use of tools such as SPC, MSA, CP and FMEA, the way of implementing PPAP and process verification.

9 Research Methodology

In order to collect information on the target group's perception of the importance and awareness of the application of the standard for the quality management system – IATF 16949, an anonymous survey will be conducted. The target group is employees in the automotive supply chain, more precisely employees of a company engaged in the production of parts in the automotive industry. The survey will be distributed to employees in the sectors of production, development and design, quality planning, procurement, incoming control, logistics and administration, with the aim of gathering information on the relevant group.

The reason for distributing the questionnaire to different organizational sectors is to represent the diversity of understanding and understanding of the management system depending on the organizational position and experience of the respondent. The ultimate goal of the research is to analyze the understanding of the quality management system and the IATF 16949 standard by employees in an organizational system. The

research will enable an analysis of the level of awareness of engineers and workers in a factory belonging to the automotive supply chain, their attitude to the importance of these standards and product quality, as well as the overall maturity of the organization in relation to the applied quality management system.

The survey will be created in English, with the help of the "Google Forms" web application and distributed to respondents via email. In order to reach the most specific conclusions, a four-point scale was used in the survey to fill in the answers. The survey was modeled on already created surveys similar to the research area (Bravi et al., 2019; Hanzličková, 2019). The questions in the survey are listed in Table 2.

Table 2. Questions from the survey

No.	Question
1	In which role are you working currently?
2	How long have you been working in the automotive industry?
3	Do you know what a quality management system is?
4	What do you think, how important is a quality management system for the automotive industry in general?
5	What do you think, how important is the proper implementation and functioning of a quality management system for the quality of the final product?
6	How does a quality management system influence product quality?
7	Do you know what IATF 16949 is?
8	Can you please name some organizations that developed a standard for an automotive quality management system?
9	Do all organizations that are working in the automotive industry need to be certified according to the IATF 16949?
10	Is IATF 16949 different than ISO 9001?
11	Who can certify the organization in accordance with the IATF 16949?
12	What is your opinion, how much implementation (and certification) according to IATF 16949 influences the quality of the product?
13	Do you know what Customer Specific Requirements are?
14	Is IATF 16949 equally important for the car manufacturers and automotive suppliers?
15	What can be done to improve the effects of the IATF 16949?

10 Research Results

The survey was distributed on 11.09.2025 and responses were collected until 25.09.2025. The survey was distributed in electronic form to 90 e-mail addresses, of which 56.67% and 51 respondents responded to the survey. The distributed survey was completed by employees from different sectors with different levels of experience.

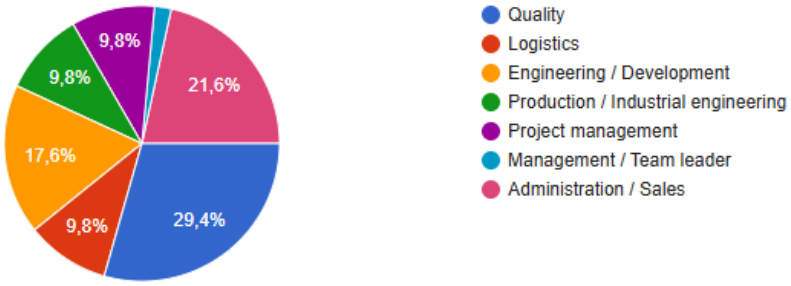


Fig. 5. Survey – field of work

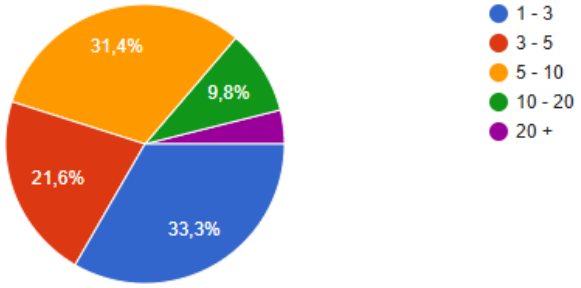


Fig. 6. Survey – years of work in the automotive industry

The largest number of respondents is employed in the quality sector, which includes various positions such as quality planners, internal auditors, quality engineers and laboratory technicians. Also, the largest number of respondents are at the beginning of their careers (1 to 3 years of work), but this group is almost equal to the group of more experienced people who have been employed in the automotive industry for 3 to 5 years.

A significant number of respondents did not know what the "Quality Management System" actually represents and what it consists of, and by analyzing the results in the table, it can be concluded that this result came from employees within the administration and sales sector, but also from the logistics sector. It can be concluded that employees who have no contact with production, development and quality as an engineering discipline are not adequately educated about the existence and importance of the organization's quality system. Although these employees "support" the basic process with their work, they must be aware of the importance of the quality system and its order. This failure could be attributed to the organization's management failing to initiate and plan adequate employee training for all employees.

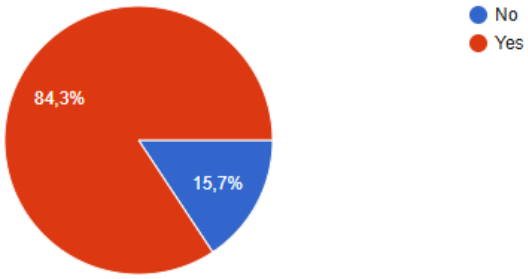


Fig. 7. Survey – Understanding the Quality Management System

When asked to name some of the standards organizations that develop standards for quality management systems for the automotive industry, respondents gave fairly uniform answers, which can be grouped as follows:

- A significant number of respondents indicated the IATF organization (standard) (IATF 16949);
- The majority of respondents who cited this standard also cited the organization VDA (VDA 6);
- A small number of respondents cited AIAG;
- A significant number of respondents did not give any answer;
- Also, some of the respondents cited the organization ISO and DIN.

The appearance of the VDA as a frequent response can be explained by the fact that the survey was conducted in an organization of German origin, whose customers are predominantly German companies that often refer to some of the VDA standards. Such organizations also supply so-called non-VDA customers (Ford, Stellantis, Renault), but the AIAG standard is usually only known to employees who have contact with this group of customers.

Interestingly, a significant number of respondents did not recognize the IATF 16949 standard as a standard for quality management systems in the automotive industry. To the question "Do you know what IATF 16949 is?", 31.4% of respondents gave a negative answer (Fig. 8).

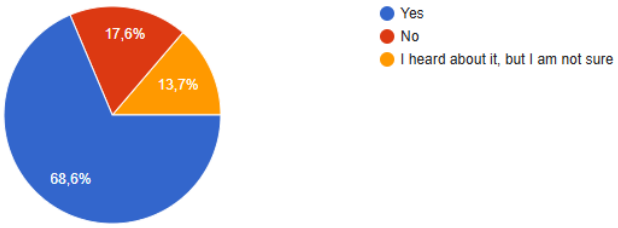


Fig. 8. Survey – Knowledge of IATF 16949

The number of respondents who do not know whether certification according to the IATF 16949 standard is necessary in the automotive industry exceeds the number of

respondents who do not know what this standard is, which means that even respondents who are aware of the importance of this standard are not sure whether its application is necessary (the use of the word mandatory would not be appropriate because it could be concluded that it is the law that "obliges" organizations to apply this standard, which is certainly not the case). As many as 22 respondents (over 40%) answered with "Do all organizations working in the automotive industry have to be certified according to IATF 16949?", and the same number answered that the application of the standard was imposed by commercial contracts and project documentation, which is certainly one of the correct answers (*of which there were several*). The second correct answer, that having a certificate is a prerequisite for applying to a customer's project, was chosen by fewer respondents. It was possible to choose multiple answers.

More than half of the respondents, 52.9%, did not know which organization could certify an organization according to the IATF 16949 standard. By reviewing the tabular presentation of the responses, it is concluded that even all employees in the quality sector did not know that certification can only be carried out by a body accredited for this purpose by the IATF office (representative). The project management staff did not provide 100% correct answers. A large number of respondents said "I don't know".

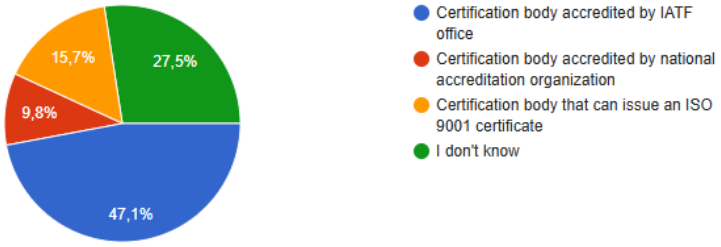


Fig. 9. Survey – IATF 16949 Certification

When asked whether the IATF 16949 standard is equally important for automakers and component manufacturers, as many as 33% of respondents answered "I don't know", and 63% think it is.

When asked about the impact of the implementation of IATF 16949 in an organization on the level of product quality, the majority of respondents stated that the implementation of this standard is important, but that there are many other factors that affect the level of product quality (Fig. 10).

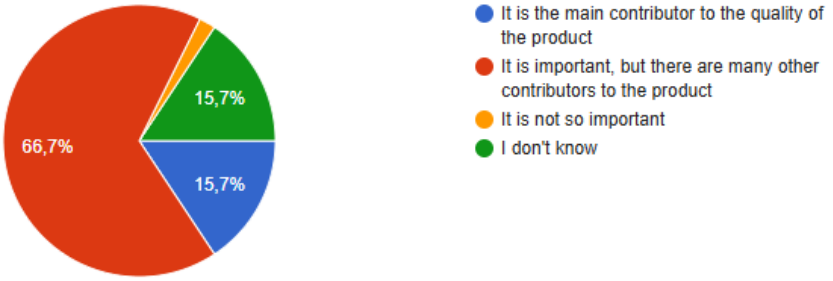


Fig. 10. Survey – Contribution of IATF 16949

11 Conclusion Based on Literature Review

By reviewing the relevant literature, it was first shown what constitutes a quality management system in general, and then which organizations develop standards in this area intended for the automotive industry. Understanding what constitutes a standard for a quality management system is essential for further understanding of the work. A quality management system standard is a set of guidelines and good practices for planning, managing and improving quality. Also, by showing the automotive industry through numbers, the importance of this industry to the global economy, but also to the daily lives of millions of people around the world.

By analyzing the various organizations that develop standards in the automotive industry, one can see the complexity of the functioning of the automotive industry and the "battle of standards" that takes place in this industry. During the 1980s and 1990s, various companies developed standards for quality management systems based on the ISO 9001 standard in an effort to better master the quality of products throughout the supply chain, thus creating great pressure on suppliers who had to meet the requirements of a large number of customers. With the development of IATF 16949 (originally ISO/TS 16949), for the first time, a universally accepted specification (standard) was created that adapted the requirements of ISO 9001 to the needs of the automotive industry, introduced new concepts such as PPAP, FMEA and others in order to more efficiently achieve the improved effect sought by car manufacturers.

However, with the development of the IATF 16949 standard, the automotive industry has not established a universality of interpretation of the requirements for the implementation of this standard, thus leaving room for the continuation of the polarization of the industry, which is very noticeable today. In the field of standardization of quality management systems, the greatest influence has the regional organizations for standardization AIAG and VDA, which literally divide the automotive industry into two parts, thus challenging the idea that officially stands behind the establishment of these organizations – achieving synchronization of requirements in the industry and reducing diversity. The greatest progress towards achieving this goal is the cooperation on the development of the FMEA manual, which is fully harmonized by the two organizations. The struggle of standards and the aspiration of car manufacturers to prevail in the world market is also seen through the existence of Customer Specific Requirements (CSRs),

which often refer to the standards of the specific association to which the car manufacturer belongs or independently developed standards, which must be applied in addition to IATF 16949 if the supplier wants to do business with the customer.

12 Possibility for Further Theoretical Research

In order to fully understand the global system of standardization of quality management systems in the automotive industry, it would be necessary to theoretically analyze the Japanese and Korean automotive industries in more detail, as well as the increasingly powerful and important in the global market – the Chinese automotive industry and the relevant organizations of these countries. This information is somewhat more difficult to come by, especially information about the Chinese automotive industry. Also, there is room for a deeper analysis of the requirements of the standards and manuals developed by the aforementioned standardization organizations – IATF, VDA and AIAG. A comparative analysis of the standards developed by the VDA, AIAG and IATF groups with the aim of finding similarities and differences in the many standards and manuals could also make a significant contribution.

Analysis of technical standards and their categorization may also be of interest. There are countless standards that prescribe the technical specifications of car parts or the car itself, as well as tests to validate the quality of components and the final product. In this area, ISO and IEC play a dominant role. Standards for chargers for electric vehicles are of great importance for the automotive industry and can practically shape its course in the future in view of the increased tendency towards the electrification of vehicles. By the same logic, it would be possible to consider standards for propellants, tyres and spare parts, and even other infrastructure and signalling standards that are extremely important for the global functioning of transport and transport in general. There is also a wide range of standards developed for the purpose of assessing vehicle safety, which can also be very interesting to explore and consider the similarities and differences between standards developed by different organizations dealing with this area. A very important aspect of the automotive industry is the so-called mandatory standardization, i.e. standardization that arises from regulations and laws imposed by associations of countries or countries individually. In one of the titles of this paper, regulations are mentioned as an important aspect, with a focus on European industry.

13 Conclusion Based on Research Results

First of all, it is necessary to note certain limitations of the conducted research and opportunities to improve it. As the research was conducted within one organizational system on a limited number of respondents, and in a short period of time, it is not possible to generalize the presented conclusions for the automotive industry globally. The conducted research can be a good basis for further, more comprehensive research on the perception of the IATF and the IATF 16949 standard, beyond the borders of one organizational system or one country. The company in which the research was conducted is a relatively young organizational system and relatively negative conclusions

might be different in another more mature organizational system, which can also be an interesting direction of research.

By analyzing the results of the conducted research, it can be clearly concluded that depending on the organizational sector to which the employee belongs, the knowledge and perception of the standards for quality management systems differ. Employees in the administration sector have significantly less knowledge about the standard and its application compared to their colleagues in the quality sector. The survey also confirmed the aforementioned persistent polarization of the automotive industry – as the survey was conducted in a company of German origin, employees were significantly more informed about the standards of the VDA organization, while AIAG was rarely mentioned. There is also a lack of a deeper understanding of IATF 16949, as even employees who know what this standard is and what it is intended for are not sure how its implementation is imposed, who can carry out the certification, and what exactly is the difference from ISO 9001. These failures can be attributed to inadequate staff education, isolation of employees within their processes without understanding the organization as a system, inadequate employee training, and lack of internal mechanisms for monitoring employee knowledge. In order to raise employees' awareness of the quality management system and its standards, the organization, as the most effective solution, can hire internal experts in order to implement trainings, the complexity and detail of which may depend on the sector in which the training is conducted. A simpler but more expensive solution may be to hire an external organization specializing in education on QMS and other topics. By organizing this training, employees will not only be provided with additional information, but in this way it will increase the level of awareness of employees about their role in the organization and deepen the understanding of the organization's processes and thus improve the work of employees.

References

1. ACEA: About ACEA (2025), https://www.acea.auto/about-acea/, last accessed 2025/08/25.
2. AIAG: About AIAG (2025), https://www.aiag.org/about, last accessed 2025/08/25.
3. AIAG: AIAG Quality (2025), https://www.aiag.org/quality, last accessed 2025/08/25.
4. Binder, A., Rae, J.: Automobile industry. Encyclopedia Britannica (2024), https://www.britannica.com/technology/automotive-industry/Ford-and-the-assembly-line, last accessed 2025/08/20.
5. Bozola, P.M., Nunhes, T.V., Barbosa, L.C.F.M., Machado, M.C., Oliveira, O.J.: Overcoming the challenges of moving from ISO/TS 16949 to IATF 16949: recommendations for implementing a quality management system in automotive companies. Benchmarking: An International Journal 30(9), 3699–3724 (2023)
6. Bravi, L., Murmura, F., Santos, G.: The ISO 9001:2015 quality management system standard: Companies' drivers, benefits and barriers to its implementation. Quality Innovation Prosperity 23(2), 64–82 (2019)
7. Cagnin, F., Oliveira, M.C.D., Cauchick Miguel, P.A.: Assessment of ISO 9001:2015 implementation: focus on risk management approach requirements compliance in an automotive company. Total Quality Management & Business Excellence 32(9–10), 1147–1165 (2021)

8. Christopher, M.: Logistics and Supply Chain Management. Irwin Professional Publishing, London (1994)
9. CLEPA: About CLEPA (2025), https://clepa.eu/who-we-are/, last accessed 2025/08/25.
10. Cromer, O., Purry, K.: Automobile. Encyclopedia Britannica (2025), https://www.britannica.com/technology/automobile, last accessed 2025/08/19.
11. De Vries, H.: Standardization: A Business Approach to the Role of National Standardization Organizations. Kluwer Academic Publishers, Boston (1999)
12. DIN: DIN Standards (2025), https://www.din.de/en/about-standards/din-standards/, last accessed 2025/09/01.
13. Đošić, A., Miletić, K., Đurić, M.: Overview of the German standard VDA 6 for quality management systems in the automotive industry. Proceedings of the International Congress on Process Industry – Processing 32(1), 391–397 (2019)
14. European Automobile Manufacturers' Association (ACEA): New car registrations 2024 (2024), https://www.acea.auto/pc-registrations/newcar-registrations-1-1-n-october-2024-year-to-date-batteryelectric-sales-4-9, last accessed 2025/08/18.
15. European Automobile Manufacturers' Association (ACEA): The automobile industry pocket guide 2024/2025 (2025), https://www.acea.auto/publication/the-automobile-industry-pocket-guide-2024-2025/, last accessed 2025/08/18.
16. Fahmi, K., Mustofa, A., Rochmad, I., Sulastri, E., Wahyuni, I.S., Irwansyah, I.: Effect of ISO 9001:2015, ISO 14001:2015 and ISO 45001:2018 on operational performance of automotive industries. Journal of Industrial Engineering & Management 2(1), 13–25 (2021)
17. Filipović, J., Đurić, M.: The Basics of Quality. Faculty of Organizational Sciences, Belgrade (2009)
18. Ford: The Model T (2024), https://corporate.ford.com/articles/history/the-model-t.html, last accessed 2025/08/19.
19. Ford, H.: Henry Ford on mass production. Encyclopedia Britannica (2024), https://www.britannica.com/money/Henry-Ford-on-mass-production, last accessed 2025/08/19.
20. Goicoechea, I., Fenollera, M.: Quality Management in Automotive Industry. DAAAM International, Vienna (2012)
21. Hanzličková, D.: Purchasing tool ensuring the conformity of ISO 9001 and base automotive requirements along the supply chain with focus on tier 2 and tier 3 suppliers. Doctoral dissertation, Wien (2019)
22. Hoyle, D.: Automotive Quality Systems Handbook. Elsevier, Oxford (2000)
23. Hoyle, D.: Quality Management Essentials. Elsevier, Oxford (2007)
24. IATF: About IATF (2025), https://www.iatfglobaloversight.org/about-iatf/, last accessed 2025/08/25.
25. IATF: About IATF Oversight Offices (2025), https://www.iatfglobaloversight.org/about-iatf-oversight-offices/, last accessed 2025/08/25.
26. International Automotive Task Force: IATF 16949:2016 – Quality management system requirements for automotive production and relevant service parts organizations. IATF (2016)
27. ISO: About ISO (2025), https://www.iso.org/about, last accessed 2025/08/25.
28. ISO: ICS – Road vehicles engineering (2025), https://www.iso.org/ics/43/x/, last accessed 2025/08/25.
29. ISO: ISO/TS 16949 (2025), https://www.iso.org/standard/52844.html, last accessed 2025/08/25.

30. Janjić, G., Tanasić, Z., Tomić Stojković, R.: Development of a quality management model in the automotive industry with a focus on the quality of incoming products (Part I). Transactions of FAMENA 46(1), 115–126 (2022)
31. Jaško, O., Čudanov, M., Jevtić, M., Krivokapić, J.: Fundamentals of Organization and Management. Faculty of Organizational Sciences, Belgrade (2014)
32. Laskurain-Iturbe, I., Arana-Landín, G., Heras-Saizarbitoria, I., Boiral, O.: How does IATF 16949 add value to ISO 9001? An empirical study. Total Quality Management & Business Excellence 32(11–12), 1341–1358 (2021)
33. Leggett: Pressing issues for automotive supply chains. Just Auto (2024), https://www.just-auto.com/features/pressing-issues-for-automotive-supply-chains/, last accessed 2025/08/28.
34. Mentzer, J.T., DeWitt, W., Keebler, J.S., Min, S., Nix, N.W., Smith, C.D., Zacharia, Z.G.: Defining supply chain management. Journal of Business Logistics 22(2), 1–25 (2001)
35. Mijatović, I.: Standardization 1. Faculty of Organizational Sciences, Belgrade (2015)
36. Petkova, P., Petkova, M., Jekov, B., Kovatcheva, E.: Literature review of key performance indicators for supplier quality management in automotive electronics industry. In: 2023 8th International Conference on Smart and Sustainable Technologies (SpliTech), pp. 1–5. IEEE (2023)
37. Sekulović, A., Čabarkapa, J.: Standardization and automotive: Transition to IATF 16949:2016 standard in the automotive supply chain. Proceedings of the International Congress on Process Industry – Processing 32(1), 373–383 (2019)
38. United Nations Economic Commission for Europe (UNECE): Vehicle Regulations (2025), https://unece.org/transport/vehicle-regulations, last accessed 2025/09/01.
39. United Nations Industrial Development Organization (UNIDO): Role of standards – A guide for small and medium-sized enterprises. United Nations, Vienna (2006)
40. VDA QMC: About the VDA 6.X standards (2025), https://vda-qmc.de/en/vda-zertifizierungen/vda-6-x/, last accessed 2025/08/25.
41. VDA QMC: VDA Certifications (2025), https://vda-qmc.de/en/vda-zertifizierungen/, last accessed 2025/08/25.
42. VDA QMC: Welcome to the VDA QMC (2025), https://vda-qmc.de/en/ueber-uns/, last accessed 2025/08/25.
43. Warmuz, K.: Evolution of the automotive industry (2025), https://makolab.com/insights/understanding-the-automotive-industry, last accessed 2025/08/20.
44. Waseem, M., Yusoff, Y.M.: The effect of total quality management practices on supply chain performance in the automobile industry. Multidisciplinary Science Journal 7(2), 2025077 (2025)
45. Yılmaz, A., Acaroglu, C.: IATF 16949 certification of a company in the automotive sector. Turk Journal Park System, Istanbul (2024)
46. Živković, N., Glogovac, M.: Quality Management. Faculty of Organizational Sciences, Belgrade (2015)
47. Živković, N., Glogovac, M.: Integrated Management Systems – Authorized Materials from Lectures and Exercises. Faculty of Organizational Sciences, Belgrade (2022)

Original Research Article *https://doi.org/10.5281/zenodo.23477117*

Closing the Governance Gap: Why ISO/IEC 42001 Completes the ISO Management System Trilogy

Nikola Vojtek[1][0000-0001-9380-7046] and Ivana Mijatović[2][0000-0002-7624-0356]

[1] Smart Consulting Agency, Belgrade, Serbia
[2] University of Belgrade, Faculty of Organizational Sciences, Belgrade, Serbia
nikola@smartconsulting-agency.com

Abstract. The adoption of artificial intelligence is significantly increasing in almost all industries, and companies are racing to be the leaders in this area. Management systems standards provide a structured framework for organizations to ensure consistency, control risks, and align processes with strategic objectives in rapidly evolving environments. AI could provide a lot of benefits for the companies in this context, but it could also introduce risks that are not completely addressed by existing management system standards. While ISO 9001 focuses on quality management and ISO/IEC 27001 on information security, both lack proper mechanisms to govern AI-specific challenges such as model behavior, data bias, and autonomous decision-making. This paper positions ISO/IEC 42001 as a necessary extension that could minimize and close the governance gap. Through a comparative analysis of ISO 9001, ISO/IEC 27001, and ISO/IEC 42001, the study highlights differences and shows how these standards together form a complementary framework and guidelines for modern teams and companies. The paper further outlines how ISO/IEC 42001 integrates with existing management systems and supports the implementation of structured AI governance. It concludes by emphasizing the importance of ISO/IEC 42001 for ensuring trust, accountability, and responsible use of AI in increasingly complex environments.

Keywords: Artificial Intelligence, ISO/IEC 42001, AI Governance.

1 Introduction

The rapid advancement of artificial intelligence is significantly transforming almost all industries. Companies are increasingly investing substantial resources to integrate AI capabilities into their operations, aiming to reduce costs, enhance efficiency, and improve decision-making accuracy. In recent years, these efforts have led to profound shifts across industries, driven by the deployment of AI models and the widespread adoption of AI-based tools.

Considering the AI impact on processes and business, the need for structured governance of AI-related initiatives is becoming critical. Artificial intelligence could provide significant benefits to the companies, but on the other hand, it is also introducing a

group of completely new risks for business and operations. A lot of companies are participating in the race to develop and implement AI, but usually the proper cost vs benefit case is not defined properly, or it is in favor of costs [1]. The impact of these risks could be reduced significantly through system approach, regulations [2] and structured governance frameworks [3].

Traditionally, companies are mostly focused on quality management (ISO 9001) and information security (ISO/IEC 27001), but these systems and frameworks lack proper mechanisms to govern AI-specific risk such as model behavior, data bias, and autonomous decision-making. First mentioned system is allowing companies to focus on defining what quality means for them, what are mechanisms in place for quality assurance, how to achieve consistency in processes, introduce and maintain continuous improvement aspect. When it comes to the security of information, following guidelines of ISO/IEC 27001, companies define and apply specific controls based on different levels of information importance and impact. Artificial intelligence is introducing different aspects that are not covered with the first two systems, such as dynamic models, probabilistic outputs, data-driven bias, overfitting, as well as data scarcity [4].

Considering the limitations of the management systems in the context of risks introduced in the last several years by AI, it is clear that existing frameworks need to be expanded and extended to properly embrace the artificial intelligence domain. ISO/IEC 42001 completes the ISO Management System trilogy by expanding focus on the governance of AI systems, highlighting the risk-based approach, as well as introducing transparency, accountability and controls over data and models. The third standard is providing guidelines for companies to deal with uncertainty and probabilistic behavior. It fits very well into the processes and security aspects, making it ideal candidate for the integrated system, in case when company is characterized with the high rate of AI adoption.

The contribution of this paper is related to the new integrated governance framework that operationalizes the relationship between ISO 9001, ISO 27001, and ISO/IEC 42001. The framework connects three dimensions: processes, roles and responsibilities, and control mechanisms. While existing studies and standards primarily address quality management, information security, or AI governance independently, this research introduces a structured comparative approach that demonstrates how traditional management systems can be extended to support governance of adaptive AI-driven environments. In addition to the comparative analysis, the paper contributes by identifying governance gaps that emerge when companies adopt AI systems without corresponding expansion of existing management structures and controls.

This paper aims to provide a comparative analysis of ISO 9001, ISO/IEC 27001, and ISO/IEC 42001, to highlight differences and shows how these standards together form a complementary framework for teams and companies. In section two, a literature review was provided. ISO/IEC 42001 AI Management System (AIMS) core elements of

the standard were described in section 3. Following that, comparative analysis is provided in the section 4, and integrated framework was proposed in section 5. Practical Implications for Organizations, conclusion and future work were provided respectfully in sections 6 and 7.

2 Literature Review

AI has been considered with a strategic importance for the companies, with a strong impact on economy and business. Expectations are growing, and companies are expecting to improve productivity with the implementation of artificial intelligence [1]. There is a lot of uncertainty within AI initiatives, thus it is very hard and it could be unclear what would be the ROI at the end [5].

By either developing AI models, or using AI-based tools, there are a lot of risks, data and ethical challenges that companies are facing [6]. Companies usually have established, even before AI, some specific principles that will take into consideration the ethical aspects, but as highlighted further in [7], principles alone cannot guarantee ethical AI. There is a need for mechanism that will ensure transparency across models and data, as well as accountability. During the AI lifecycle, there is a lot of uncertainty that will push companies to make decisions that are not straightforward [8]. On the other side, for the AI models to be trained, a lot of data will need to be collected and provided. In some situations, companies could reach towards synthetic data generation, to solve them the problem of data scarcity [9]. But that would need to be properly documented, since it could significantly impact on the results provided by the model. On the other hand, there could be the situation with the biased data, that could result in the false calibration of the model [10].

A lot of discussions are happening recently on the global level by the different institutions and bodies, related to the AI governance. It is clear that establishing a proper management system for artificial intelligence is becoming a necessity [2]. Existing risk management mechanisms and processes are no longer relevant and are failing to capture AI-related risks without expansion [11]. Companies are exposed to the cyber threats, the IT systems are becoming more vulnerable, as well as the software. This is due to the software architecture, mechanisms for storing the data, and the way how two systems communicate in a digital world.

Traditional governance and risk management models are characterized by the process and security-based approaches. Without expansion in a proper direction, they are becoming obsolete. In [12] authors proposed a unified ethical framework for artificial intelligence, that is combining existing guidelines into five major principles. According to their paper, these principles are beneficence, non-maleficence, autonomy, justice, and explicability. With the proposed unification, they managed to clearly highlight the limitations of the fragmented and traditional approaches.

From both academic and practical views, there is a lack of initiative towards the integrated and unified AI governance approach. Some of the reasons could be related to the uncertainty that AI is introducing, as well as since currently AI governance is split across security, regulations, data governance and ethics. Furthermore, artificial intelligence is interdisciplinary, which means that it combines software engineering, data science, business processes and strategy. Finally, there is a lack of comparative analysis conducted in this area, that could highlight the core differences based on which specific action items for the implementation could be generated. This also represents one of the major motivations for the authors for tackling this topic, and proposing the integrated framework based on the comparative analysis results.

To address the identified research gap, this paper applies a qualitative comparative analysis approach based on a criteria-driven analytical framework for system management. The comparative analysis was conducted with the focus on governance, AI lifecycle coverage, risk orientation, control mechanisms and organizational impact. These criteria were derived from management systems available literature and AI governance principles.

3 ISO/IEC 42001 – AI Management System (AIMS)

ISO/IEC 42001 is a first management standard for the artificial intelligence, and it was introduced as a response to the increased AI adoption and risks that are emerging in this area for companies. Its focus is on the AI governance, and the implementation of the controls that ensure transparency and accountability across business operations. Through the different lifecycle phases such as planning, development, deployment and monitoring, it helps companies to establish proper checkpoints to ensure reliable management system for AI.

Artificial intelligence Management System (AIMS) represents a structured framework for governing the lifecycle of the AI systems. It ensures that AI-related risks are properly managed, that model and dataset versioning is in place together with change history, and defines specific set of responsibilities across organizational roles involved in the AI lifecycle. AIMS is designed in a way to be easily integrated into existing processes and aligned with existing management systems.

Overview of the main elements and aspects of AIMS is provided below in the following figures.

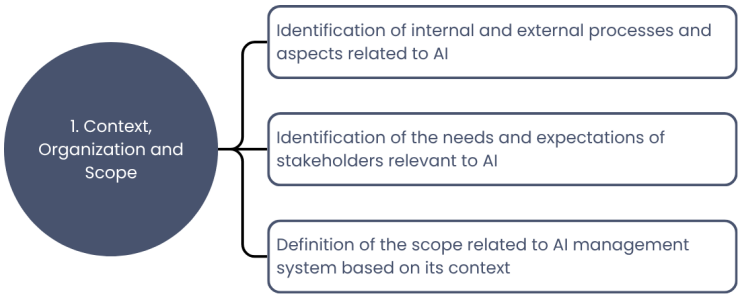


Fig. 1. Context, Organization and Scope

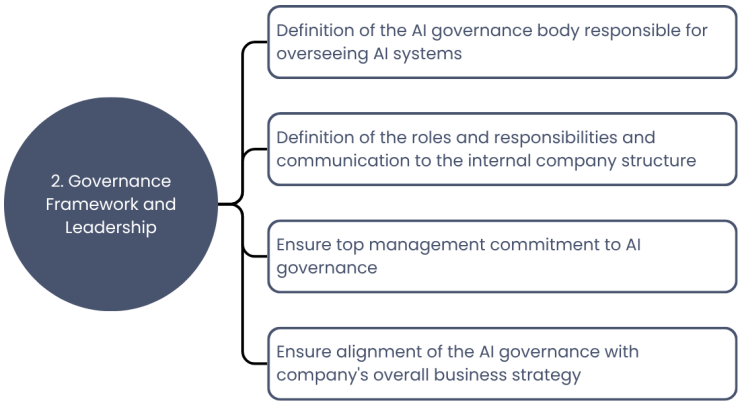


Fig. 2. Governance Framework and Leadership

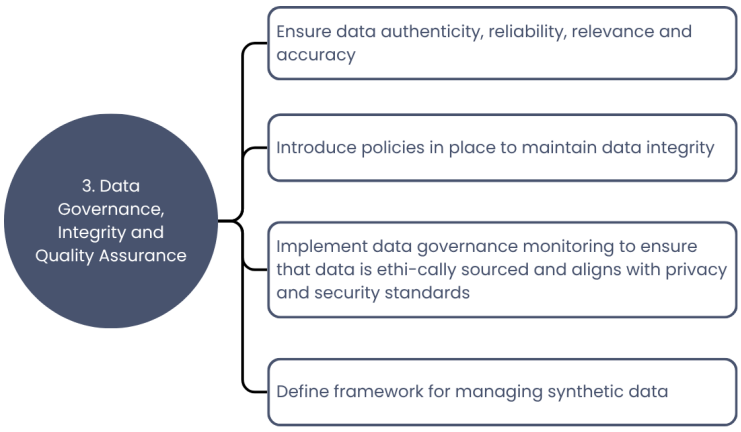


Fig. 3. Data Governance, Integrity and Quality Assurance

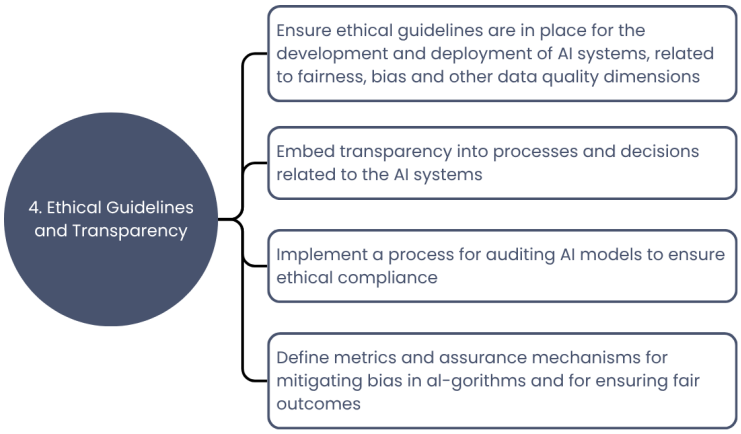


Fig. 4. Ethical Guidelines and Transparency

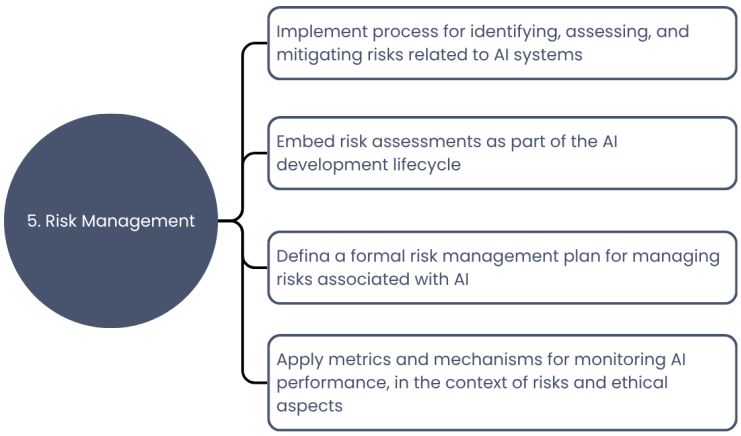


Fig. 5. Risk Management

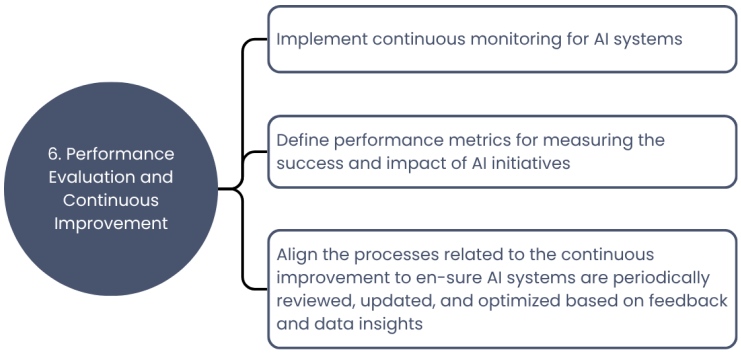


Fig. 6. Performance Evaluation and Continuous Improvement

AIMS is ensuring that core aspects of the AI governance are managed properly. These aspects are further examined in the context of key capabilities that company needs to implement, and provided in the next table.

Table 1. ISO/IEC 42001 Aspects, Principles and Capabilities

Aspect	Principle	Capability
Risk	Manage AI system based on the risk level	Identify, assess and mitigate AI risks
Transparency and Explainability	Make sure to understand AI related decision	Document model and data changes and decisions
Accountability and Responsibility	Define clear ownership of AI systems	Assign roles and responsibilities
Lifecycle Management	Define end-to-end processes for AI governance	Establish controls from data collection, through model development, deployment and monitoring
Continuous Monitoring and Improvement	Continuously evaluate evolving AI system	Implement proper mechanisms for performance tracking, model updates and risk reassessment

AIMS differs from the traditional management systems, due to its needs to address different challenges. Significant amount of uncertainty is present in the multiple stages of the model development and data preparation. Companies are facing data scarcity which is leading them towards relying on synthetically generated data. Also, models are evolving constantly, generating the probabilistic outputs, which further increase the complexity in business and operations.

4 Comparative Analysis

In the previous section, we already highlighted that each of the standards for management systems address different aspects. ISO 9001 is focused on quality management, processes and consistency. ISO 27001 is concentrated around information security and data protection. Finally, ISO/IEC 42001 is considering AI governance, models, data, transparency and decisions.

We are starting the comparative analysis with key aspects provided in the next table.

Table 2. Comparative analysis of key aspects of ISO 9001, ISO 27001 and ISO/IEC 42001

Aspect	ISO 9001	ISO 27001	ISO/IEC 42001
Primary Focus	Quality and Processes	Information security	AI governance
Type of System	Process driven, deterministic	Security-oriented, data-centric	Adaptive, probabilistic, data-driven
Risk Focus	Processes and operational risks	Security and data-related risks	AI-specific risks
Lifecycle Coverage	Process lifecycle	Information lifecycle	AI lifecycle (model, data)
Decision-making	Human-driven	Secured and data usage related	AI-assisted or AI-driven
Outputs	Predictable and repeatable	Controlled and secured	Probabilistic and evolving
Adaptability	Low (stable processes)	Moderate (security updates)	High (models evolving)
Key challenge	Process inefficiency	Cyber threats and data breaches	Lack of control and transparency over intelligent systems

The evolution of the management system is clearly seen from this table, starting from the processes, through the data and finalizing with intelligence. Every aspect is shifting towards the artificial intelligence, and adaptive governance structures capable of managing uncertainty and autonomous deci-sion-making. While ISO 9001 and ISO 27001 primarily operate within predictable boundaries, ISO/IEC 42001 addresses environments characterized by probabilistic outputs, evolving models, and continuous learning mechanisms. This has some major organizational implications. Traditional governance mechanisms are generally designed around static processes and predefined controls, whereas AI systems require continuous monitoring, dynamic reassessment of risks, and governance mechanisms capable of responding to changing model behavior over time. As a result, companies that are adopting AI cannot rely solely on conventional quality and security management approaches, since these systems were not originally designed to govern adaptive intelligent systems.

Next step is to compare the main sections in related standards, as provided in the next table.

Table 3. Comparison of main sections in standards ISO 9001, ISO 27001 and ISO/IEC 42001

Main Sections	ISO 9001	ISO 27001	ISO/IEC 42001
Scope	Quality of products and services	Information assets protection	Governance and lifecycle of AI systems
Context of the Organization	Understanding external and internal factors affecting quality	Define information security context	Idenfity AI-related context and usage
Leadership	Focus on customer and committed to quality management	Committed to the information security policies	Committed to AI governance and responsible AI usage
Planning	Quality objectives definition	Risk identification and treatment definition	AI-related risks identification, impact assessment, and mitigation
Resources and Support	Competence, awareness, communication	Training and awareness	AI competencies, data management, model documentation
Operation	Process control and delivery	Security controls implementation	AI-related controls implementation, system design, development, deployment and monitoring
Performance evaluation	Internal audits	Security monitoring, internal audits, incident evaluation	AI model performance evaluation, bias detection
Improvement	Continuous improvement of processes	Continuous improvement of security mechanisms	Continuous improvement of governance mechanisms and AI systems
Control mechanisms	Standardization of processes	Security controls	AI-related controls and transparency

Looking into the comparison of the key aspects and main standard sections, it is clear that, together, these standards form a foundation for the unified governance model that will cover processes, data, and intelligent decision-making. Unlike ISO 9001 and ISO 27001, which primarily focus on process standardization and information security, ISO/IEC 42001 introduces governance requirements associated with explainability, model transparency, ethical considerations, and continuous evaluation of AI lifecycle.

5 Closing the Governance Gap: Integrated Framework

With the aim of closing the governance gap, integrated framework is provided in this section. It consists of the expansion of the three key management system dimensions:

- processes,
- roles and responsibilities, and
- controls.

First, we are considering the expansion of the existing processes within the traditional management systems, which is provided in the next table.

Table 4. Integrated Framework: Expansion of the existing processes

Process area	ISO/IEC 42001 AIMS expansion
Planning	Include in the planning: AI-specific use cases Assess AI risks together with other risks
Design & Development	Design, train, validate, and test the AI models
Data Management	Data quality assurance, bias detection procedure, dataset governance, synthetic data handling policy
Operations	Consider AI model deployment, traceability of any decisions being made
Monitoring & Evaluation	Monitoring of AI system behavior, Model performance tracking
Continuous Improvements	Define steps for Model retraining and continuous AI risk reassessment

Next, we are considering in which ways companies need to expand their current roles within the traditional management systems in the context of adding specific responsibilities related to AI governance. Proposed solution is provided in the next table.

Table 5. Integrated Framework: Expansion of the existing roles

Role	ISO/IEC 42001 AIMS expansion
Top Management	Governance oversight and AI-specific policies Ensure responsible AI usage and accountability AI governance lead responsibilities
Process owners	Identify AI aspects in processes Validate AI-supported decisions and conduct reviews from ethical perspective
IT / Security Team	Ensure secure AI infrastructure is in place
Risk / Compliance Officers	Assess and manage AI risks Ensures regulatory, ethical, and governance compliance
AI-Specific Roles	Model owner - Responsible for maintaining model integrity and data quality assurance

Framework is completed by extending the existing controls that are in place for the ISO 9001 and ISO 27001 with more specific AI-related controls from the AIMS. Proposed mapping of the controls is provided in the next table.

Table 6. Integrated Framework: Expansion of the existing controls

Control	ISO/IEC 42001 AIMS
Process	Expand Standardized procedures and audits with Adaptive control of AI processes and model lifecycle
Security	Expand Access control, data protection with mechanisms for protecting the training data, models, and AI pipelines
Risk	Expand risk controls with AI-specific risk controls (bias, model overfitting, unintended outcomes)
Performance	Update existing KPI tracking, process metrics and information radiators with model accuracy, robustness, fairness, and performance monitoring
Transparency and Traceability	Expand documentation and records wherever it is appropriate with explainability, decision traceability, and model documentation
Continuous improvement	Update process optimization procedure to include regular model retraining and recalibration activities, as well as continuous governance updates

The expansion of traditional controls with AI-specific governance mechanisms demonstrates that existing management systems are insufficient when applied to intelligent systems without adaptation.

At the end, it could be concluded that these three expanded dimensions are enabling companies to fully operationalize the integrated governance framework, and maximize the use of the existing traditional management systems in the context of the AI governance.

6 Practical Implications for Companies

The purpose of the integrated framework is to expand the management system structure into three directions: (1) processes, (2) roles, and (3) controls. Companies should not drastically change their existing processes. Instead, updates should be carefully considered to incorporate necessary elements that enable the proper AI governance within processes. This initiative should not focus primarily on updating the documentation, but it should consider specific activities that are impacted by AI aspects.

In the first step, AI-related initiatives will be properly integrated into existing processes. Moving further with the updates, companies should review their existing roles and responsibilities. The required changes will depend on the usage of AI (context), industry, domain, organizational structure and company-specific factors. Mid-management should be playing a critical rol, and act as a bridge between strategic objectives, operational initiatives and execution.

After processes are updated accordingly, and roles are expanded with AI-related responsibilities, the third step is for a company to define and implement control mechanisms. Majority of organizations have already in place some kind of a risk management process, as well as security checks and established mechanisms for the performance measurement and metrics. Thus, updates should be related to the monitoring of model performance, bias detection, data quality validation, as well as ensuring that every AI-driven decision is explained adequately. Also, controls should be aligned with continuous evaluation of the AI system.

6.1 Example of Integrated Governance in Practice

In this section, we are presenting one example of implementing additional controls into the existing processes for the IT company in the biometric industry, developing software for digital identity proofing. The use case is limited to the controls provided in the next table.

Table 7. Example of Integrated Governance in Practice

Area	Existing Control	AI Governance GAP	ISO/IEC 42001 Control
OCR Acuracy	QA testing	No model drift monitoring	AI model performance monitoring
Data Management	Encryption/access control	No bias detection	Data audits and evaluation
Release Process	Change management	No retraining governance in place	Model lifecycle monitoring

As indicated, in this scenario related to the OCR accuracy of the model, company was performing regular testing activities, by a proper AI model performance monitoring in production needed to be put in place in order to detect potential model drift on time. Similarly, when it comes to the data management, basic controls were related to encryption and access approvals to the data set, but a proper periodical audits and evaluation of data needed to be established. Finally, the release process itself was organized without taking into consideration model lifecycle monitoring activities, which was updated with the appropriate and corresponding control.

7 Conclusion and Future Work

Development of artificial intelligence is significantly transforming almost all industries, and companies are increasingly investing substantial resources to integrate AI capabilities into their operations. Main objectives are related to cost reduction, improvement in the efficiency and effectiveness of operations and achieving better accuracy in decision-making. AI is introducing significant risks due to its impact on processes and business, thus a need for structured governance of AI-related initiatives is becoming a necessity.

While ISO 9001 focuses on quality management and ISO/IEC 27001 on information security, both lack proper mechanisms to govern AI-specific challenges such as model behavior, data bias, and autonomous decision-making. To properly embrace the artificial management domain, existing frameworks need to be expanded. ISO/IEC 42001 completes the ISO Management System trilogy by expanding focus on the governance of AI systems, highlighting the risk-based approach, as well as introducing transparency, accountability and controls over data and models.

This paper aims to provide a comparative analysis of ISO 9001, ISO/IEC 27001, and ISO/IEC 42001. It highlights the differences and shows how these standards could be combined together to a complementary framework for teams and companies involved in artificial intelligence initiatives. Based on the analysis, an integrated framework is proposed for expanding the management system structure into three directions, which are processes, roles, and controls. Proposed framework is enabling companies to fully operationalize the integrated governance framework, and maximize the use of the existing traditional management systems in the context of the AI governance. Finally, it is important to neglect that these extensions are focused not just on technical aspects, but primarily on the organizational aspects.

Potential future work should consider empirical validation of the proposed framework, additional extension towards the AI regulations such as EU AI Act, implementation specifics for agile teams, as well as introduction of automation in certain aspects of the AI-extended processes and controls.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Brynjolfsson, E., Rock, D., Syverson, C.: Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics. In: Agrawal, A., Gans, J., Goldfarb, A. (eds.) The Economics of Artificial Intelligence, pp. 23–57. University of Chicago Press (2019)

2. European Parliament and Council: Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union 2024(1689), 1–144 (2024)
3. International Organization for Standardization: ISO/IEC 42001:2023 Artificial Intelligence Management System – Requirements. ISO, Geneva (2023)
4. Vojtek, N., Smudja, B., Milošević, P., Dragović, I.: Leveraging generative adversarial networks in project management for task estimation. In: Proceedings of the 52nd Symposium on Operational Research – SYM-OP-IS 2025, Palić, Serbia, 7–10 September 2025, pp. 400-405 (2025)
5. Bughin, J., Seong, J., Manyika, J., Chui, M., Joshi, R.: Notes from the AI frontier: Modeling the impact of AI on the world economy. McKinsey Global Institute, 1–40 (2018)
6. Jobin, A., Ienca, M., Vayena, E.: The global landscape of AI ethics guidelines. Nature Machine Intelligence 1(9), 389–399 (2019)
7. Mittelstadt, B.: Principles alone cannot guarantee ethical AI. Nature Machine Intelligence 1(11), 501–507 (2019)
8. Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., Chaudhary, V., Young, M., Crespo, J.-F., Dennison, D.: Hidden technical debt in machine learning systems. Communications of the ACM 66(8), 74–83 (2023)
9. Frid-Adar, M., Klang, E., Amitai, M., Goldberger, J., Greenspan, H.: Synthetic data augmentation using GAN for improved liver lesion classification. IEEE Transactions on Medical Imaging 38(3), 677–685 (2019)
10. Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., Galstyan, A.: A survey on bias and fairness in machine learning. ACM Computing Surveys 54(6), 1–35 (2021)
11. Radanliev, P., De Roure, D., Walton, R.: Artificial intelligence and cyber risk: A new approach for managing cyber risk. Journal of Cyber Policy 5(1), 1–19 (2020)
12. Floridi, L., Cowls, J.: A unified framework of five principles for AI in society. Harvard Data Science Review 1(1), 1–15 (2019)

Original Research Article *https://doi.org/10.5281/zenodo.23477145*

Context and Identity in higher education institutions in Industry 5.0

Nedeljka Živković[1][0000-0001-6663-7467], Isidora Milošević[2][0000-0001-5539-8253], Ana Kićanović[3][0000-0002-1924-844X] and Olesea Plotnic[4][0000-0001-9368-7806]

[1] Belgrade Business and Arts Academy of Applied Studies, Kraljice Marije 73, 11000 Belgrade, Republic of Serbia
[2] University of Belgrade, University of Belgrade, Technical Faculty in Bor, Vojske Jugoslavije 12, 19210 Bor, Republic of Serbia
[3] University of Belgrade, Faculty of Organizational Sciences, Jove Ilića 154, 11000 Belgrade, Republic of Serbia
[4] Moldova State University, Faculty of Law, 67 M. Kogălniceanu Street, Block 2, MD-2009, Chisinau, Republic of Moldova

nedeljka.zivkovic@bpa.edu.rs

Abstract. A better understanding of the Context and Identity of organisations contributes to their ability to achieve sustainable business success. Elements of ISO 9001:2015 such as the Context and Identity of the organisation encompass relevant interested parties as well as external and internal issues. In addition, the Identity of organisations includes vision, mission, values, and culture. Managers in all organisations should proactively consider the Context and Identity of the organisation, including higher education institutions, especially in the new environment of Industry 5.0. In this regard, the aim of the paper is to identify the Context and Identity variables in higher education institutions within the context of Industry 5.0 using descriptive statistics. The research was conducted in the Republic of Serbia and its region through a questionnaire on the basis of which 374 responses were collected. The results and the conclusions derived from them show the highest mean value for meeting the requirements of customers and other relevant interested parties. This paper can be useful for managers of higher education institutions in gaining a better understanding of the complexity of the Context and Identity of the organisation, particularly in the era of a global social phenomenon such as Industry 5.0.

Keywords: Context and Identity of an organisation, Higher education institutions, Industry 5.0., ISO 9001, ISO 9004

1 Introduction

In the contemporary market environment, organisational systems face increasingly strong competitive pressure. In addition, globalisation imposes new requirements and challenges on organisational systems. One such challenge is the requirement to increase the efficiency and effectiveness of organisational operations. However, for survival in

a highly competitive market, meeting the requirements for increasing efficiency and effectiveness is often not sufficient. This means that organisational systems should invest additional efforts to ensure the sustainability of operations as well as survival in the global competitive market [1, 2]. One approach to achieving the sustainability of operations may be a focus on the application of quality management system standards, which can improve organisational performance. Organisational performance can be defined as the outcomes of operations [3].
Adopting quality management system standards, or focusing on specific elements of these standards, can help organisations achieve sustainability and survival in a market. The application of these elements, which are integrated within the guidelines for achieving sustained success of operations, increases the capability of organisational systems to achieve sustained success of operations. Organisations worldwide invest significant efforts to improve business performance. The implementation of the ISO 9004 standard, or certain parts of this standard, is one of the ways that can enable the mapping of a path for improving organisational performance [4]. According to the available literature, there is a positive relationship between the implementation of quality management systems and organisational performance [5].

All organisations, whether manufacturing, service, or combined, can implement the ISO 9004 standard, which also applies to higher education. In recent decades, quality in higher education has become a priority issue for all management structures of higher education institutions [6]. This is confirmed by numerous studies that have examined quality management systems in higher education. However, measuring the performance of quality management systems in the field of higher education is a very complex task and requires an understanding of the context and complexity of quality management system (QMS) in higher education, as well as consideration of QMS measurement approaches from different perspectives of relevant interested parties [7]. For measuring and evaluating the development of QMS in higher education, understanding the factors that influence the assurance of QMS is of great importance [8]. The achievement of sustained success in the operations of higher education institutions is also influenced by understanding the Context and Identity of the organisation. Context and Identity of the organisation are elements of the widely accepted ISO 9001:2015 standard, and the application of these elements can contribute to improving performance in higher education institutions, especially in the new era of industrial changes characterised by a new industrial revolution. The aim of the paper is to identify the variables of Context and Identity of higher education institutions using descriptive statistics in order to enable the ranking of variables by importance, in doing so, the variables are sorted according to mean values.

1.1 Literature review

All organisations, regardless of their manufacturing or service nature, should implement the ISO 9001:2015 standard in order to ensure a higher level of management of the quality management system [9]. By applying quality management system standards, organisations can improve organisational performance and, in this way, also respond more easily to increasing competition in the global competitive market [10]. Context

and Identity of an organisation are elements of the ISO 9001:2015 standard as part of the international standard for quality management systems [11]. Ultimately, the implementation of quality management system standards directs organisations towards achieving business excellence. It should be noted that there have been different versions of this standard ISO 9001, with the current edition published in 2015 as ISO 9001:2015. The application of this standard in full or in its individual parts provides confidence to all interested parties of the organisation [11]. The part of this standard that relates to a better understanding of the requirements of all interested parties is the Context and Identity of the organisation [11], and this approach is generally applicable to organisations. A significant body of research such as [12, 13, 2], has shown that there is a close relationship between the implementation of the ISO 9001 standard and organisational performance, however, there is a need for deeper research in different applications [5].

Aspects of the organisational Context element include relevant interested parties as well as internal and external issues [11]. Relevant interested parties are individuals or groups that have or may have an influence on the operations of an organisation [14]. Quality is a very significant aspect when it comes to interested parties [15, 16]. Regarding internal and external issues, an organisation must identify those that are relevant to it and ensure a clear understanding of them [17, 18]. External issues relate to legal, societal, and other contextual factors, while internal issues relate to organisational factors [18]. The Identity of an organisation encompasses its values, culture, vision, and mission [18]. While the mission represents the primary purpose of existence, the vision represents the desired future state. Mission and vision statements represent strategic management tools in higher education institutions, which are highly significant for communication as well as for creating the image and identity of higher education institutions [19]. These statements also represent institutional strategies of higher education institutions and are directed towards the activities, values, and organisational culture of the institution [20]. All organisations should have an organisational culture aligned with elements such as values, vision, and mission [18], which also applies to higher education institutions, particularly in the contemporary business conditions of the new industrial revolution.

1.2 Context and Identity of higher education institutions

The Context of an organisation element helps higher education institutions better understand internal and external issues but also a shared understanding of stakeholders' issues [21]. Relevant stakeholders of higher education institutions are, among others, students, professors, parents of students, employers, the relevant ministry, the state, and others [22]. In order to meet the requirements of different stakeholders of higher education institutions, the adoption of the "Context of an organisation" element some of the main reasons may be attracting and retaining students, increasing competition, and similar internal and external factors [23]. As already stated, one of the factors of Context is relevant stakeholders whose expectations are constantly changing. In such conditions, the perceived level of quality of current, former (alumni), and potential students is also influenced by the reputation of the higher education institution [24], and the adoption of these elements by the management structures of higher education

institutions helps in striving for an strengthened and improved reputation [25]. Expected improvements are possible both at the institutional level and at the level of individual processes. In this regard, the primary process in higher education institutions is the teaching process which should be integrated with the mission, vision, value and culture in higher education institutions [26]. The mission and vision statement represents strategic management tools of higher education institutions which are very significant for communication but also creating image and identity of higher education institutions [19]. These statements also represent institutional strategies of higher education institutions and are oriented toward the activities, values, and organisational culture of the institutions [20].

1.3 Industry 5.0 in higher education institutions

While the previous industrial revolution was focused exclusively on technological achievements, and was therefore often subject to criticism [27, 28], Industry 5.0 is focused not only on technological achievements but also on humans and improved collaboration between people and technology [29]. [30] point out that this industrial revolution is predominantly focused on human creativity. Industry 5.0 is a new industrial revolution that has received considerable attention in recent years and represents a global social phenomenon, and it is believed that it can improve organisational performance as well as contribute to increased productivity in organisations [31]. Industry 5.0 implies the use of the latest technological achievements, primarily artificial intelligence, Big Data and Analytics, Internet of Things, Autonomous robot, Machine learning and similar [32], while the primary goal is the satisfaction of human needs, resilience and sustainability of business operations [33]. As in all fields, including higher education, Industry 5.0 holds a significant place, especially since there are indications that the human role could be partially replaced by automated systems [28]. There is a lack of research on the application of Industry 5.0 in organisations in general, and particularly in higher education institutions, which represent the fundamental pillar of the development of every society [28].

2 Methodology

The study involved employees at higher education institutions in the Republic of Serbia and the surrounding region. Furthermore, a questionnaire was used in this study as the primary instrument for data collection. The distribution of the questionnaire was carried out via email to employees at higher education institutions. The research variables were adopted directly from the ISO 9004:2018 standard. The study included both public and private higher education institutions. The questionnaire consisted of two parts: the first part was dedicated to demographic characteristics, while the second part was focused on the professional section. A total of 374 responses were collected using the questionnaire. Variables related to the ISO 9004:2018 elements of Context and Identity in higher education institutions were analysed using descriptive statistics to enable their ranking by importance, based on mean values. For the descriptive statistics analysis, SPSS v.27.0 was used.

3 Results

The first part of the research covers the demographic characteristics of the respondents. Analysing the demographic structure, 52.94% of respondents were female, while 47.06% were male. According to work experience structure, 1.87% of respondents have less than 1 year of work experience, 27.54% have 1–10 years of experience, 38.24% have 11–20 years of experience, while 32.35% have more than 20 years of work experience. Furthermore, analysing the structure according to job position in higher education institutions, 16.84% of respondents belong to administrative staff, the same percentage are teaching assistants, while 66.32% are professors. Based on institutional ownership, 90.64% of respondents are from public higher education institutions, while only 9.36% are from privately founded higher education institutions.

The second part of the research includes descriptive statistics of the analysed variables related to the ISO 9004:2018 elements of organisational Context and Identity. The questions covered by the Context of the organisation were: meeting the requirements of customers and other relevant interested parties in relation to the Industry 5.0 environment, and addressing internal and external issues in the new Industry 5.0 environment, while the question related to the Identity of the organisation referred to the way in which vision, mission, culture, and values are established in the environment of the new industrial revolution. Table 1 presents the analysis of descriptive statistics of the observed variables. These elements represent strategic tools of higher education institutions [19].

Table 1. Analysis descriptive statistics

VARIABLES	*Code*	*Mean*	*Std. Deviation*	*Mediana*	*Mode*	*Range*
To what extent are the requirements of customers and other relevant interested parties being met in higher education institutions?	K&I_1	3.51	1,022	4.00	4	4
To what extent are internal and external issues being addressed?	K&I_2	3.48	1,058	4.00	4	4
How are vision, mission, culture and values established?	K&I_3	3.39	1,070	3.00	3	4

4 Discussion

Analysing the obtained research results, the highest value (3.51) was recorded for the variable related to meeting the requirements of customers and other interested parties,

which is in line with [15, 16]. Meeting the requirements of customers and other relevant interested parties is highly significant for all higher education institutions. For achieving sustainability of HEIs' operations, it is highly important to create new value for all relevant stakeholders of HEIs [26]. Also, in the field of higher education, authors agree that HEIs should establish an appropriate performance measurement system for potential improvements, including all relevant stakeholders [34]. Based on performance measurement, highly significant information is obtained that can be used for potential improvements. Furthermore, the next variable refers to the consideration of internal and external issues that are important for strategic orientation [17], as well as for the functioning and implementation of teaching and non-teaching processes in higher education institutions. When considering internal and external issues, top management should take into account key risks that could affect the sustainability of HEIs' operations, as well as opportunities [18]. The obtained mean value is 3.48. In addition, a slightly lower value was obtained for understanding the way in which mission, vision, values, and culture are established. This result is consistent with the findings of [35], who point out that higher education institutions have still not fully aligned their strategies with the new requirements of Industry 5.0. By actively monitoring and adopting Industry 5.0 trends and requirements, as well as consistently responding to these new demands through the redefinition of HEIs' strategies and visions aimed at meeting the new requirements of relevant stakeholders, primarily students, a solid foundation can be established for overcoming the challenges brought by this new industrial revolution [35, 28]. The new industrial revolution involves the redesign of strategic values such as mission, vision, values, and culture in higher education institutions, in order to meet the new expectations of students and other relevant stakeholders of higher education institutions [32].

5 Conclusions

The aim of the paper was to identify variables related to Context and Identity in higher education institutions by applying descriptive statistics to enable their ranking by importance, based on mean values. Descriptive statistics help researchers gain a better understanding of a dataset [36, 37]. Understanding the Context and Identity of higher education institutions is highly significant for their management. In this regard, understanding and considering priorities related primarily to meeting the requirements of students and other interested parties is of key importance for the functioning of higher education institutions [26], in relation to the Industry 5.0. According to [17], the consideration of internal and external issues represents the strategic orientation of organisations. It is expected that this social change called Industry 5.0 will encourage higher education institutions to consider how to utilize the modern technologies in order to meet new expectations of students (such as sustainability-related skills) as well as other relevant stakeholders brought by Industry 5.0 [28].

This research also has several limitations. One limitation is that only three questions from the original structure of the ISO 9004:2018 standard were included in the study. The next limitation relates to the application of descriptive statistics itself, which

enables a better understanding of the characteristics of the sample, however, further application of additional statistical techniques is recommended. For example, the application of regression analysis could provide additional information and insights regarding the analysed variables, which may guide future research. Furthermore, future research may include other elements of the ISO 9004:2018 standard.

Acknowledgments. This research is based upon work from COST Action CA22124 - ECO4ALL, "EU Circular Economy Network for All: Consumer Protection through reducing, reusing, repairing," supported by COST (European Cooperation in Science and Technology—www.cost.eu), and analysed the Ministry of Science, Technological Development and Innovations of the Republic of Serbia, Contract No. 451-03-137/2025-03/ 200131.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Glogovac, M., Ruso, J., Arsić, S., Rakić, A., & Milošević, I. Leadership for quality 4.0 improvement, learning, and innovation. Engineering Management Journal, 35 (3), 313–329 (2023).
2. Siltori, P. F., Simon Rampasso, I., Martins, V. W., Anholon, R., Silva, D., & Souza Pinto, J. Analysis of ISO 9001 certification benefits in Brazilian companies. Total Quality Management & Business Excellence, 32 (13–14), 1614–1632 (2021).
3. Vetchagool, W., Augustyn, M. M., & Tayles, M. ISO 9000, activity based costing and organizational performance. Total Quality Management & Business Excellence, 32 (3–4), 265–288 (2021).
4. Bravi, L., & Murmura, F. Evidences about ISO 9001: 2015 and ISO 9004: 2018 implementation in different-size organisations. Total Quality Management & Business Excellence, 33 (11–12), 1366–1386 (2022).
5. Sfreddo, L. S., Vieira, G. B. B., Vidor, G., & Santos, C. H. S. ISO 9001 based quality management systems and organisational performance: a systematic literature review. Total Quality Management & Business Excellence, 32 (3–4), 389–409 (2021).
6. Bravo, L. G., Nistor, N., Ramírez, B. C., Soto, I. G., Contreras, M. V., Vives, M. N., & Robles, P. M. Higher education managers' perspectives on quality management and technology acceptance: A tale of elders, mediators, and working bees in times of Covid-19. Computers in Human Behavior, 131, 107236 (2022).
7. Matosas-López, L., Aguado-Franco, J. C., & Gómez-Galán, J. Constructing an instrument with behavioral scales to assess teaching quality in blended learning modalities. Journal of New Approaches in Educational Research, 8 (2), 142–165 (2019).
8. Chaiya, C., & Ahmad, M. M. Success or failure of the Thai higher education development—Critical factors in the policy process of quality assurance. Sustainability, 13 (17), 9486 (2021).
9. Makolov, V. Context of organization and quality management. In IOP Conference Series: Earth and Environmental Science (Vol. 272, No. 3, p. 032216). IOP Publishing (2019, June).

10. Sari, Y., Wibisono, E., Wahyudi, R. D., & Lio, Y. From ISO 9001: 2008 to ISO 9001: 2015: Significant changes and their impacts to aspiring organizations. In IOP Conference Series: Materials Science and Engineering (Vol. 273, No. 1, p. 012021). IOP Publishing. (2017, November).
11. International Organization for Standardization. 9001: 2015. Quality management systems–Requirements. ISO. (2015).
12. Wei, J., & Wang, X. Impact of quality management system on operational efficiency in agro-processing SMEs. Total Quality Management & Business Excellence, 1–39 (2025).
13. Camango, C., & Cândido, C. J. ISO 9001 maintenance, decertification and recertification: a systematic literature review. Total Quality Management & Business Excellence, 34 (13–14), 1764–1796 (2023).
14. Martin, J., Elg, M., & Gremyr, I. The many meanings of quality: towards a definition in support of sustainable operations. Total Quality Management & Business Excellence, 36(3-4), 185-198 (2025).
15. Bettinson, E., Young, K., Haven-Tang, C., Cavanagh, J., Fisher, R., & Francis, M. Employers' conceptions of quality and value in higher education. Higher Education, 87 (5), 1393–1409 (2024).
16. Clemons, R., & Jance, M. Defining Quality in Higher Education and Identifying Opportunities for Improvement. Sage Open, 14 (3) 1-11 (2024).
17. Wolniak, R. Context of the organization in ISO 9001: 2015. Zeszyty Naukowe. Organizacja i Zarządzanie/Politechnika Śląska. 133, 121-136 (2019).
18. International Organization for Standardization. 9004:2018. Quality management-Quality of an organization-Guidance to achieve sustained success. ISO. (2018).
19. Tamássy, R., Géring, Z., Harmat, V., Knúlné Pál, E., & Rakovics, M. Navigating diverse demands: mission and vision statements of top-ranked European higher education institutions. Studies in Higher Education, 51(6), 1354-1374 (2026).
20. Bedoya-Dorado, C., Murillo-Vargas, G., & Gonzalez-Campo, C. H. Sustainability in the mission and vision statements of Colombian Universities. International Journal of Sustainability in Higher Education, 23(1), 67-86 (2022).
21. Jingura, R. M., Kamusoko, R., & Tapera, J. Critical analysis of the applicability of the ISO 9001 Standard in higher education institutions. International Journal of African Higher Education, 6(1), 97–120. (2019).
22. Živković, N. Model zrelosti sistema menadžmenta kvaliteta visokoškolskih ustanova (Doctoral dissertation, Univerzitet u Beogradu-Fakultet organizacionih nauka). (2025).
23. Ramese, W. K. D., Jayasooriya, S. D., & Lakmal, A. H. Adopting ISO 9001: 2015 & ISO 21001: 2018 EOMS to the Sri Lankan Higher Educational Institutions: A Review of Literature. KDU Journal of Multidisciplinary Studies, 6(1) 145-157 (2024).
24. Masserini, L., Bini, M., & Pratesi, M. Do quality of services and institutional image impact students' satisfaction and loyalty in higher education?. Social Indicators Research, 146 (1), 91–115 (2019).
25. Almuhaideb, A. M., & Saeed, S. Fostering sustainable quality assurance practices in outcome-based education: Lessons learned from ABET accreditation process of computing programs. Sustainability, 12 (20), 8380 (2020).
26. Serdar Asan, Ş., & Tanyaş, M. Integrating Hoshin Kanri and the balanced scorecard for strategic management: The case of higher education. Total Quality Management, 18 (9), 999–1014 (2007).

27. Ruso, J., Milošević, I., Stojanović, A., Glogovac, M., Rakić, A., & Arsić, S. Regional variations in digital transformation: the impact of industry 5.0 on business quality performance. Total Quality Management & Business Excellence, 35 (15–16), 1879–1899. (2024).
28. Hashim, M. A. M., Tlemsani, I., Mason-Jones, R., Matthews, R., & Ndrecaj, V. Higher education via the lens of industry 5.0: Strategy and perspective. Social Sciences & Humanities Open, 9, 100828 1-16 (2024).
29. Waheed, A., Alharthi, M., Khan, S. Z., & Usman, M. Role of industry 5.0 in leveraging the business performance: investigating impact of shared-economy on firms' performance with intervening role of i5. 0 technologies. Sage Open, 12 (2) (2022).
30. Pelaez-Sanchez, I. C., Glasserman-Morales, L. D., & Rocha-Feregrino, G. Exploring digital competencies in higher education: Design and validation of instruments for the era of Industry 5.0. In Frontiers in Education (Vol. 9, p. 1415800). Frontiers Media SA. (2024).
31. Tlili, A., Huang, R., & Kinshuk, X. Metaverse for climbing the ladder toward 'Industry 5.0'and 'Society 5.0'?. The Service Industries Journal, 43 (3–4), 260–287 (2023).
32. Rosak-Szyrocka, J. Engineering The Future of Higher Education: A VOSviewer Analysis of Smart University Trends in The Digitalization and Industry 5.0 Era. Management Systems in Production Engineering. 33(1), 29–39. (2025).
33. Supriya, Y., Bhulakshmi, D., Bhattacharya, S., Gadekallu, T. R., Vyas, P., Kaluri, R., ... & Mahmud, M. Industry 5.0 in smart education: Concepts, applications, challenges, opportunities, and future directions. IEEE Access, 12, 81938-81967 (2024).
34. Makki, A. A., Alqahtani, A. Y., Abdulaal, R. M., & Madbouly, A. I. A novel strategic approach to evaluating higher education quality standards in University Colleges using multi-criteria decision-making. Education Sciences, 13 (6), (2023).
35. Arregi Lopez, A., Igartua, J. I., Retegi, J., Eguren, J. A., & Ibarra, D. Addressing the challenges associated with Industry 5.0 in higher education: a training concept approach. The TQM Journal, 38(4), 713-731. (2026).
36. Fisher, M. J., & Marshall, A. P. Understanding descriptive statistics. Australian Critical Care, 22(2), 93-97 (2009).
37. Thompson, C. B. Descriptive data analysis. Air medical journal, 28(2), 56-59 (2009).

Original Research Article *https://doi.org/10.5281/zenodo.23477202*

Integrating Artificial Intelligence into Quality Management: Opportunities, Challenges, and Future Directions

Mladen Đurić[1][0000-0002-8219-2899]

[1] University of Belgrade, Faculty of Organizational Sciences, Serbia

Abstract. The paper is dedicated to the analysis of the relationship that occurs between artificial intelligence (AI) and quality management (QM) in organisations, based on relevant studies. The main areas in which AI is proved to be transforming QM are: automated defect detection, visual inspection, rework prediction, root cause identification, and customer satisfaction analysis. Despite great improvement in efficiency and accuracy that AI brings, several challenges follow, mainly related to transparency, data quality, and ethical standardisation. An integrated four-level framework for the synergy of AI and QM is proposed.

Keywords: Artificial Intelligence, Quality Management, Organisations.

1 Introduction

In 2025, 88% of businesses surveyed in a McKinsey research were using artificial intelligence (AI), and this novelty became the new normal (Singla, et al. 2025). At the same time, quality management (QM) has been changing under the influence of technological improvements that Industry 4.0 has brought. In fact, the traditional QM approaches were proved to be insufficient to deal with growing demands of contemporary markets, in terms of both supply- and demand-side factors. The advancements in technology have enabled managers to control quality in new ways – through the use of AI, and this topic has gained a lot of attention recently. Therefore, the following paper aims to (1) systematise knowledge on AI applications in QM, (2) identify key application areas, (3) analyse relevant challenges, and (4) propose future research directions.

2 Applications of Artificial Intelligence in Quality Management

The Fourth Industrial Revolution has resulted in significant benefits for the domain of quality control and management, in particular through the use of big data. Batarseh et al. (2021) identified five subareas of AI relevant to quality assurance (QA): machine learning (ML), deep learning, reinforcement learning, computer vision, and natural

language processing (NLP). They connect it to AI and explain AI assurance as a "process that is applied at all stages of the AI engineering lifecycle ensuring that any intelligent system is producing outcomes that are valid, verified, data-driven, trustworthy and explainable to a layman, ethical in the context of its deployment, unbiased in its learning, and fair to its users" (Batarseh, Freeman and Huang 2021, 2).

Recent studies reveal various ways in which AI is entering QM domain. Most studies mention: automated defect detection, visual inspection, rework prediction, root cause identification, and customer satisfaction analysis. The most visible implementation is detecting defects and determining the quality of a product. Sariyer et al. (2025) have developed a model to classify products depending on their quality levels, with the use of multilayer perceptron (MLP), and achieved 96% accuracy – significantly more useful than a binary (defect/non-defect) classification. Similarly, Kausik et al. (2025) have shown that supervised methods (SVMs, CNNs, Random Forests) achieve high accuracy in classifying defects, and that unsupervised approaches (K-Means, DBSCAN) are useful in anomaly detection, which was a conclusion they got through the analysis of more than 300 studies. Regarding quality control, image – that is, visual inspection is particularly important for the use of AI in QM. A study from 2024 showed that 80% of large enterprises use AI-enhanced systems, which is one of the largest technologies chosen for the sake of improving efficiency and reliability (Jain 2024). Chouchene et al. (2020) paid significant attention to the question of automated visual systems for vehicle non-conformity control. They highlighted that AI can significantly decrease costs of quality control (visual control of different components can be done in 3-20 seconds) and improve results – accuracy can be significantly higher than in the case of human-led control (Chouchene, et al. 2020). In terms of rework predictions and root cause identification, Sariyer et al. (2025) gave additional contribution – with MLP, they also predicted rework quantities with 98% performance and applied Association Rule Mining (Apriori) to identify root causes, with 80% confidence. For risk management, AI can be implemented through (neuro-)fuzzy models to identify risks and mitigate the process-related ones (Malik and Singh 2020). Ultimately, AI can be – and is often – used for customer satisfaction analysis. Even ten years ago, Yussupova et al. (2016) outlined an AI-based customer satisfaction research, using sentiment analysis of more than 635000 hotel reviews, with almost 94% accuracy. In 2025, Alkhaldi suggested that using AI to analyse customer feedback can greatly help a business implement a customer-centric strategy, further improving the proactiveness of a company.

3 Challenges and Future Research Directions

In 2013, Allahbakhsh et al. predicted that AI will be deeply integrated in quality control mechanisms, but that it will also raise certain issues. They believed those would be mainly related to the challenge of cooperation between humans and AI (Allahbakhsh, et al. 2013). In a systematic literature review from 2024, Cevallos et al. identified five main challenges of AI adoption in QM of small and medium enterprises, which are tied to ignorance or insufficient skills in the AI domain, lack of

resources to build the infrastructure, data security and data privacy-related concerns, cultural barriers (team rejecting the change), and technical difficulties to integrate AI in the existing system (Cevallos, et al. 2024). Surely, ethical concerns are to be added on this list, as Tosic et al. (2023) warned about the proliferation of the standards, since this potentially creates more confusion. Almost a decade ago, Bryson and Winfield (2017) suggested that standardisation bodies should address the issues of ethical dimensions, so that QA can have a social side, not only be tied to technical performance. While there are numerous initiatives, there is no one universally accepted solution. ISO/IEC 42001:2023 contains structured approaches to manage AI systems, with focus on risk mitigation, governance, and accountability. In more recent studies, there is a particular emphasis on operationalising fairness, transparency, and lifecycle-based assurance, aligned with the EU AI Act (de Borba and Brinkhues 2025, McCormack and Bendechache 2026).

Ultimately, most organisations actually lack dedicated AI governance structures, which would be in charge of overseeing the implementation of various AI solutions. In other words, even when some of challenges are resolved, the company does not have a clear mechanism of accountability, and this is the core of the work McCormack and Bendechache, who in 2026 present their Trustworthy AI Maturity Model (TAIMM), with the idea of aiding the creation of ethical and legally compliant AI systems within organisations (McCormack and Bendechache 2026). Beyond the challenges identified in the literature, additional concerns deserve attention: most notably, the temporal dimension. This is a relevant question, as the environment changes continuously – from both demand (customers expectations, trends), and supply side (materials, production conditions, etc). As a consequence, there is a great need for an AI-based QM approach to be agile-first to keep its benefits.

As a logical suggestion to the outlined challenges would be the call for future research to focus on the quality standards, as they would need to be very dynamic, and regulatory bodies should be able to at least keep up the pace with the industry. Furthermore, the idea of relying on the same AI solutions for quality control across a supply chain would help avoid the problem of too many different AI solutions between entities. Ultimately, the human-AI collaboration must not be neglected, as it resolves many challenges and opens the door to many more applications, still unimaginable.

4 Conclusion

The relationship between AI and QM in organisations is gaining attention. The key application areas, as identified through the research of literature are automated defect detection, visual inspection, rework prediction, root cause identification, and customer satisfaction analysis, and all speak about AI's transformative potential in QM. At the same time, significant challenges persist, including ethical standardisation complexity, lack of governance structures, and the interaction between human and AI. AI has a dual role in QM: it is a tool for enhancing quality processes, but also a subject of quality assurance itself. Organisations seeking to integrate AI into QM

should therefore address not only the technical implementation, but also governance, ethics, and workforce readiness. Future research should focus on dynamic quality standards that evolve alongside AI capabilities, cross-organisational AI-QM models within supply chains, and human-AI collaborative frameworks that balance automation with domain expertise.

References

1. Alkhaldi, F.M.: AI-driven customer feedback analytics: Unlocking latent insights for strategic business agility. Al-Zaytoonah Journal of Business **1**, 1–15 (2025)
2. Allahbakhsh, M., Benatallah, B., Ignjatovic, A., Motahari-Nezhad, H.R., Bertino, E., Dustdar, S.: Quality control in crowdsourcing systems: Issues and directions. IEEE Internet Computing **17**(2), 76–81 (2013)
3. Batarseh, F.A., Freeman, L., Huang, C.H.: A survey on artificial intelligence assurance. Journal of Big Data **8**(1), 1–30 (2021)
4. Bryson, J., Winfield, A.: Standardizing ethical design for artificial intelligence and autonomous systems. Computer **50**(5), 116–119 (2017)
5. Cevallos, M.G.O., Jaramillo, G.A.L., Vélez, J.L.B., Zambrano, M.I.U., Montesdeoca, L.D.Z., Palomeque, M.A.B.: Implementation of artificial intelligence in quality management in SMEs: Benefits and challenges. Evolutionary Studies in Imaginative Culture Impact **8**(2), 1489–1500 (2024)
6. Chouchene, A., Carvalho, A., Lima, T.M., Charrua-Santos, F., Osório, G.J., Barhoumi, W.: Artificial intelligence for product quality inspection toward smart industries: Quality control of vehicle non-conformities. In: 9th International Conference on Industrial Technology and Management (ICITM), pp. 127–131. IEEE, Oxford (2020)
7. de Borba, D., Brinkhues, R.: Artificial intelligence in business process management: Challenges, opportunities and strategies for alignment with ISO 42001. Review of Artificial Intelligence in Education **6**, 1–21 (2025)
8. Jain, D.: Artificial intelligence in quality control systems: A cross-industry analysis of applications, benefits, and implementation frameworks. International Journal of Scientific Research in Computer Science, Engineering and Information Technology **10**(6), 1321–1333 (2024)
9. Kausik, A.K., Rashid, A.B., Baki, R.F., Maktum, M.M.J.: Machine learning algorithms for manufacturing quality assurance: A systematic review of performance metrics and applications. Array **26**, 1–30 (2025)
10. Malik, V., Singh, S.: Artificial intelligent environments: Risk management and quality assurance implementation. Journal of Discrete Mathematical Sciences and Cryptography **23**(1), 187–195 (2020)
11. McCormack, L., Bendechache, M.: The Trustworthy AI Maturity Model (TAIMM): Integrating ethics and regulation across the AI lifecycle. Journal of Responsible Technology **21**, 100156 (2026)
12. Sariyer, G., Mangla, S.K., Kazancoglu, Y., Ocal Tasar, C., Luthra, S.: Data analytics for quality management in Industry 4.0 from a MSME perspective. Annals of Operations Research **350**(2), 365–393 (2025)
13. Singla, A., Sukhovetsky, A., Yee, L., Chui, M.: The state of AI in 2025: Agents, innovation, and transformation. McKinsey & Company, https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai, last accessed 2025/11/05

14. Tosic, B., Mijatovic, I., Kicanovic, A.: Ethical aspects of AI and the role of standardisation in IT: Are we facing the Brussels effect? In: XIX International Symposium Unlocking the Hidden Potentials of Organization Through Merging of Humans and Digitals, pp. 871–875. University of Belgrade – Faculty of Organizational Sciences, Zlatibor (2023)
15. Yussupova, N., Kovács, G., Boyko, M., Bogdanova, D.: Models and methods for quality management based on artificial intelligence applications. Acta Polytechnica Hungarica **13**(3), 45–60 (2016)

Original Research Article *https://doi.org/10.5281/zenodo.23477215*

Implementation of the DMAIC Methodology for Quality Improvement in Aviation Industry: Case Study of "Belgrade Airport"

Mila Marković[1][0009-0005-7745-2341] Mladen Đurić[1][0000-0002-8219-2899] and Jovana Dimitrijević[2][0009-0009-3812-941X]

[1] Faculty of Organizational Sciences, University of Belgrade
[2] Independent Researcher, Belgrade, Serbia

Abstract. The aim of this paper was to demonstrate how the DMAIC framework can be applied across different industries and to show that, although it originated in the manufacturing sector, it is no longer associated exclusively with production processes. Given that DMAIC is increasingly used today to improve service and infrastructure-related processes, this paper examines its application through the example of the company "Belgrade Airport."

The paper is divided into two parts: theoretical and practical. The first part presents the origins of the DMAIC methodology and explains the main activities within each phase of the DMAIC framework (Define, Measure, Analyze, Improve, Control), as well as the tools most used in process analysis and improvement. Emphasis is placed on the ESG standards framework, together with a brief overview of what these standards entail.

The practical part of the paper is based on a case study concerning the management of heating, ventilation, and air conditioning in the terminal building at Belgrade Nikola Tesla Airport. The problem of non-optimal energy consumption during periods of low space occupancy was identified, and, through the application of the DMAIC methodology, the root causes of the problem were analyzed and appropriate improvements were proposed. Special emphasis was placed on the implementation of automated HVAC system control in accordance with actual space occupancy, which would contribute to achieving compliance with ESG requirements.

Keywords: DMAIC, aviation industry, HVAC, quality improvement, Belgrade Airport.

1 Introduction

The paper may be viewed as comprising two separate yet interconnected parts: a theoretical review of the DMAIC framework and its Lean Six Sigma foundations, and a practical application through a case study of "Belgrade Airport."

In today's competitive business environment, organisations face continuous pressure to improve processes and reduce costs while maintaining service quality. Structured methodologies such as Lean Six Sigma provide a systematic framework for addressing

these challenges, with DMAIC representing one of its most important process improvement tools applicable not only in manufacturing but across a growing range of operational contexts [15].

Energy efficiency and sustainability are of particular importance in large infrastructure systems such as airports, where HVAC systems account for a significant share of total energy consumption. Non-optimal management of these systems leads to increased operating costs, negative environmental impacts, and difficulty in meeting ESG objectives, a challenge central to the operations of airports owned by the company "VINCI."

Against this background, this paper addresses a research gap in the application of structured process improvement methodologies to airport infrastructure management. While DMAIC has been widely employed in manufacturing and general service settings, its application to HVAC energy management in an airport terminal environment, with ESG compliance as an explicit evaluative criterion, has not been systematically addressed in the literature. The central research question guiding this study is therefore: can the DMAIC methodology be effectively applied to reduce non-optimal HVAC energy consumption in an airport terminal, and what insights does this application yield for ESG-oriented operational improvement? The objective of the paper is to answer this question through a structured case study of Belgrade Nikola Tesla Airport. The contribution is twofold: it extends the existing body of DMAIC applications to a new and practically relevant domain, and it demonstrates that ESG requirements can be meaningfully integrated into the DMAIC process, not merely as background context, but as a guiding framework for defining quality requirements, evaluating improvement alternatives, and monitoring outcomes across all five phases.

2 Brief Introduction to DMAIC Methodology

The DMAIC methodology represents a continuous improvement cycle consisting of five phases, Define, Measure, Analyze, Improve, and Control, applicable to both process and product quality improvement [7]. As a core component of the Lean Six Sigma approach, it enables organisations to systematically diagnose and resolve quality issues by targeting their root causes [3, 10]. DMAIC is particularly suited to processes that function adequately but fall short of optimal performance, where suboptimal operation increases the likelihood of errors and inefficiencies, making structured intervention necessary [5, 14].

Table 1. Phases of DMAIC Methodology [7]

Phases	Define	Measure	Analyze	Improve	Control
Activities	Defining the current situation	Identifying potential causes of the problem	Analyzing potential causes of the problem	Identifying potential solutions to the problem	Verifying the effectiveness of the solution

	Defining the desired situation after improvement	Quantifying the problem	Selecting the primary causes of the problem	Selecting and implementing the solution	Implementing control measures
Results	Defined problem and potential benefits	Root cause of the problem identified and potential causes listed	Quantification of potential causes and selection of the primary cause of the problem	Pilot solution and final implementation of the problem solution	Process control to sustain the achieved benefits

The Define phase establishes the project foundation by clearly identifying the problem, setting objectives, and determining key customer requirements [18]. Its output is a defined improvement goal, developed through tools such as process mapping, stakeholder analysis, and the Lean Six Sigma project charter [2, 14, 17].

The Measure phase ensures that subsequent analysis and decisions are grounded in accurate, systematically collected data rather than assumptions [12, 18]. It addresses key measurement concepts including data types, variability principles, and repeatability and reproducibility analyses [2, 14].

In the Analyze phase, data on the key process characteristics are examined to identify the root causes of the problem, drawing on both Six Sigma's focus on process variation and Lean's structured approach to waste elimination — encompassing seven waste categories: overproduction, waiting, transportation, process inefficiencies, inventory, motion, and defects [13, 16, 17].

The Improve phase generates and implements solutions aimed at eliminating the identified root causes, relying on team-based techniques such as brainstorming to support collaborative problem-solving [14, 17].

In the Control phase, the changes made are formalised: improved processes are documented and standardised, results are verified against project objectives, and monitoring mechanisms are established to sustain the achieved improvements [17].

The structured, phase-based logic of DMAIC makes it particularly well suited to complex operational environments where problems are multi-causal and improvement goals span both technical and organizational dimensions. This flexibility underpins its application in the case study presented in the following section.

3 Application of the DMAIC Methodology in the Company “Belgrade Airport”

The aviation industry forms the backbone of the modern global economy, enabling international trade, tourism, and global connectivity, while simultaneously facing significant operational pressures related to fuel price fluctuations, geopolitical instability, and regulatory changes [8]. As one of the fastest-growing sectors for over a century, aviation has also generated substantial environmental consequences, including extensive

fossil fuel consumption, waste management challenges, and a growing negative impact on air quality and climate [9].

These pressures have placed ESG performance at the centre of strategic priorities for airport operators worldwide. Sustainability today requires proactive efforts across economic, environmental, and social dimensions, with the ESG framework providing a structured basis for evaluating an organisation's long-term responsibility and overall business value [11]. The aviation industry is consequently facing increasing scrutiny related to environmental, social, and governance risks [6].

Against this background, the following case study examines the application of the DMAIC methodology at Belgrade Nikola Tesla Airport. An internal analysis of energy management data recorded via the Building Management System and electricity metering infrastructure at Belgrade Nikola Tesla Airport during 2025 showed that, on average, 18% of total electricity consumption resulted from the non-optimal use of heating, ventilation, and air conditioning systems during periods of low occupancy.

In this case study, the DMAIC framework is applied to the management of heating, ventilation, and air conditioning in the terminal buildings of “Belgrade Airport.” The identified problem refers to the non-optimal operation of the HVAC system during periods of low occupancy, such as night-time hours and off-season periods, resulting in unnecessary energy consumption. Since this directly affects operating costs, CO_2 emissions, and progress towards ESG objectives, the improvement effort focuses on aligning HVAC operation with actual space occupancy in order to achieve more efficient and sustainable terminal management.

3.1 Application of the “Define” Step in the Company “Belgrade Airport”

An organisation’s structure provides the framework through which work is divided, responsibilities are assigned, and coordination between different functions, employees, and organisational units is achieved. In this way, it supports the efficient functioning of the organisation and the achievement of shared objectives [1].

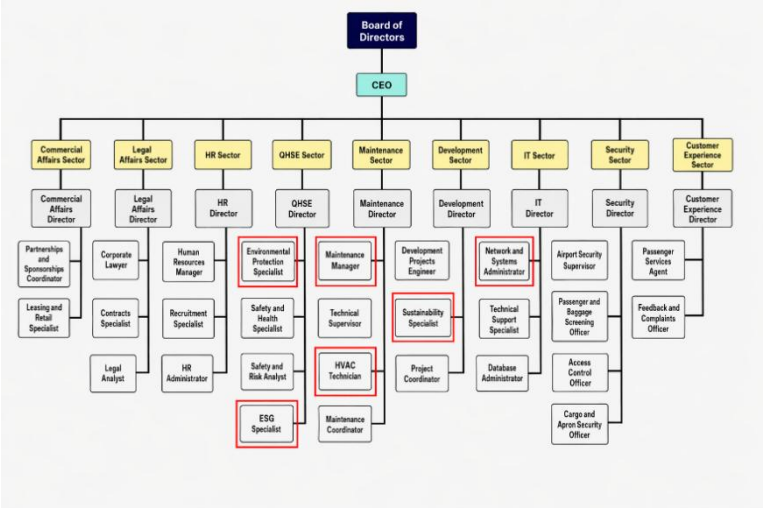


Fig. 1. Organizational structure and DMAIC team composition at "Belgrade Airport".

Figure 1 shows the organisational structure of "Belgrade Airport", where operations are divided into nine sectors, each managed by a sector director and coordinated under the Chief Executive Officer. The Board of Directors has the highest level of authority over the company's activities. Since this study focuses on the management of heating, ventilation, and air conditioning, the Maintenance Sector is identified as the organisational unit most relevant to the analysed process.

Lean Six Sigma Project Charter. A well-structured Project Charter establishes the project scope, defines the available resources, and sets clear boundaries all of which are prerequisites for a successful DMAIC implementation [4].

The Lean Six Sigma Project Charter outlines a DMAIC project aimed at optimizing HVAC system operations at "Belgrade Airport" in order to reduce energy consumption, lower CO2 emissions, and decrease operational costs in line with ESG objectives. The project focuses on addressing non-optimal system performance during periods of low occupancy and involves a multidisciplinary team within the Maintenance Sector. The project timeline is defined, starting on January 16, 2026, and ending on March 16, 2026, with key milestones including the completion of the measurement phase (January 28), analysis phase (February 1), improvement phase (February 7), and final control phase (March 16). Overall, the project is expected to enhance operational efficiency and deliver the greatest benefits to "Belgrade Airport".

CTQ (Critical-to-Quality) Tree. The HVAC system (Heating, Ventilation, and Air Conditioning system) refers to all systems that regulate temperature, airflow, and air quality within buildings. The customer requirements and CTQ parameters were defined based on the expected needs of passengers and employees in the terminal, drawing on

inputs from the Airport Experience function, passenger feedback mechanisms, and service quality monitoring indicators, with particular focus on thermal comfort, indoor air quality, and energy-efficient HVAC operation.

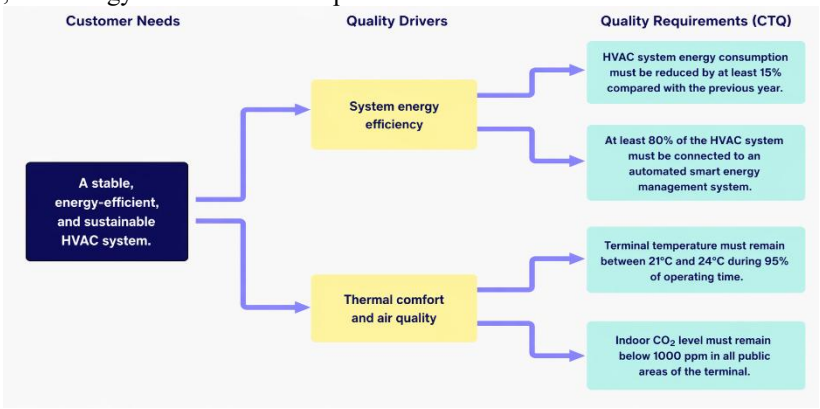


Fig. 2. CTQ (Critical-to-Quality) tree for the HVAC system.

Fig. 2 presents a CTQ (Critical-to-Quality) tree for the HVAC system. It identifies the main customer need, a stable, energy-efficient, and sustainable HVAC system, and links it to the key quality drivers, namely energy efficiency and thermal comfort with air quality. These drivers are further translated into measurable quality requirements related to energy consumption, automation, temperature control, and indoor CO2 levels.

Analysis of Customer Requirements. The following analysis identifies the main HVAC-related requirements from the perspective of both passengers and airport operations staff, assigning measurable targets and priority rankings to each.

Customer: In this case, the customer is the client, i.e. the company "Belgrade Airport"		**Output:** Optimal operation of the HVAC system, leading to cost reduction and better alignment with ESG principles.			
Customer Requirements	**Measure**	**Target**	**Importance to the Customer**	**Customer Satisfaction**	**Priority Ranking (Importance × Satisfaction)**
1. Reduction of HVAC system energy consumption in the terminal building.	Total annual electricity consumption of the HVAC system (kWh).	Reduce energy consumption by at least 15% compared to the previous year.	9	3	27
2. Improvement of energy efficiency through automated and centralized HVAC system management.	Percentage of the HVAC system connected to the BMS.	At least 80% of the HVAC system connected to the BMS.	3	3	9
3. Ensuring pleasant and healthy conditions for passengers and employees in the terminal.	Temperature (°C) and CO_2 concentration (ppm) in the public areas of the terminal.	Temperature between 21–24°C and CO_2 level below 1000 ppm.	9	9	81
	1 – low	**3 – medium**	**9 – high**		
What is the customer's main complaint? What would they like us to do for them?					
Legend Importance to the Customer: 1 = Not important 2 = Important 3 = Very important Customer Satisfaction: 1 = Very satisfied 2 = OK. Could be better 3 = Not satisfied, improvement is required					

Fig. 3. Analysis of customer requirements for HVAC system management at "Belgrade Airport".

Fig. 3 presents an analysis of customer requirements related to HVAC system management at “Belgrade Airport”. It identifies the main customer requirements, defines the corresponding measures and target values, and evaluates them according to customer importance and satisfaction. Based on this assessment, the figure highlights the priority areas for improvement, with the greatest emphasis placed on providing comfortable and healthy conditions in the terminal. The third requirement, namely 'Ensuring pleasant and healthy conditions for passengers and employees in the terminal,' has the highest overall ranking, which reinforces the dual imperative of this project: operational energy reduction must not compromise passenger comfort, a balance that directly reflects the social dimension of the company's ESG commitments and informs the quality targets defined in the subsequent Measure phase.

3.2 Application of the “Measure” Step in the Company “Belgrade Airport”

The Measure phase builds directly on the quality requirements established in the Define phase, translating each CTQ requirement into a concrete, measurable indicator with a defined target value, measurement method, and monitoring frequency. The measurement targets were operationalised on the basis of existing airport service standards and BMS operational data, with baseline values for each indicator derived from electricity metering records for the preceding operational period, against which post-improvement measurements will be compared. Table 2 presents the resulting monitoring and measurement plan for each defined quality requirement.

Table 2. Monitoring and Measurement Plans for Each Quality Requirement

No.	Quality Requirement	What is Measured?	Measurement Method	Measurement Frequency
1.	HVAC system energy consumption must **be reduced by at least 15%** compared with the previous year.	Total electricity consumption of the HVAC system.	Reading data from electricity meters and the BMS system.	Monthly and/or annually.
2.	**At least 80% of the HVAC system must be connected** to an automated smart energy management system (BMS).	Percentage of HVAC equipment connected to the BMS.	Verification of integration with the BMS.	One-time measurement, with the possibility of periodic checks.
3.	**Terminal temperature** must remain **between 21°C and 24°C during 95%** of operating time.	Air temperature in the public areas of the terminal.	Digital temperature sensors connected to the BMS.	Continuous, with daily and monthly reports.
4.	Indoor CO2 level must remain below 1000 ppm.	CO2 concentration in the air.	CO2 sensors installed in public areas and connected to the BMS system.	Continuous, with automatic alarms in case of exceedance.

3.3 Application of the "Analyze" Step in the Company "Belgrade Airport"

Building on the measurement data collected in the previous phase, the Analyze phase focused on identifying the root causes of non-optimal HVAC operation. The Brainstorming method was used to systematically generate and examine potential causes of the identified problem, with the full DMAIC team contributing domain-specific insights from the Maintenance, Operations, and ESG sectors.

Overview of the Potential Causes of the Problem. The brainstorming session yielded twelve potential causes, grouped across technical, operational, and organisational dimensions. Table 3 presents the full list as an input to the subsequent cause-and-effect analysis.

Table 3. Overview of the Potential Causes of the Problem

No.	Potential Causes of the Problem
1.	The HVAC system operates in the same mode regardless of the actual occupancy of the space.
2.	Insufficient integration of HVAC equipment with the BMS system.

3. The HVAC system is not synchronized with the flight schedule and passenger flow.
4. Lack of automated procedures for reducing system operation outside peak hours.
5. Insufficient staff training in managing energy-efficient HVAC operating modes.
6. Limited cooperation between the technical, operational, and ESG sectors.
7. Variable passenger movement patterns and unpredictable terminal occupancy.
8. Misalignment of HVAC system operation with actual weather conditions, such as sudden drops or rises in outdoor temperature.
9. Inadequate thermal insulation in certain parts of the terminal building due to structural constraints.
10. Faulty or inaccurate temperature and CO2 sensors.
11. Insufficient monitoring and analysis of historical energy consumption data.
12. Lack of clearly defined responsibility for energy consumption optimization.

Presentation of the Potential Causes of the Problem Using a Cause-and-Effect Diagram. The cause-and-effect diagram aims to enable the DMAIC team to identify possible causes of the observed problem by using the Brainstorming method, visualize potential causes, and illustrate the relationships between them [13].

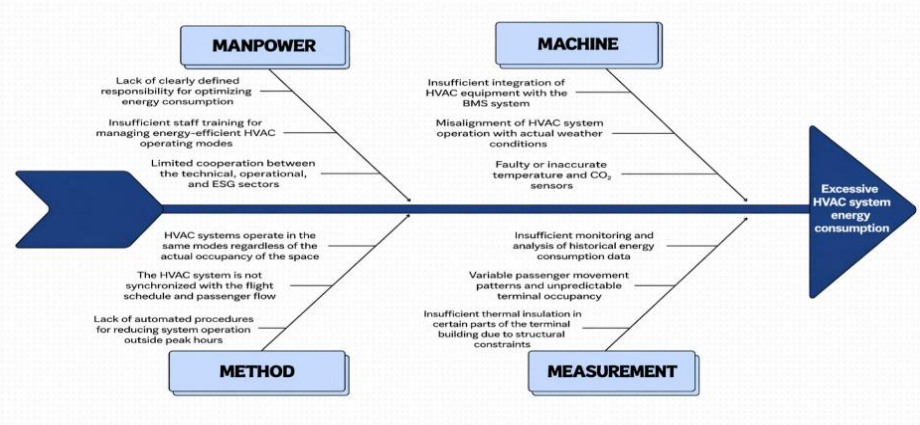


Fig. 4. Cause-and-effect (Ishikawa) diagram — non-optimal HVAC operation during periods of low occupancy.

Fig. 4 presents an Ishikawa diagram (cause-and-effect diagram) showing the potential causes of the problem of non-optimal operation of the heating, ventilation, and air conditioning system during periods of low space occupancy. The diagram served as the basis for the subsequent quantitative prioritisation of causes, in which each identified cause was assigned a frequency of occurrence by the project team based on operational records and expert assessment.

Overview of the Frequency of Occurrence of the Potential Causes of the Problem Before the Introduction of Improvements. Based on the team assessment described

above, Table 4 presents the identified causes ranked in descending order of their estimated frequency of occurrence, providing the basis for Pareto prioritisation.

Table 4. Frequency of Occurrence of the Potential Causes of the Problem Before the Introduction of Improvements

	Potential Causes of the Problem	Frequency	Cumulative Frequency	Share (%)	Cumulative Share (%)
1.	HVAC systems operate in the same mode regardless of the actual occupancy of the space.	208	208	41.6%	41.6%
2.	The HVAC system is not synchronized with the flight schedule and passenger flow.	152	360	30.4%	72.0%
3.	Lack of automated procedures for reducing system operation outside peak hours.	44	404	8.8%	80.8%
4.	Insufficient integration of HVAC equipment with the BMS system.	28	432	5.6%	86.4%
5.	Insufficient monitoring and analysis of historical energy consumption data.	19	451	3.8%	90.2%
6.	Variable passenger movement patterns and unpredictable terminal occupancy.	12	463	2.4%	92.6%
7.	Insufficient staff training in managing energy-efficient HVAC operating modes.	10	473	2.0%	94.6%
8.	Misalignment of HVAC system operation with actual weather conditions, such as sudden drops or rises in outdoor temperature.	8	481	1.6%	96.2%
9.	Faulty or inaccurate temperature and CO_2 sensors.	7	488	1.4%	97.6%
10.	Limited cooperation between the technical, operational, and ESG sectors.	5	493	1.0%	98.6%
11.	Lack of clearly defined responsibility for optimizing energy consumption.	4	497	0.8%	99.4%
12.	Inadequate thermal insulation in certain parts of the terminal due to structural constraints.	3	500	0.6%	100%
	Total	500	/	100%	/

The results indicate that three causes collectively account for over 80% of the identified occurrences, while HVAC operation that does not respond to actual occupancy accounts for 41.6% on its own. This points to a concentrated problem structure suitable for Pareto-based prioritisation.

Pareto Diagram Before Improvement Implementation. The Pareto analysis confirmed that three causes collectively account for approximately 80% of the identified problem: non-occupancy-responsive HVAC operation (41.6%), lack of synchronisation with the flight schedule (30.4%), and the absence of automated off-peak reduction procedures (8.8%). Notably, all three priority causes are directly linked to predictable airport operational patterns, passenger flow variability, flight schedule structure, and seasonal occupancy cycles, which confirms that the problem is systemic rather than isolated to a single technical failure.

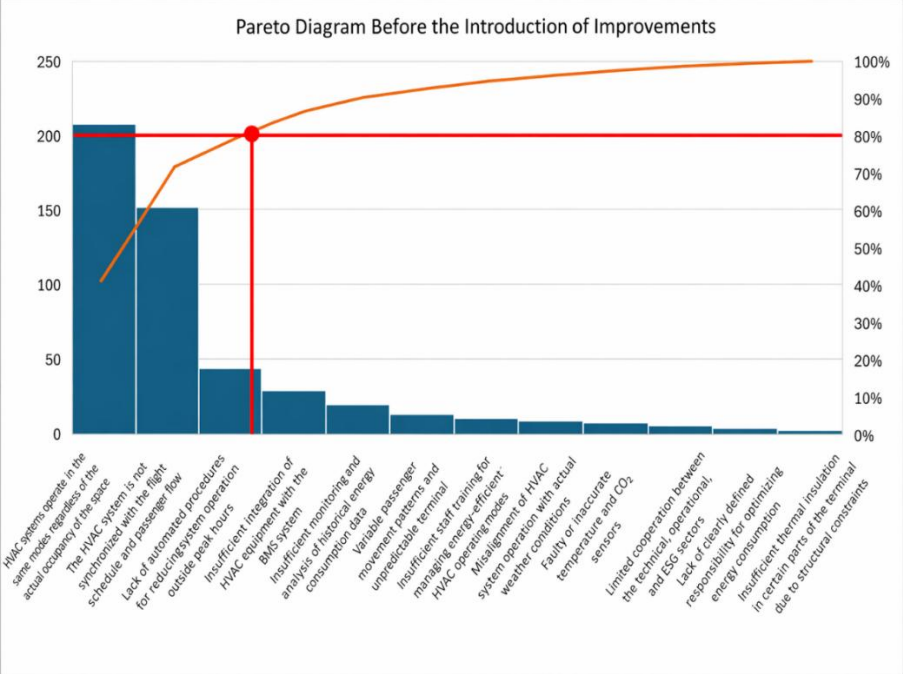


Fig. 5. Pareto diagram of cause frequencies before improvement implementation.

Fig. 5 presents the Pareto diagram, confirming the 80/20 distribution and visually identifying the three priority causes that form the basis for the subsequent improvement phase.

Matrix of Potential Solutions. The three priority causes identified through the Pareto analysis served as the basis for generating potential improvement solutions. For each cause, two alternatives were proposed through team brainstorming and evaluated using a scoring matrix. Each solution was assessed across three criteria — implementation cost, ease of implementation, and estimated share of the problem addressed — on a 1–3–9 scale where 9 represents the most favourable outcome. Scores were assigned by the DMAIC team based on operational expertise, and the total ranking was calculated as the product of the three scores.

Table 5. Matrix of Potential Solutions

No.	Which Cause Does It Relate To?	Potential Solution	Cost Ranking	Ease of Implementation	% of the Problem Solved	Total Ranking
1.	HVAC systems operate in the same modes regardless of the actual occupancy of the space.	**1.1.** Implement HVAC system control based on actual space occupancy using presence sensors.	9	3	9	243
		1.2. Establish separate HVAC operating modes for different terminal zones in accordance with their purpose and load.	3	3	9	81
2.	The HVAC system is not synchronized with the flight schedule and passenger flow.	**2.1.** Adjust the HVAC system to operate in enhanced mode only during flight arrival and departure periods, and in reduced mode during all other periods.	1	9	9	81
		2.2. Align the HVAC operating schedule with the daily flight schedule in order to avoid full system capacity when the terminal is lightly occupied.	1	9	3	27
3.	Lack of automated procedures for reducing system operation outside peak hours.	**3.1.** Introduce automatic switching of the HVAC system to a reduced operating mode during night-time and off-peak periods.	3	3	9	81
		3.2. Define HVAC operating time schedules that automatically reduce heating, cooling, and ventilation when the terminal is less occupied.	1	9	3	27

Solution 1.1 achieved the highest total ranking (243), significantly outperforming all alternatives. Beyond its numerical superiority, this solution directly and permanently addresses the highest-frequency cause through full automation, eliminating dependency on manual scheduling adjustments. Unlike rule-based alternatives (Solutions 2.1 and 2.2), occupancy-based control responds dynamically to actual terminal conditions, a

critical advantage given the variable and seasonally unpredictable nature of passenger flow at Belgrade Nikola Tesla Airport. The moderate ease-of-implementation score (3) reflects hardware installation requirements, but this is a one-time constraint offset by long-term operational adaptability and alignment with ESG energy reduction objectives.

3.4 Application of the "Improve" Step in the Company "Belgrade Airport"

Following the selection of Solution 1.1 as the optimal improvement measure, the project team conducted a projected reassessment of cause frequencies in order to estimate the anticipated impact of occupancy-based HVAC control. The reassessment was based on a structured expert evaluation by the DMAIC team, taking into account the expected operational effects of the proposed solution on each identified cause.

Overview of the Frequency of Occurrence of the Potential Causes of the Problem After the Introduction of Improvements. Table 6 presents the projected frequencies after the proposed improvement, based on the DMAIC team's assessment of how Solution 1.1 is expected to reduce the occurrence of each identified problem, both directly and through its indirect effects on related causes.

Table 6. Frequency of Occurrence of the Potential Causes of the Problem After the Introduction of Improvements

No.	Potential Causes of the Problem	Frequency	Cumulative Frequency	Share (%)	Cumulative Share (%)
1.	HVAC systems operate in the same mode regardless of the actual occupancy of the space.	147	147	40.38%	40.38%
2.	The HVAC system is not synchronized with the flight schedule and passenger flow.	109	256	29.95%	70.33%
3.	Lack of automated procedures for reducing system operation outside peak hours.	34	290	9.34%	79.67%
4.	Insufficient integration of HVAC equipment with the BMS system.	23	313	6.32%	85.99%
5.	Insufficient monitoring and analysis of historical energy consumption data.	16	329	4.4%	90.39%
6.	Variable passenger movement patterns and unpredictable terminal occupancy.	10	339	2.75%	93.14%

7.	Insufficient staff training in managing energy-efficient HVAC operating modes.	8	347	2.2%	95.34%
8.	Misalignment of HVAC system operation with actual weather conditions, such as sudden drops or rises in outdoor temperature.	6	353	1.65%	96.99%
9.	Faulty or inaccurate temperature and CO_2 sensors.	5	358	1.37%	98.36%
10.	Limited cooperation between the technical, operational, and ESG sectors.	3	361	0.82%	99.18%
11.	Lack of clearly defined responsibility for optimizing energy consumption.	2	363	0.55%	99.73%
12.	Inadequate thermal insulation in certain parts of the terminal due to structural constraints.	1	364	0.27%	100%
	Total	364	/	100%	/

Pareto Diagram After Improvement Implementation. Fig. 6 illustrates the projected reduction in cause frequencies following the implementation of Solution 1.1, with all twelve causes showing a decrease in estimated occurrence. The three priority causes still account for approximately 80% of the remaining projected occurrences, which is consistent with the pre-improvement Pareto structure. This indicates that, although the selected solution is expected to produce a meaningful overall reduction, the relative dominance of the main causes remains. Therefore, the monitoring mechanisms defined in the Control phase are essential for comparing actual post-implementation outcomes with the projections presented in this analysis.

Table 7. Fig. **6.**

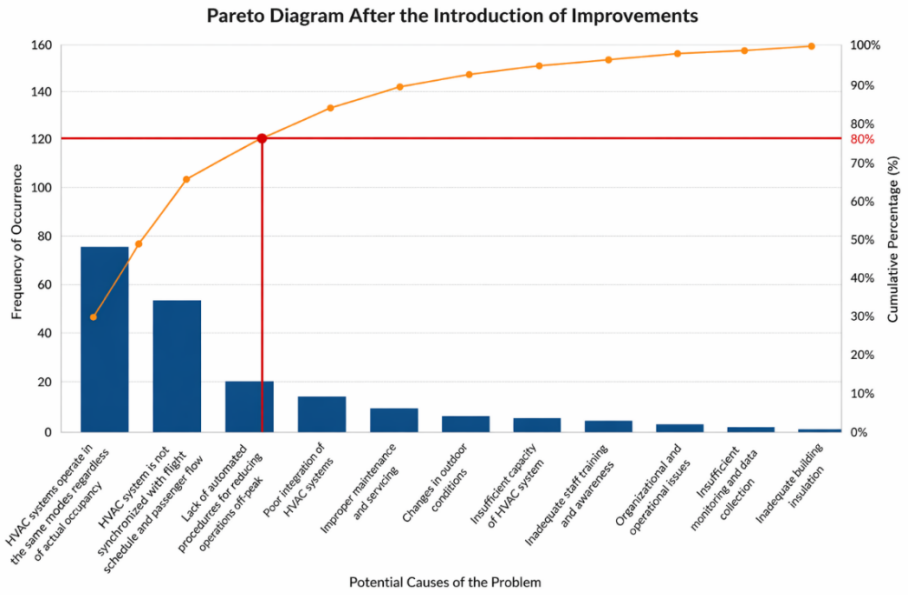


Fig. **6.** Pareto diagram of projected cause frequencies after improvement implementation.

3.5 Application of the "Control" Step in the Company "Belgrade Airport"

The solution selected as optimal in this case is the implementation of HVAC system control based on actual space occupancy using presence sensors. This solution ensures that the HVAC system no longer operates continuously at the same intensity, but automatically adjusts its mode of operation depending on whether people are present in the space and how many there are, as detected by presence sensors.

When the space is highly occupied (e.g., during check-in periods or in gate areas before boarding), the HVAC system operates in normal or enhanced mode, aiming to maintain a stable temperature and increase ventilation. When the space is partially occupied, the system switches to a moderate mode, with reduced airflow and lower heating or cooling intensity. When there are no people in the terminal building (e.g., at night or in unused areas), the HVAC system operates in a reduced mode.

Presence sensors detect whether people are present in a specific zone and transmit this information to the Building Management System (BMS). Based on occupancy data, the BMS automatically adjusts the HVAC system's operation, increasing or decreasing heating, cooling, and ventilation according to actual needs.

The control logic is designed to respond to the inherently dynamic nature of airport operations, in which terminal occupancy varies significantly across daily cycles, seasonal demand patterns, and irregular flight schedules. By linking HVAC operation directly to real-time occupancy data through the BMS, the proposed solution embeds adaptive control within the airport's broader operational framework — ensuring the

improvement remains effective under varying conditions without requiring continuous manual adjustment. This systemic approach moves the improvement beyond a single corrective measure and positions it as a sustainable component of the airport's energy management infrastructure, directly supporting long-term ESG performance targets.

Methods for Monitoring Results After the Implementation of Improvements.
The following methods will be used to monitor the effects of the implemented solution:

- **Monitoring the electrical energy consumption of the HVAC system** – Data on electricity consumption will be regularly collected by tracking electricity meters and the Building Management System (BMS). The terminal building can be divided into zones, allowing energy consumption to be measured per zone. The collected data (daily, monthly, annually) will be compared with baseline data to determine whether a reduction in electricity consumption has been achieved;
- **Analysis of HVAC system operating time by modes** – The BMS system enables automatic recording of HVAC operating modes over time. It collects data on how long the HVAC system operates in different modes, both when the terminal building is empty and when it is occupied, making it possible to identify periods when the system operates at full capacity unnecessarily;
- **Monitoring air quality levels** – The most commonly monitored indicators include carbon dioxide concentration, relative humidity, and temperature. CO2 is considered one of the most important indicators of ventilation adequacy and air freshness in indoor spaces.

Future Plans for Work and Improvement. The following targets represent projected benchmarks derived from the frequency reduction analysis conducted in the Improve phase. They will be validated against actual BMS and electricity metering data during the monitoring period.

- Reduce the electrical energy consumption of the HVAC system by at least 15% within the next 12 months;
- Maintain CO2 concentration below 1,000 ppm and temperature within the defined comfort range for at least 95% of the time over the next 12 months;
- Reduce the operating time of the HVAC system at full capacity during periods without user presence by at least 30% within 9 months.

4 Conclusion

This paper has demonstrated that the DMAIC methodology can be effectively applied to airport infrastructure management, specifically to the problem of non-optimal HVAC energy consumption during periods of low terminal occupancy. Through a structured five-phase process, the study identified three root causes responsible for approximately 80% of observed energy inefficiency: non-occupancy-responsive system operation (41.6%), lack of synchronisation with flight schedules (30.4%), and the absence of

automated off-peak reduction procedures (8.8%). The selected improvement — occupancy-based HVAC control via presence sensors — is projected to reduce total problem occurrences by approximately 27%. These findings extend the existing DMAIC literature [3, 10, 12, 14], confirming the methodology's applicability to complex, multi-causal infrastructure settings. A distinguishing feature of this application is the explicit integration of ESG requirements across all five DMAIC phases — not as contextual background, but as a guiding framework for defining quality requirements, evaluating improvement alternatives, and establishing monitoring targets — representing a contribution that extends beyond applying DMAIC in a new domain.

While the specific parameters of this study reflect the operational context of Belgrade Nikola Tesla Airport, the methodological approach is transferable to other airport environments and, more broadly, to any large infrastructure system characterised by variable occupancy patterns and ESG-related energy reduction obligations. From a practical perspective, the monitoring plan and projected targets defined in the Control phase provide a concrete foundation for ongoing performance tracking and process enhancement, contributing to the broader goal of sustainable airport infrastructure management.

Several limitations should be acknowledged. Post-improvement frequency data are based on projected team assessments rather than empirically measured outcomes, as the solution had not yet been fully implemented at the time of writing. The study also relies on a single-case design and expert-based cause frequency estimation, which introduces subjectivity into the prioritisation process. Future research could address these limitations through longitudinal monitoring of post-implementation outcomes and comparative studies across multiple airport facilities.

Disclosure of Interests

The author declares no competing interests relevant to this study.

References

1. Ahmady, G.A., Mehrpour, M., Nikooravesh, A.: Organizational structure. Procedia Soc. Behav. Sci. 230, 455–462 (2016). https://doi.org/10.1016/j.sbspro.2016.09.057
2. Baraka, J.-C.M., Yadavalli, V.S.S.: The use of DMAIC Six-Sigma for productivity "gap" reduction in manufacturing production line. In: Proceedings of the 2nd African International Conference on Industrial Engineering and Operations Management, pp. 961–972. IEOM Society International (2020)
3. Baro, M., Piña, M., Reyes, A.: 6 Sigma and DMAIC method: basic tool teaching and application for beginning practitioners in automotive assembly. J. Optim. Ind. Eng. 16(1), 89–96 (2023). https://doi.org/10.22094/JOIE.2023.1976488.2029
4. Bloja, M.-D., Moica, S., Veres, C.: Lean Six Sigma in the energy service sector: a case study. Procedia Manuf. 46, 352–358 (2020). https://doi.org/10.1016/j.promfg.2020.03.051
5. Cronemyr, P.: DMAIC and DMADV – differences, similarities and synergies. Int. J. Six Sigma Compet. Advant. 3(3), 193–209 (2007). https://doi.org/10.1504/IJSSCA.2007.015070

6. Feng, Z.: ESG risk assessment in the aviation industry: a comparative analysis of Airbus, Boeing, and Bombardier's financial and sustainability strategies. In: Proceedings of the 3rd International Conference on Financial Technology and Business Analysis, pp. 184–185. EWA Publishing (2024)
7. Ifrim, A.M., Bițan, G.E., Maier, D., Fogoroș, T.E.: Improving the performance of organizational innovation processes by applying the Six Sigma methodology. In: Proceedings of the 14th International Conference on Business Excellence, pp. 1098–1108. Sciendo (2020). https://doi.org/10.2478/picbe-2020-0103
8. Jash, U., Nidhi, C.S., Prasad, N., Satyaprasad, D., Sohan, J.P., Vasudeva, M.R., Aithal, P.S.: Aviation industry analysis with special reference to logistics and supply chain management. Poornaprajna Int. J. Teach. Res. Case Stud. 2(2), 386–395 (2025)
9. Jia, X., Macário, R., Buyle, S.: How do airports align with sustainability? An analysis of the world's 150 busiest airports. Sustain. Dev. 32, 7410–7425 (2024). https://doi.org/10.1002/sd.3084
10. Kaid, H., Noman, M.A., Nasr, E.A., Alkahtani, M.: Six Sigma DMAIC phases application in Y company: a case study. Int. J. Collab. Enterp. 5(3–4), 181–197 (2016)
11. Lee, S., Park, J.-W., Choi, D.: The effects of ESG management on business performance: the case of Incheon International Airport. Sustainability 15(24), 16871 (2023). https://doi.org/10.3390/su152416871
12. Mijajlevski, A.: The Six Sigma DMAIC methodology in logistics. In: Proceedings of the 1st Logistics International Conference, p. 228. Belgrade (2013)
13. Mohamad Zin, M.S., Mahmood, A.R., Wan Mahmood, W.H., Rahim, F.: Six Sigma analyze phase for new product development prototype process improvement. Int. J. Eng. Technol. 7(2.29), 79–85 (2018). https://doi.org/10.14419/ijet.v7i2.29.13135
14. Monday, L.M.: Define, measure, analyze, improve, control (DMAIC) methodology as a roadmap in quality improvement. Glob. J. Qual. Saf. Healthc. 5(2), 44–46 (2022). https://doi.org/10.36401/JQSH-22-2
15. Powell, T., Sammut-Bonnici, T.: Pareto analysis. In: Cooper, C.L. (ed.) Wiley Encyclopedia of Management. John Wiley & Sons, Chichester (2014). https://doi.org/10.1002/9781118785317.weom120202
16. Soliński, B.: Analysis of Six Sigma tools utilization in phases of DMAIC cycle. Decis. Mak. Manuf. Serv. 15, 6–7 (2021). https://doi.org/10.7494/dmms.2021.15.6892
17. Sony, M., Naik, S.: Six Sigma with C-K theory for innovations in operational excellence: a case study. Benchmarking Int. J. 26(7), 2105–2121 (2019). https://doi.org/10.1108/BIJ-08-2018-0241
18. Vishwakarma, J., Kumar, S.: Navigating the skies: an analysis of ESG practices in the airline industry. Sci. Temper 17(1), 5501–5502 (2026)

THEMATIC SECTION

Redesigning Finance for Permanent Disruption

Resilient and Intelligent Organizations
in Times of Permanent Disruption

SYMORG 2026 • CONFERENCE PROCEEDINGS

Original Research Article *https://doi.org/10.5281/zenodo.23478734*

When Brand Valuation Fails: Geopolitical Risk Blindness and Private Equity Due Diligence

Marko Jovanović[1][0009-0009-2646-4325], Maja Đurica[2][0009-0004-0217-3340], Miha Marič[3][0009-0001-6689-7944](✉)

[1] Invigo Capital Group, Belgrade, Serbia;
[2] BAPUSS, Belgrade, Serbia
[3] Faculty of Organizational Sciences, University of Maribor, Slovenia
miha.maric@um.si

Abstract. The standard brand valuation methodologies, including the main ISO 10668 standard, do not consider geopolitical risk adjustments at the brand's revenue stream level. This leaves the brand valuation as a standalone procedure in the jurisdictional concentration risk calculation. This paper examines this gap through the case of Pink Panda, a digital beauty brand acquired by Studio Moderna Group in 2016, whose valuation did not account for the important factor of the group's infrastructure dependency on two geopolitical volatile markets, Russia and Belorussia. The authors direct involvement in the following Studio Moderna's insolvency procedure and independent due diligence, the paper identifies three valuation failures: a common blind spot in all major brand valuations methodologies, company's going concern concept on revenues excluding the risk adjustments on the politically unstable markets and valuing a brand as a standalone asset from the group operations. The following collapse of Studio Moderna in August 2023, while Pink Panda's survival in ten European markets confirmed both the valuation gaps and the private equity kill factor assessments that had identified that risk before the insolvency procedure. The paper's main recommendations are the proposals of geopolitical concentration disclosure, and the private equity kill factor assessments in ISO 10668 standard as a mandatory path for the brand valuations strategy.

Keywords: Brand Valuation, Private Equity Due Diligence, Earnings Capitalization, Going Concern, Geopolitical Risk, Jurisdictional Concentration Risk, Management, Intangible Assets, ISO 10668.

1 Introduction

This paper examines the case of geopolitical market misalignment that was not considered at the time of the brand acquisition. The case is Pink Panda, a Croatian digital beauty brand acquired by Studio Moderna back in 2016, whose valuation was overperformed along with the Groups bankruptcy in August 2023, while the brand itself remained operational.

The key problem was that the standard brand valuations, such as Discounted Cash Flow (DCF) model rely more on the historical figures and cannot predict the war conflicts, sanctions (including EU sanctions [19]), or trade disputes. This led to terminal growth rates being completely miscalculated. This did not just impact their financial statements, but also one of their biggest export markets, as confirmed by public insolvency documentation and founder statements [8], [25]. From the valuation perspective the mistake of WACC and Beta coefficient adjustments were omitted [7].

The structure of this paper is structured in the following way: Section 2 provides an overview of the relevant theoretical foundations of the brand valuation methods and its shortcomings. Section 3 details the methodological approach of the geopolitical risk in corporate valuations. Section 4 reports the Pink Panda´s case and our findings. Section 5 is the private equity perspective of the due diligence and methodologies that were missed. Section 6 presents the paper's discussion on the findings, and Section 7 is the conclusion.

2 Theoretical Part of the Brand Valuation Methods and Their Limits for Pink Pandas valuation

There are several most important brand valuations used in similar scenarios: cost - based approach, market-based approach, royalty relief method, excess earnings method, DCF, and Composite Index model [14].

2.1 Royalty Relief Method

The royalty relief method involves accurately calculating the hypothetical royalty fees that would be saved by owning an intangible asset rather than licensing it by ISO 10668 [12]. This limitation explains why the most commonly used valuation technique was not applicable in our case. In the model they had the assumptions of the projected revenues from 2016 – 2026, without any geopolitical risk adjustments. Accordingly, no geopolitical risk adjustment to the precise royalty rate was correctly applied.

"The royalty relief method is a combination of the market and income valuation approaches. It reflects the market approach in its use of similar licensing deals to calculate an appropriate royalty rate (from databases like RoyaltyRange), and it mirrors the income approach by using estimates of revenue, growth rates, tax rates and discount rates as a basis for value" [19]. This method could be possibly used for these kinds of valuations, but it requires to know precisely the CEE digital beauty sector and to assume the business continuity in the future [12].

2.2 DCF Based Brand Valuations

DCF approaches critically depend on the assumptions with respect to the economic useful life of the intangible asset, projected sales, what represents a reasonable royalty rate, any expenses borne by the owner of the intangible asset, and the appropriate discount rate [14, p.321].

A simple version of DCF would estimate the value (V) in all the valuations of a brand as:

$$V = Ct /(r – g) \quad (1)$$

where Ct = the cash flows from the brand, g = the steady state growth rate of these cash flows, and r = a reasonable discount rate. Some valuations are based on the questionable premise that cash flows represent the royalties from the use of the brands, which ignores ongoing marketing expenses required to maintain the brand value and geopolitical risks. The consultants' original application of DCF earnings capitalization assume an infinite brand useful life and geographically business continuity, which both were not the case [14, p.322].

From the private equity perspective this method was the case in the Pink Pandas valuation, chosen by the default and not as it was the optimal one as per the brands specific history. It has both good and bad parts when applied to early brand valuations, it needs to have the precise and reliable future revenues projections and scenario analysis for the geopolitical disruptions [19, p.229]. The geopolitical part here was not made in the sensitivity's analysis, and the continuity of all markets was supposed, which led to the wrong terminal value calculation (g calculation number in the particular case). In line with the applicable confidentiality obligations, some figures from our internal valuation documentation are described in methodological terms rather than numerically.

2.3 ISO 10668 and ISO 20671 as the statutory relevant standards for the brand valuations, critical point

Until now, there is no study that covered the systematic absence of geopolitical risk adjustments from the brand valuation perspective, nor has the private equity due diligence literature proposed some mechanics for this structural gap. Skalický [29] proposed a risk difference adjustment for the brand uncertainty and Salinas and Ambler [4] assumed that valuation method selection explains the context as asset characteristics, but neither incorporates geopolitical market concentration as one of the most important systematic variable. When reading both standards none of them is referring to the geopolitical scenarios, nor the market concentration disclosures. And they all have the same assumption going concern concept of all the markets where the brand is present in the future projections. This together with the non-jurisdictional risk adjustments represent the gap that was emphasized in the paper both from the academic and business private equity perspective.

From the major brand valuations methods four were not applicable to the Pink Panda case: the cost-based approach, as this is for the stable old tradition companies that have more than 50 years of tradition and huge advertising and marketing costs that can be capitalized (not Pink Pandas case); the market based approach was out of scope because there was no market comparable as Pink Panda was niche digital DTC beauty brand, without a revenue over €1.5m [research team's analysis, insolvency proceedings]; the excess earning method is widely used in small-to-medium business transactions, legal disputes, and tax valuations when market data is unavailable, which was not the case,

from the pure technical reasons and composite index method complicated the scoring methodology and can't be properly allocated to geopolitical risks that Pink Pandas brand already carries.

Table 1. Brand Valuations Methods: Applicability and geopolitical reasons in our Pink Panda acquisition model. Authors' own conclusions based on [12],[14], and case analysis.

Method	Applicable in our case	Inapplicability Reasons	Geopolitical risk
Cost-Based	No	Just six years of the operations	Measures just the historical costs, not the future one under the geopolitical risk
Market - Based	No	No comparable CEE digital brand	Going concern concept in all the future projections with the % growth
Royalty Relief	No	No licensing income in the same product portfolio	Royalty rate comes from the steady markets, no WACC adjustments
Excess Earnings	No	Short profitability history for the brand isolation	On a geographically stable markets with the predictable inflations
DCF	Yes, by default	Practically applicable with some adjustments	Terminal value problem, as also based on the going concern assumption
Composite index	No	No consumer scoring data available	No scores for the contingency situations until they arrive

As Table 1 shows, all the methods cannot be applicable to the early-stage startup brands in the same way, and they share the same structural limitation: none includes the geopolitical risk adjustments and going concern concepts.

3 Geopolitical Risk in Corporate Valuation: A Missing Variable

The country risk is well-established in corporate DCF analysis, it is not incorporated at all into brand valuation frameworks, and it represents a gap that ISO 10668, does not recognize [12]. The misalignment here is the jurisdictional concentration risk that needs to be added in the calculation and future projections however Skalicky [29] proposed a risk difference adjustment, although this was not a systematic variable [29]. Geopolitical Risk in Standard DCF.

Country risk is a well-established variable in corporate DCF frameworks. Damodaran (2022) identifies three primary methods for estimating the country equity risk premium: the default spread on a dollar-denominated sovereign bond, the sovereign CDS spread adjusted for the US baseline, and the default spread derived from

the local currency sovereign rating [17]. These are applicable in the mature markets to prepare the ground for the jurisdiction discount rate.

In politically unstable markets, this figure increases investors' rate of return through WACC and reduce the present value of the projected cash flow given in the projection. This approach has a huge limitation for the brand valuation perspective: it can be measured and applied on the consolidated company's level, not on the individual brand revenue stream in the specific jurisdiction [17].

For Studio Moderna Group where the material share of its revenue comes from a high-risk jurisdiction [31] this will impact on their overall WACC and will affect other countries within the group, especially the mother company. DCF method among all the others assume of no territory discrepancy and treat brand cash flow the same way.

3.1 Geographic Revenue Concentration as Risk Amplifier

The material part of the brand revenue or its infrastructure that supports that revenue is coming from one high risk jurisdiction, the country risk cannot be diversified. The standard WACC calculation treats all the jurisdictions the same, but the brand valuation needs to account this on the revenue stream levels [17]. However, none of the current methodologies require market concentration disclosure. In our case, the brand valuation was done without any adjustment for the group's operational dependency on Russia and the WACC difference can vary by material number of base points, direct methodology misses [8].

3.2 The Gap and Research Question

The brand valuation methodology does not inherit the corporate DCF logic and that leads directly on the over or under brand valuation. That directly means that brands that have geopolitical risk revenues are valued the same as those do not [8]. This gap is not an isolated incident, it shows the disconnection between corporate finance, and M&A where the country risk is well known standard and brand valuations on the other side where it remains absent [8]. This leads to the formulation of our research question: Does geopolitical market presence represent a blind systematic spot in brand valuation methodologies, and can private equity due diligence serve as a secondary defence mechanism in those cases?

4 Case Study: Studio Moderna Group and the Pink Panda Brand

Following Yin, [6] this section explains how the methodology was implemented. The main topic of this paper is that the brand valuation conducted by the external consultants ignored the geopolitical infrastructure of one of the biggest Studio Moderna´s markets.

The research data relies on two different types of evidence. The figures from Eurofound, AJPES regarding the financial position, media reports, interviews on the national TVs and newspapers are based on the publicly available sources that were quoted throughout the paper. The quotes regarding the valuation methods, valuation

multiple 6-7x, expansion plan figures, sensitivity analysis, due diligence percentage, Russia and Belorussia market share rely on the research teams' access to the insolvency proceedings and particular due diligence documentation, subject to NDA obligations. At those points the paper describes those findings in methodological way without the precise figures.

Pink Panda was selected because a professional brand valuation, an operational business collapse, and an independent due diligence process all occurred within the same Group in two years period, enabling the direct comparison between the assumptions and outcomes.

4.1 Corporate Context and Brand Origin

SM Group structure and geographic Expansion. The Group was founded in Slovenia in 1992 by Sandi Češko and Livija Dolanc. It became a leading omni-channel direct-to-consumer retailer in CEE. Their famous brands were Dormeo, Delimano, and Walkmaxx, and the company grew initially through television shopping, then expanding into e-commerce and retail stores across 25 countries [21]. After significant expansion, the group faced financial difficulties, before the insolvency problem in August 2023 [10] [11].

The Russian and Belorussian market was an important part of Studio Moderna´s sales strategy, with three subsidiaries present on that market prior to the invasion in 2022 [8]. According to the publicly announced Groups figures, 10% of total revenues of €500m [31] came from these three subsidiaries, however more important was the infrastructure dependence on the logistics, local operational know-how and distribution on those markets [28]. The operational dependency had several forms. The central warehouse storing jewellery prototypes was physically in Russia and Belorussia, as well with the technical know how and licences for the jewellery designs. The logistics and supply chain for all the groups brands was managed and invoiced through the same entities. When those markets were closed in January 2022, the whole supply chain system for the rest of the world has collapsed. The geopolitical exposure to those jurisdictions was not implemented in the valuations and special risk adjustments in Pink Pandas brand acquisitions.

Pink Panda Brand Beginning and Acquisition. The Pink Panda brand was established in 2010 by Maja Švener and Gregor Koprivnik, young Slovenian entrepreneurs.

The brand had a constant growth before the Studio Moderna´s purchase and expanded just from Slovenian market into several CEE countries. Their revenue grew from €301k in 2012 to €1,500k in 2015, making a compound annual growth rate of almost 70% over that period [18]. The main source of that growth came from the digital marketing platforms on Facebook and Google channels, not from the retail or wholesale brand demand, a distinction material to this brand valuation methodology [9].

In 2016, when Studio Moderna acquired the brand, their management made an expansion plan that targets twelve more European markets by 2018, with projected revenues of additional €5.2 million [authors' analysis based on insolvency proceedings documentation]. The most significant finding here for us, when analysed the insolvency procedure, was that nine from twelve of these markets did not have the significant

operational history in this industry. So, our question here was challenging the growth assumptions and pointing on missing the downside of the need for the sensitivities analysis. [authors' analysis based on insolvency proceedings documentation].

4.2 Brand Valuation at Acquisition, Where the Model Failed

The Methodology and Assumptions. The method applied in this brand acquisition was earnings capitalization based simply on the projected adjusted EBITDA figures. According to the research team's engagement in the subsequent insolvency proceedings and independent due diligence process performed by an independent advisor, the valuation can be described without reference to some of the figures: the multiple applied in the valuation was 6-7x future earnings capitalization. That relied on the twelve new markets expansion plan, missing the historical information for nine of twelve [authors' analysis, insolvency process and independent due diligence documents]. However, having this misleading information about the growth market, the Groups management has adopted this plan.

The Financial Mistakes in the Model. The revenue model made the hockey stick trajectory, that assumed the parallel expanded growth for all twelve markets at the same time, while missing the probability-weighted scenario for the market risk for the operationally empty historical markets [authors' analysis based on financial documentation checked during the insolvency process]. This approach is not in line with usual private equity valuation practice, which requires probability-weighted scenarios when historical data is not applied [15, p.229].

A paid digital traffic cannot be the only relevant input for the revenue model and projections. Advertising expenses increased proportionally with the revenue increase, which indicates that the revenue growth was not increased by the organic brand demand itself, but the investments in the paid acquisition as well. This means that Pink Panda revenue growth was dependent on the advertising spending and does not have its independent demand growth, that the income-based brand valuations methods have as a base assumption [14, p.322].

The Geopolitical Mistake in the Model. A standard DCF brand valuation relies on the following formula:

$$V=Ct/(r-g) \tag{2}$$

where Ct stands for the projected brand cash flows, r for the discount rate, and g is the terminal growth rate [14, p.321]. In the Pink Panda brand valuations, all three variables had the wrong assumptions: no clear projections for the Ct, there was no jurisdiction specific risk premium, and the perpetual growth was assumed for all the markets in Studio Moderna Group [authors' analysis, insolvency proceedings]. These were the structural omissions that implied the wrong future cashflows in the jurisdictions.

One of the most important elements in the calculation was missed in the risk adjustments that the brand infrastructure, that relies on the supplier contracts, logistic and in house operational know how was materially concentrated in the markets of Russia and Belorussia [authors' analysis, insolvency proceedings]. Zero scenario

analysis was made in the sensitivity analysis for the potential distortion of those markets, nor the operational logistic brand support was projected for the Pink Panda´s brand in any of the new entering markets, excepts partially in the parent entity, who carried the operational burden on other huge portfolio of brands [authors' analysis, insolvency proceedings]. That lead to the wrong terminal value calculation on the uniform market going concern assumption [14, p.322].

4.3 The History of Studio Moderna´s collapse and how Pink Panda Survived

In 2022 just before the beginning of the war in Ukraine, Studio Moderna was losing the business control on Russian, Belorussian and Ukrainian markets. So, the first one divested was the Russian entity in January 2022, following the Belorussian entity in couple more weeks, that had a logistic support for all the brands within the group. This was the logical move after the RUB/EUR depreciation of 36% [18] just in January that year, making huge losses in the Group's financial statements and material decline in the Group's total inventory and intangible assets in just single month. Table 2 below describes the key events from the Pink Panda purchase to the formal insolvency proceedings, showing that the brand valuation on the acquisition date can make huge impacts on the company's projection in the near future if not treated in the right way, as we described in our paper.

Table 2. Studio Moderna Group Timeline of the Key Events that lead to the insolvency (2016-2024) [8],[10].

Date	Event	Source used
2010	Pink Pandas founded by young Slovenian entrepreneurs	[13]
2016	Studio Moderna Group acquires the brand and made it part of its portfolio	[18]
31 Dec 2021	Studio Moderna is losing the Russian and Belorussian market share and operations	[10]
6 Jan 2022	Russian and Belorussian companies were closed and in the same month Groups result and assets materially declined	[authors' analysis]
Feb 2022	EU sanctions towards Russia, Council Regulation (EU) No 357/2022	Public record and [19]
April 2022	Private equity funds exit as the shareholders and the old owner repurchase the Group	[20]
Jul 2023	Groups retail division sold to Matratzen Concord	[31]
August 2023	Ljubljana District Court opens the bankruptcy procedure	[8]
2024	Pink Panda brand is operational in ten countries	[13]

Table 2 confirms the paper's core thesis: brand valuation cannot be treated as a standalone asset when the Group carries the material geopolitical risk.

5 Private Equity Perspective: Due Diligence as an Additional Risk Filter

This section explains the private equity perspective of the due diligence as an additional risk assessment framework to the regular brand valuation methodology. It describes the kill- factor methodology used in the third party independent due diligence process related to Studio Moderna Group, and examines how this approach identified the geopolitical risk that the original common brand valuations methods had not taken into account,

5.1 The Context of the Independent Due Diligence Procedure

An independent advisor was engaged to perform an independent due diligence process before the formal insolvency proceedings, as a potential investor related to the Studio Moderna Group. As per the confidentiality agreement signed, some of the findings of that procedure were described in more methodological terms. The team performed the killer factor framework that refers to the identification of critical risks, "red flags," or "deal-breakers" that cause the investors abandon an acquisition [24]. The outcome was that geopolitical revenue concentration in a just one high-risk jurisdiction represented a disqualifying criterion, independent of projected future revenues, WACC or projections. On this basis the potential financing was not proceeded, and the Group entered bankruptcy half year later, which validated the assessment.

5.2 Kill – Factor Methodology Used as an Opponent to the Geopolitical Risk Assessment

In this case assessment we have used the structured private equity due diligence tool, that upfront defines the potential risk categories that leads to the potential investment elimination, no matter how good IRR, EV/EBITDA multiplier and DCF valuation look. This is not a purely quantitative method it also a qualitative one in the disturbed market analysis.

5.3 Due Diligence Findings

After the conducted due diligence procedure, the analysis identified three main structural gaps that were tested by the kill-factor criteria. First is related to the geopolitical infrastructure presence on three high-risks markets [authors' analysis, independent due diligence]. Second was the wrong growth assumption for the exit strategy, not adjusted for the jurisdictional disruption risk. And the final one is the standalone Pink Panda´s valuation without the concept of the infrastructural operating dependency on the same high-risks markets [research team’s analysis, insolvency proceedings and independent due diligence].

The main recommendation for the future is that similar brand valuations and regulatory bodies need to be to include the geopolitical concentration as the mandatory part of the disclosure condition in the ISO 10668 standard [12], not as the optional one. Killer factor logic is the base for the operational implementation.

6 Discussion

6.1 Structural Blindspot in all Methodologies

One thing common for all six major brand valuation methods, as summarized in Table 1 [12], is the structural omission in all of them no matter what assumptions they used for the brand calculations. That mistake is the wrongly formulated brand lifecycle and going concern method throughout the projected period. This is not the personal mistake of the individual valuator; it is a methodological gap that needs to be very deeply considered [4]. Not having this in mind will lead to more future mistakes for the brands that will have the groups infrastructural dependency on high-risk jurisdictions.

6.2 Geographical Concentration as a Hidden Risk

Pink Panda case can be a relevant example of the empirical confirmation that the brand value can be contingent on the group's infrastructure from the politically unstable jurisdictions [authors' analysis, insolvency proceedings]. None of the six valuation methods contains the right mechanism to track this contingency and put it as a relevant figure into the calculation process, although the connection between brand value and company's performances are deeply explained [27].

6.3 Private Equity Due Diligence process as a Controlling Mechanism

During the due diligence phase the due diligence team applied kill- factor method as described in Section 5 which identify immediately the geopolitical concentration risk through qualitative analysis check list. That did not require the financial model, sensitive or scenario analysis, just the publicly available data [8],[20]. Standard valuation is historically based, and private equity methodology is future oriented.

7 Conclusion

This paper pointed out the structural absence of the geopolitical risk calculation and adjustments in the brand valuation methodology through the case of Slovenian digital beauty brand Pink Panda. The acquisition by the local key player Studio Moderna at that time relied on the insufficient brand valuation methodology that missed the fact that Pink Panda has a significant group infrastructure dependent on high-risk political jurisdictions. While this paper is limited to a single case perspective, all three findings point in the same conclusion: market continuity treatment cannot be a prearranged assumption, it needs to be treated as a stress test variable in the brand valuation process,

The future research would examine a bigger sample of the similar brand valuations particular in the geopolitically unstable jurisdictions between 2017 and 2023.

This case is not the exception; this can be applied to all the income-based brand valuation models whose revenues are driven from the geographically unstable jurisdictions with higher political risks. Until the main standard ISO 10668 requires mandatory concentration disclosure risk and political risk adjustments, equivalent errors will be made in the comparable transactions and that can lead to the same outcome that the brand will outlive the company that acquired it.

Disclosure of Interests. The authors declare a potential conflict of interest: one author was professionally engaged in the insolvency and valuation proceedings and independent due diligence process related to Studio Moderna Group Slovenia, which forms the empirical basis of this paper. This involvement is explicitly acknowledged in Section 5 of this paper. Readers should know that this professional involvement in the due diligence process is a potential source of bias. The whole research team has pointed out this by separating public evidence from the NDA privileged access throughout the paper and by describing the confidential findings in the methodological terms, rather than independently verifying each.

References

1. Aaker, D.A.: Managing Brand Equity. Free Press, New York (2021)
2. Dong, F., Doukas J.: The Role of Intangible Assets in Shaping Firm Value, European Financial Management 31(4), 1325–1353 (2025) https://doi.org/10.1111/eufm.12547
3. France, S.L., Davcik, N.S., Kazandjian, B.J.: Digital brand equity: The concept, antecedents, measurement, and future development. J. Bus. Res. 192, 115273 (2025). https://doi.org/10.1016/j.jbusres.2025.115273
4. Belesis, N., Kampouris, C., Andreas, F., Malliari, N.: The impact of marketing brand value on a company's financial performance. Int. Rev. Manag. Mark. 16(2), 263–272 (2026). https://doi.org/10.32479/irmm.22354
5. Ash, K.: Brand equity valuation after merger and acquisition: A review of evidence. J. Serv. Sci. Manag. 16, 133–143 (2023). https://doi.org/10.4236/jssm.2023.162009
6. Yin, R.K.: Case Study Research and Applications, 6th edn. SAGE Publications, Thousand Oaks (2018)
7. Bremmer, I.; Managing Risk in an Unstable World. Harvard Business Review, 83(6), 102-111 (2005)
8. Eurofound: Studio Moderna, Bankruptcy in Slovenia. European Restructuring Monitor, Factsheet 200235. Eurofound, Dublin (2023). https://apps.eurofound.europa.eu/restructuringevents/detail/200235 (Accessed 10 April 2026)
9. Davis, J.: Measuring Marketing: 103 Key Metrics Every Marketer Needs, pp. 220–223, 238–244. John Wiley & Sons (Asia), Singapore (2025)
10. AJPES: Studio Moderna d.o.o. financial data. Agency of the Republic of Slovenia for Public Legal Records. https://www.ajpes.si (Accessed 11 April 2026)
11. eBonite.si: Studio Moderna d.o.o. v stečaju. Financial indicators 2021–2022. https://www.ebonitete.si/studiomodernadoo/ (Accessed 10 April 2026)
12. International Organization for Standardization: ISO 10668 Brand Valuation: Requirements for Monetary Brand Valuation. ISO, Geneva (2020)

13. Pink Panda: Company profile. https://si.linkedin.com/company/konverzija-d.o.o. (Accessed 7 April 2026)
14. Salinas, G.: The International Brand Valuation Manual. Wiley, Chichester (2009)
15. Paugam L., : Brand Valuation, Routledge, London (2024)
16. Leite L., Brand valuation: how convergent (or divergent) are global brand rankings and how correlated is brand value to enterprise value? Journal of Marketing Analytics ,110-115 (2024) https://doi.org/10.1057/s41270-023-00234-6
17. Damodaran, A.: Country Risk: Determinants, Measures and Implications. NYU Stern Working Paper (2022). pages.stern.nyu.edu/~adamodar (Accessed 7 April 2026)
18. Central Bank of Russia: Official exchange rates. Bank of Russia, Moscow (2022). https://www.cbr.ru/eng/currency_base/daily/ (Accessed 10 April 2026)
19. European Union: Council Regulation (EU) No 357/2022 of 28 February 2022. Official Journal of the European Union, L 63, 1–6 (2022). https://eur-lex.europa.eu/legal-content/EN/TXT/ ?uri=CELEX:32022R0357 (Accessed 11 April 2026)
20. BusinessWire: Studio Moderna Holdings acquired by Studio Moderna Founders (2022) https://www.businesswire.com/news/home/20220413005224/en/Studio-Moderna-Holdings-Acquired-by-Studio-Moderna-Founders (Accessed 11 April 2026)
21. Studio Moderna: Brand portfolio. https://www.studiomoderna.com/brands (Accessed 10 April 2026)
22. Kunder, D.: The role of branding in shaping business valuation and profitability: A study of brand equity's influence on financial performance. Res. Arch. Rising Scholars (2025). https://doi.org/10.31124/advance.25143756
23. Adıgüzel, H., Ergen, A.: Corporate sustainability reporting and its influence on brand value: A sectoral analysis of top brands in an emerging market. Sustainability 17, 10108 (2025). https://doi.org/10.3390/su172210108
24. Mineola Search Partners: Hard to kill — a four-part framework for evaluating small business acquisitions (2025). https://mineolasearchpartners.com/2025/12/11/hard-to-kill-a-four-part-framework-for-evaluating-small-business-acquisitions (Accessed 10 April 2026)
25. Češko, S.: Interview on Studio Moderna insolvency. N1 Info Television, Belgrade, August 2022;
26. Bloomberg Adria: Studio Moderna uradno v stečaju. 28 August 2023. https://si.bloombergadria.com/posel/poslovanje/38445/studiomodernauradnovstecaju/news (Accessed 8 April 2026)
27. Naila, M., Hakim, M.S.: Unveiling the link between brand value and firm performance: A systematic literature review. J. Account. Finance Manag. 6(4) (2025). https://doi.org/10.38035/jafm.v6i4
28. PwC UK: Studio Moderna Holdings acquired by founder, Sandi Cesko. PwC Deals (2022) https://www.pwc.co.uk/services/deals/recent-deals/ studio-moderna-holdings-acquired-by-founder-sandi-cesko.html (Accessed 10 April 2026)
29. Skaličký, R.: Brand valuation: An innovative approach based on the risk difference. Oeconomia Copernicana 12(2), 415–438 (2021). https://doi.org/10.24136/oc.2021.014
30. Contzen, T., Werry, N.A.F.: The impact of intangibles — unlock unseen value. Deloitte Legal (2025). https://www.deloittelegal.de/dl/en/services/legal/perspectives/the-impact-of-intangibles.html, (Accessed 8 April 2026)
31. Slovenia Times: Storied retailer files for bankruptcy (2023). https://sloveniatimes.com/39584/storied-retailer-files-for-bankruptcy (Accessed 8 April 2026)

Original Research Article *https://doi.org/10.5281/zenodo.23478789*

Linking Financial Literacy to Financial Well-being: Evidence from Serbia

Sladjana Barjaktarovic Rakocevic[1][0000-0002-8114-1256], Rade Rakocevic[2][0009-0002-6653-813X], Nela Rakic[1][0009-0009-3170-8064]

[1] University of Belgrade, Faculty of Organizational Sciences; [2]Senzal Capital

Abstract. Scientific studies show that financial literacy, essential to the financial well-being of individuals and society, remains inadequate in both developing and emerging countries. This research is the first to assess financial literacy in Serbia within this scope. The findings emphasise the need to implement financial literacy programs in educational institutions, as formal education provides significant benefits. These results can help educational institutions, individuals, and policy-makers foster sustainable economic growth.

Keywords: financial literacy, financial well-being, financial education

1 Introduction

Financial literacy is a crucial skill that individuals should develop throughout their lives. It plays a significant role in fostering personal and societal financial well-being, inclusion, and empowering individuals as well as micro, small, and medium enterprises [8]. Grasping basic financial and macroeconomic ideas such as budgeting, saving, inflation, debt management, interest rates, and investing, it is vital for personal financial independence and overall health. The importance of financial literacy is especially pronounced in developing countries like Serbia, which are still transitioning from a post-communist era to a modern, democratic, and developed society.

Since 2020, the world has faced numerous health, geopolitical, and macroeconomic challenges that have profoundly changed people's perspectives on life. Meanwhile, technological progress has made it easier to access global financial products and tools from anywhere, emphasizing the importance of financial literacy for reaching financial well-being through a diverse international investment portfolio.

Financial literacy is a person's ability to understand basic financial concepts and make decisions that support their financial goals. OECD [19] describes it as a mix of awareness, knowledge, skills, attitudes, and behaviors essential for making informed, sound financial choices. When integrated with improved access to financial services and adequate consumer protection, it creates a comprehensive approach to boosting financial resilience and well-being.

According to the European Commission, financial literacy involves the knowledge and skills essential for making key financial decisions. The Financial Consumer Agency of Canada defines financial literacy as the blend of knowledge, skills, and confidence required for responsible financial decision-making. Knowledge encompasses

understanding both personal and broader financial matters; skills allow for applying this knowledge in everyday situations; confidence offers self-assurance in making important decisions; and responsible financial choices involve leveraging one's knowledge, skills, and confidence to choose options that fit one's unique circumstances.

Financial literacy has become a focal point in scientific research due to its increasing significance. International organizations like the OECD, the World Bank, the European Commission, and the Federal Reserve Bank, along with governments from countries such as Canada, Australia, Singapore, the Netherlands, and Scandinavian nations, recognize its importance. Regardless of a nation's level of financial market development, improving financial literacy should be a universal goal to boost financial understanding [10]. Latest OECD data [22] covering 39 countries shows an average financial literacy score of 60 out of 100, highlighting the ongoing need to promote financial education worldwide.

Financial well-being refers to the balance between income and expenses that fosters a stress-free life and supports the attainment of personal financial objectives. It includes maintaining both current and future living standards and enjoying financial independence [2]. The Consumer Financial Protection Bureau (2015) describes financial well-being as a state where individuals can fulfill their current and future financial obligations, feel secure about their financial prospects, and make choices that enhance their quality of life. Additional definitions are offered by researchers and international bodies [22].

To the authors' knowledge, no scientific study on the relationship between financial literacy and financial well-being among adults has been published in Serbia so far. The only relevant research is the 2014 S&P Global Financial Literacy Survey, which examined financial decision-making. This survey included five questions about risk diversification, inflation, interest, and two on compound interest [9]. The number of Serbian respondents is unknown. Serbia's financial literacy score was 38% out of 100, below the EU average of 52% and the US score of 57%. Another 2018 study assessing OECD financial literacy among 15-year-olds found Serbia's scores were significantly lower than the OECD average [16].

This study provides a preliminary investigation into financial literacy in Serbia, exploring its influencing factors and impact on financial well-being. It assesses financial literacy levels among highly educated professionals and considers personal factors affecting it. The research also examines how financial literacy relates to these individuals' financial well-being. In Serbia, financial literacy is absent from primary and secondary education curricula, only appearing in university economics and business courses. The authors concentrate on a well-educated professional cohort. Furthermore, the researchers adopted a top-down approach, recognizing that Serbia lacks an institutional framework for financial literacy education. They assume that this well-educated professional group has already acquired basic financial knowledge through their careers and other educational programs. The study utilized a quantitative method to collect data on respondents' financial literacy levels, including demographic information, perceived financial literacy, and perceived personal financial well-being. The relationship between

financial literacy components and financial well-being was analyzed statistically. Future research plans include involving different focus groups, adopting a long-term perspective, and monitoring changes in financial literacy over time.

Serbia currently does not have a comprehensive national strategy for financial literacy, and this research could serve as a foundational step toward creating one. Therefore, key priorities for improving financial literacy should focus on aspects like education level, financial goals, and investment experience. These findings will help direct government and institutional initiatives to tackle financial literacy challenges in Serbia. Additionally, these findings will benefit authorities and researchers in other developing and emerging countries facing similar challenges related to low financial literacy.

2 Theoretical background and hypothesis development

2.1 Financial literacy

Modern finance offers a broad array of products and services that are becoming more accessible to the public. The growth of digital technologies and their increasing role in the financial sector have generated new needs and expectations among users. This also raises concerns about financial literacy—specifically, how well individuals understand financial products and services. Economic literacy is defined as the ability to make informed decisions related to financial planning, debt management, and pension planning, as well as wealth accumulation [13]. Various socioeconomic and demographic factors, such as age, gender, education level, and income, influence financial literacy, a connection supported by numerous studies [7,13]. Among the advantages of higher financial literacy, some authors suggest that it can lead to increased savings and better risk management [4]. Furthermore, improved financial literacy among households can result in greater use of financial services, which boosts competition among providers and enhances financial intermediation. Ultimately, this fosters financial development and helps reduce economic volatility.

Additionally, research in Italy [10] explored the links between financial literacy, knowledge of sustainable finance, sustainable development, and ESG (environmental, social, and governance) factors. The study found positive relationships among financial literacy, sustainable finance, and sustainable investment choices. Evidence shows that financial literacy influences financial inclusion and saving behaviors [15], with some researchers affirming that increased financial literacy benefits financial inclusion [1]. The research further indicates that how financial literacy affects individual investment decisions, with more knowledgeable investors feeling more confident about their financial choices.

2.2 Financial well-being

Financial literacy remains alarmingly low even in advanced economies like the G7 and G20. Nonetheless, increased financial literacy benefits both individual financial health and the overall economy [14]. It is particularly important for those saving for retirement, planning passive income streams, and developing financial resilience [6].

The researchers highlight that managing unexpected expenses and financial shocks is crucial for attaining financial well-being.

Financial capability, often synonymous with financial literacy, includes three key components: financial knowledge, behavior, and skills [28]. Research indicates that financial behavior has the most substantial impact on overall financial well-being, with skills and knowledge playing a secondary role. An individual's attitude towards money significantly influences their financial well-being, alongside factors like financial planning, risk tolerance, and behavior [3]. Moreover, engaging in long-term investing helps promote sustainable financial well-being.

Some researchers view financial well-being as the primary result of financial literacy [22]. They show that an individual's attitude toward personal finances directly affects their financial well-being. Furthermore, they argue that this principle applies at the societal level, where the overall financial literacy of the population significantly affects national longevity, over-indebtedness, quality of life, and future pension income.

Financial well-being refers to an individual's understanding of their daily financial status and their capacity to handle short-term setbacks, meet financial objectives, and maintain enough financial flexibility for decision-making and leisure activities beyond basic needs [29]. It also depends on how people perceive their values and whether they feel those values are being satisfied at a given moment. Effective money management and financial well-being are associated with greater self-control over spending, the ability to manage immediate pleasures, and a proactive mindset [23].

A survey of young workers and families with children found that personal financial control is more important to them than the total amount of money they have [27]. Better financial well-being is linked to improved personal financial management, while debts such as loans and credit card obligations tend to undermine it. Research by [25] shows that the younger generation views financial well-being as encompassing three main aspects: preserving their status, enhancing their lifestyle, and achieving financial freedom.

Older adults are seen as financially secure if they can sustain their current lifestyle in the future. Many groups do not have a clear financial plan to support their ongoing well-being. Additionally, there is a strong link between financial self-efficacy and overall financial health. Financial self-efficacy also correlates very highly with current financial stress and expected future financial security [5].

Based on the literature discussed, we have developed a general hypothesis to guide research on financial literacy.

Hypothesis: Financial literacy has a significant impact on financial well-being.

3 Research design

This study used empirical research methods, where the authors created an online questionnaire to gather data on various facets of financial literacy. These included its components, financial knowledge, behavior, attitudes toward finance and investment, personal financial management, financial goals, and perceived financial well-being.

Conducted in Serbia with adults aged 18 and over, the study contrasts with the OECD toolkit (OECD, 2022), which focuses on objective financial literacy. Instead, this research emphasizes subjective financial literacy, based on respondents' perceptions. To date, Serbia has not conducted an objective evaluation of adults' financial literacy using the OECD framework.

Moreover, various studies employ different methods to evaluate financial literacy [11]. Because financial literacy is a complex concept that is difficult to define and measure, researchers use a range of items and calculation methods across the literature to assess individuals' financial literacy levels [21]. Objective measures assess actual financial knowledge, whereas self-reported measures reflect an individual's perceived financial literacy. This distinction is crucial because financial decisions and behaviours are shaped by both actual knowledge and personal beliefs about one's financial understanding.

In autumn 2024, the authors distributed questionnaires primarily via professional social networks, chat groups, and email. They received responses from 301 participants. Some potential respondents felt they were not qualified to answer questions about financial literacy, even though all participants were adults. A pilot test was conducted with 20 highly educated professionals from diverse industries, and their feedback was incorporated into the final questionnaire. Multiple linear regression was used to examine how sociodemographic factors affect financial literacy and financial well-being. Respondents rated on a four-point scale, with lower numbers indicating weaker financial literacy (Appendix).

Financial well-being was evaluated by averaging responses to question 2, which used a five-point scale; lower scores signified less agreement with statements about financial health (Appendix). The statement "I don't have enough income to cover all living expenses" was reverse-coded because it negatively correlated with the overall score. Sociodemographic variables included gender, age, education, employment, and earnings. Gender is coded as a binary variable, with 0 representing female and 1 representing male. Age is a continuous variable measured in years. Education, employment, and earnings are categorical variables, and dichotomous dummy variables are created for each category in the multiple regression, with 0 indicating non-membership and 1 indicating membership of a specific category.

4 Results

4.1. Descriptive statistics

Table 1 presents the sociodemographic characteristics of the 301 respondents who participated in the study.

Table 1. Sociodemographic characteristics of the respondents

	F	%
Gender		
Female	91	30,2
Male	210	69,8
Age		
18-25	25	8,3
26-35	84	27,9
36-45	94	31,2
46-55	90	29,9
≥56	8	2,7
Education		
Primary school	1	,3
Secondary school	39	13,0
College or Master's	236	78,4
PhD	25	8,3
Employment		
Entrepreneur	83	27,6
Employed in a private company	180	59,8
Public sector	32	10,6
NGO	1	0,3
Part-time jobs	25	8,3
Student	13	4,3
Pensioner	1	0,3
Average monthly income		
From 500 to 1,000 euros	37	12,3
From 1,001 to 1,500 euros	56	18,6
From 1,501 to 2,000 euros	66	21,9
More than 2,000 euros	142	47,2

Note: f- frequency

Source: Author's research

The survey primarily included men, accounting for 210 (69.8%), while women numbered 91 (30.2%). Most respondents had a university or master's degree, totaling 236 (78.4%), whereas only one respondent (0.3%) had finished primary school. Almost half of the participants (142, 47.2%) reported an average monthly income over €2,000.

The employment question allowed respondents to select all relevant employment types through a multiple-choice format. For each type, a dummy variable was created—0 if not specified and 1 if explicitly stated—facilitating the listing of frequencies in the table. Nearly two-thirds of respondents (180, 59.8%) work for companies, while 83 (27.6%) are entrepreneurs. The respondents' average age is middle-aged (M = 39.77; SD = 9.52), with the youngest being 19 and the oldest 69.

Table 2 presents respondents' answers regarding their perceptions of financial literacy.

Table 2. Financial literacy

	f	%
I need a lot of help	23	7,6
I need help	107	35,5
I'm competent	137	45,5
I'm very competent	34	11,3

Note: f- frequency
Source: Author's research

Nearly half of the respondents, 137 (45.5%), see themselves as financially literate, whereas 130 (43.1%) feel they need substantial or some assistance. Overall, half of the participants consider themselves either competent or require help (Md = 3.00). When asked about financial literacy, respondents most frequently rated themselves as competent (Mo = 3.00).

Financial well-being was assessed by averaging respondents' answers on a five-point scale, with lower scores indicating less agreement with statements about financial health. The statement "I don't have enough income to cover all my living expenses" is recoded because it negatively correlates with the overall measure. On average, respondents report very good financial well-being (M = 4.13; SD = 1.00). The distribution of scores on the financial well-being scale is left-skewed (skewness = -1.16) and somewhat peaked (kurtosis = 0.55). Scores range from 1.00 to 5.00. The Kolmogorov-Smirnov test indicated that the scores do not follow a normal distribution (D (301) =0.19; P≤0.001). Internal consistency, measured by Cronbach's α, is very high at 0.835.

4.2. Hypothesis testing

Hierarchical multiple regression was employed to assess the influence of financial literacy on financial well-being. The authors applied this hypothesis while controlling for potential influences of sociodemographic factors like gender, education, and age. The assumptions for the use of multiple regression have been preliminarily checked: multicollinearity, presence of extreme values, normality, linearity, homoscedasticity and independence of residuals.

The minimum tolerance value is 0.43, well above the critical threshold of 0.1, indicating no multicollinearity among the independent variables. The maximum VIF is 2.30, indicating the absence of multicollinearity. Multivariate outliers were checked

using Mahalanobis distances, and none were found. The largest Mahalanobis distance recorded is 19.80, below the critical value of 20.52 for five independent variables. The normality of the distribution was tested using the Kolmogorov-Smirnov test. The residuals do not meet the normality assumption; D (301) = 0.14; $p < 0.001$. However, no significant violations of the assumptions were observed.

The control-variable model was statistically significant, F (4, 296) = 7.43, $p < .001$, $R^2 = .091$, indicating that the control variables accounted for 9.1% of the variation in financial well-being. When financial literacy was added as an additional predictor, the model remained significant (F (5, 295) = 8.25, $p < .001$) and showed increased explanatory power. This indicates that financial literacy contributes independently and importantly to predicting financial well-being, over and above the control variables. Table 9 presents the results of how financial literacy impacts financial well-being.

Table 9. The impact of financial literacy on financial well-being

				95 % an interval of confidence for β	
	B	SE	P	Lower limit	Upper Limit
Constant	0,000	0,056	0,987	-0,110	0,110
Gender	0,217	0,060	<0,001	0,099	0,335
Age	0,019	0,058	0,746	-0,095	0,133
Education – College	0,215	0,068	0,002	0,081	0,349
Education – PhD	0,108	0,069	0,119	-0,028	0,296
Financial Literacy	0,182	0,058	0,002	0,069	0,353

β – standardized regression coefficient, SE- standard error beta, p- significance
Source: Author's calculations

Gender (β=0.22; p<0.001), college education (β=0.22; p=0.002), and financial literacy (β=0.18; p=0.002) each uniquely and significantly explain part of the variance in financial well-being. All three factors—gender, college education, and financial literacy—positively influence financial well-being.

Men tend to have higher financial well-being by 0.22 standard deviations than women. Similarly, obtaining a college education is associated with a 0.22-standard-deviation increase in financial well-being compared with having only a high school diploma. Additionally, respondents with financial literacy report a 0.18-standard-deviation higher level of financial well-being than those without financial literacy.

The findings indicate that financial literacy plays a significant role in enhancing financial well-being when sociodemographic factors are taken into account. Higher financial literacy is associated with better financial well-being. Gender and having a college degree also positively influence financial well-being. Although age and holding a PhD are linked to better outcomes, their effects are not statistically significant. These results are consistent with those reported by [12, 24].

5 Conclusion

Over the past decade, financial literacy has become increasingly important, especially in developing regions where markets are often unstable and face various financial challenges. This skill is vital because it helps individuals make informed choices and supports macroeconomic stability. Those with higher financial literacy are more likely to save and invest, as they better understand risks and can react promptly to financial issues.

This paper explores financial literacy in Serbia, a country lacking formal strategies or programs for financial education. It begins by investigating how people perceive financial literacy and examines how it influences financial well-being. The study used a survey method to target highly educated individuals residing in Serbia, who were asked to assess their financial literacy.

The results of this study may be relevant to educational institutions and government agencies and could lay the groundwork for creating official strategies to improve financial literacy across the country. However, the research has limitations. Most participants are highly educated and earn above-average incomes, with some involved in investing and finance, which could skew perceptions of financial literacy. Data collection was done through questionnaires distributed via professional networks and personal contacts, potentially biasing results toward a higher perception of financial literacy among those in business or finance sectors. Including focus groups or interviews might offer more in-depth insights into how people perceive financial literacy and overall financial well-being.

Current limitations can be addressed by expanding the sample size and incorporating participants from more diverse educational, industry, and income backgrounds. Additionally, future research is expected to encompass countries from the Western Balkans. This study acts as an initial step toward understanding financial literacy in Serbia. Ultimately, forthcoming work will aim to include quantitative assessments of financial literacy and secure support from pertinent financial institutions.

References

1. Bashiru, S., Bunyaminu, A., Yakubu, I., & Al-Faryan, M. (2023). Drivers of financial inclusion: insights from Sub-Saharan Africa. Economies, 11(5), p.146. https://doi.org/10.3390/economies11050146.
2. Brüggen, E. C., Hogreve, J., Holmlund, M., Kabadayi, S., & Löfgren, M. (2017). Financial well-being: A conceptualization and research agenda. Journal of business research, 79, 228-237.
3. Castro-González, S., Fernández-López, S., Rey-Ares, L., & Rodeiro-Pazos, D. (2020). The influence of attitude to money on individuals' financial well-being. Social Indicators Research, 148(3), pp.747-764. DOI: 10.1007/s11205-019-02219-4.
4. Cole, S., & Fernando, N. (2008). Assessing the importance of financial literacy. ADB Finance for the Poor, 9(2), 1-6.
5. Dare, S. E., van Dijk, W. W., van Dijk, E., van Dillen, L. F., Gallucci, M., & Simonse, O. (2023). How executive functioning and financial self-efficacy predict subjective financial

well-being via positive financial behaviors. Journal of Family and Economic Issues, 44(2), 232-248.
6. Estrada-Mejia, C., Mejía, D., & Córdoba, P. (2023). Financial literacy and financial wellbeing: Evidence from Peru and Uruguay. Journal of Financial Literacy and Wellbeing, 1(2), 403-429.
7. Garg, N., & Singh, S. (2018). Financial literacy among youth. International journaL of sociaL economics, 45(1), 173-186.
8. G20. (2021, July 9-10). Italian G20 Presidency Third Finance Ministers and Central Bank Governors Meeting Communiqué. Retrieved from G20Italy, pp.6.: http://www.g20italy.org/wp-content/uploads/2021/07/Communique-Third-G20-FMCBG-meeting-9-10-July-2021.pdf
9. Klapper, L., Lusardi, A., & Van Oudheusden, P. (2015). Financial literacy around the world. World Bank. Washington DC: World Bank, 2, 218-237.
10. Lanciano, E., Previati, D., Ricci, O., & Santilli, G. (2024). Financial literacy and sustainable finance decisions among Italian households. Journal of Economics and Business, p.106220. https://doi.org/10.1016/j.jeconbus.2024.106220.
11. Lim, T. S., Mail, R., Abd Karim, M. R., Ulum, Z. K. A. B., Jaidi, J., & Noordin, R. (2018). A serial mediation model of financial knowledge on the intention to invest: The central role of risk perception and attitude. Journal of Behavioral and Experimental Finance, 20, 74-79.
12. Lone, U. M., & Bhat, S. A. (2022). Impact of financial literacy on financial well-being: a mediational role of financial self-efficacy. Journal of Financial Services Marketing, 1.
13. Lusardi, A., Mitchell, O., & Curto, V. (2010). Financial literacy among the young. Journal of consumer affairs, 44(2), pp.358-380. https://doi.org/10.1111/j.1745-6606.2010.01173.x.
14. Lusardi, A., & Messy, F. A. (2023). The importance of financial literacy and its impact on financial wellbeing. Journal of Financial Literacy and Wellbeing, 1(1), 1-11.
15. Morgan, P., & Long, T. (2020). Financial literacy, financial inclusion, and savings behavior in Laos. Journal of Asian Economics, 68, p.101197. http://dx.doi.org/10.1016/j.asieco.2020.101197.
16. OECD. (2017, May 27). PISA 2015 Results (Volume IV): Students' Financial Literacy, PISA, OECD Publishing, Paris,. Retrieved from OECD: https://doi.org/10.1787/9789264270282-en.
17. OECD. (2022). OECD/INFE Toolkit for Measuring Financial Literacy and Financial Inclusion 2022. Retrieved from OECD Publishing, Paris,: https://doi.org/10.1787/cbc4114f-en (Accessed on 10.11.2024)
18. OECD. (2023, December 14). “OECD/INFE 2023 International Survey of Adult Financial Literacy”. Retrieved from OECD Business and Finance Policy Papers. No. 39, OECD Publishing, Paris: https://doi.org/10.1787/56003a32-en.
19. OECD. (2024a). Retrieved from Financial Education: https://www.oecd.org/en/topics/sub-issues/financial-education.html
20. OECD. (2024b, November 21). G20 policy note on financial well-being, OECD Publishing, Paris, https://doi.org/10.1787/7332c99d-en.
21. Ouachani, S., Belhassine, O., & Kammoun, A. (2021). Measuring financial literacy: A literature review. Managerial Finance, 47(2), 266-281.
22. Philippas, N. D., & Avdoulas, C. (2019). Financial literacy and financial well-being among generation-Z university students: Evidence from Greece. The European Journal of Finance, 26(4–5), 360–381. https://doi.org/10.1080/1351847X.2019.1701512.
23. Ponchio, M. C., Cordeiro, R. A., & Gonçalves, V. N. (2019). Personal factors as antecedents of perceived financial well-being: evidence from Brazil. International Journal of Bank Marketing, 37(4), 1004-1024. https://doi.org/10.1108/IJBM-03-2018-0077

24. Rahman, M., Isa, C. R., Masud, M. M., Sarker, M., & Chowdhury, N. T. (2021). The role of financial behaviour, financial literacy, and financial stress in explaining the financial well-being of B40 group in Malaysia. Future Business Journal, 7(1), 52.
25. Riitsalu, L., Sulg, R., Lindal, H., Remmik, M., & Vain, K. (2024). From Security to Freedom—The Meaning of Financial Well-being Changes with Age. J Fam Econ Iss, 45, pp. 56–69. https://doi.org/10.1007/s10834-023-09886-z.
26. Shah, S., Qureshi, F., Memon, F., & Uddin, M. (2024). Financial literacy and investment behavior of individuals in Pakistan: Evidence from an environment prone to religious sentiment. Journal of Behavioral and Experimental Finance, 44, p.100974. https://doi.org/10.1016/j.jbef.2024.100974.
27. Vlaev, I., & Elliott, A. (2014). Financial Well-Being Components. Soc Indic Res, 118, 1103–1123 (2014). https://doi.org/10.1007/s11205-013-0462-0.
28. Xiao, J., & Porto, N. (2022). Financial capability and well-being of vulnerable consumers. Journal of Consumer Affairs, 56(2), pp.1004-1018. https://doi.org/10.1111/joca.12418.
29. Zhang, Y., & Chatterjee, S. (2023). Financial well-being in the United States: The roles of financial literacy and financial stress. Sustainability, 15(5), p.4505. https://doi.org/10.3390/su15054505.

Appendix

Questionnaire on Financial Literacy in the Republic of Serbia

(Part of the original Questionnaire that relates to this research. The original questionnaire has been prepared with the aim of examining the elements of financial literacy, financial behavior, and financial attitudes and, in addition to the existence of financial goals, to determine their effect on the financial well-being of an individual)

1. How do you assess your competence in terms of financial literacy (Financial literacy can be defined as the ability of people to process and make informed decisions about financial planning, wealth management, investments, retirement plans, and credit)
   1 - I need a lot of help, 2 - I need help, 3 - I'm competent, 4 - I'm very competent

2. How do you see your financial well-being on a five-point scale **(**1 – strongly disagree, 2 – disagree, 3 – neither agree or disagree, 4 – agree, 5 – strongly agree):

	1 – strongly disagree	2 – disagree	3 – neither agree or disagree	4 – agree	5 – strongly agree
I cover all living expenses with my income					
I don't have enough income to cover all living expenses					
I have enough income to cover all living expenses and have above					

3. Gender
   - Female
   - Male

4. Age _____

5. Education
   - Primary school
   - Secondary school
   - College and Master's
   - PhD

6. Employment
   - Entrepreneur
   - Public sector
   - Private copany
   - NGO
   - Part-time jobs
   - Pensioner
   - Student
   - Other_____

7. Average monthly income
   From 500 to 1,000 euros
   From 1,001 to 1,500 euros
   From 1,501 to 2,000 euros
   More than 2

Original Research Article *https://doi.org/10.5281/zenodo.23478822*

The Walking Dead: Causes and Consequences of the Emergence of the Zombie Economy

Dario Maradin[1][0000-0001-9642-6038], Stella Suljić Nikolaj[2][0000-0002-0435-4047], Vesna Buterin[3][0000-0002-1090-6549], and Sara Sadrić[4][0009-0003-2730-5656]

[1,2,3,4] University of Rijeka, Faculty of Economics and Business, Ivana Filipovića 4, 51000 Rijeka, Croatia
[1] dario.maradin@efri.uniri.hr; [2] stella.suljic.nikolaj@efri.uniri.hr; [3] vesna.buterin@efri.uniri.hr; [4] sadricsara@gmail.com

Abstract. The concept of the zombie economy has become increasingly relevant in contemporary economic research and policy discussions. Zombie firms are companies that remain active despite being economically unviable, relying on continued external financial support. Unable to generate sufficient income to cover operating costs and debt obligations, they persist due to assistance from banks, governments, or other financial institutions. Their presence distorts market dynamics and negatively affects overall economic productivity. The main purpose of this study is to examine the role, causes, and consequences of zombie firms in economic systems, focusing on their implications for the economy as a whole. In this research, the primary method is a comprehensive literature review, which enables a critical synthesis of existing theoretical and empirical contributions. The research findings aim to contribute to a deeper conceptual understanding of zombie firms as entities that often reduce the efficiency of the entire economy by distorting the allocation of resources.

Keywords: Zombie Firms, Financial Distress, Economic Crises.

1 Introduction

The term "zombie firm" is increasingly attracting the attention of economists in various countries. These companies are generally considered to be in financial distress; they typically report negative profits over an extended period, are heavily indebted, and have difficulty meeting their interest obligations, yet continue to operate despite having lost all their equity. Firms may become zombies through multiple channels, which can include both external environmental and internal factors. External factors may involve macroeconomic effects such as business cycle fluctuations or changes in interest rates, or industry conditions such as the level of market concentration and competition, while internal factors consist of firms' behaviour, such as managerial decisions, debt policy, and other firm activities [1].

This paper aims to clarify the root causes of the emergence of zombie firms and to examine their long-term effects on the economic system. Through a detailed analysis of historical examples, it identifies how macroeconomic policies, financial structures,

and structural changes in economies encourage or enable the prolonged existence of zombie firms. As a relevant microeconomic topic, investigating the presence of zombie firms in national economies is significant due to the consequences or potential damage these firms can cause to the economy and society.

The main purpose of this study is to examine the role, causes, and consequences of zombie firms in economic systems, focusing on their implications for the economy as a whole. The primary research method is a comprehensive literature review, which enables a critical synthesis of existing theoretical and empirical contributions.

2 The zombie economy phenomenon: causes, development and implications for the economy

Paradoxes frequently arising in contemporary economic systems cause concern among economists, investors and policymakers. One such phenomenon that has attracted significant attention is the concept of the "zombie economy", referring to markets in which companies continue to operate despite unsustainable financial positions. Zombie companies are market participants that survive solely through external financing, whether from commercial banks, government interventions or other financial institutions. Despite this external support, these companies fail to generate positive business results or sustainably address their financial liabilities, thus remaining a persistent burden on broader economic processes.

2.1 Causes of occurrence of zombie firms

Various factors contribute to a company becoming a "zombie", including changes in market conditions, shifts in technology or consumer preferences, and poor management. Zombie companies often emerge during economic crises, when many firms face financial difficulties and survive only due to support from banks and the state. Numerous situations can lead to companies becoming illiquid and insolvent, such as a decline in demand, reduced revenue, and difficult credit conditions. To prevent a slowdown in economic growth or a possible recession, the state often intervenes with rapid monetary measures, providing support through financial packages, subsidies, and affordable loans. These quick stimulus measures allow companies to survive when, under normal market conditions, they would likely declare bankruptcy or go out of business.

2.2 Characteristics and specifics of zombie firms

The zombie economy as a phenomenon is particularly significant because it disrupts healthy market competition, slows economic growth, and hinders the efficient allocation of resources within the economy.

While there is no commonly agreed definition of what constitutes a zombie firm, and the literature has used several different definitions to capture these unproductive companies, zombie firms can broadly be defined as risky, unproductive, and unviable firms that manage to avoid immediate bankruptcy, most likely due to continued support from lenders such as banks, investors, or governments in light of misaligned incentives [2].

A central feature of the zombie economy is the existence of zombie firms, which have substantially weaker balance sheets than non-zombies within the same country and industry. Zombies invest less, are less productive and profitable, hold limited liquid assets, tend to be smaller, and face a higher risk of default. In addition, zombie firms face a slightly higher cost of debt than their non-zombie peers. Zombie companies barely manage to repay their debts and survive only through external assistance, such as state subsidies or favourable loans from financial institutions. These companies do not generate sufficient income to properly service their obligations to creditors and are therefore dependent on refinancing or debt restructuring. They are unable to effectively restructure their operations or achieve the necessary profitability for sustainable growth. This phenomenon often results from poor business models, outdated technologies, or unfavourable market conditions that limit their growth. Nevertheless, despite ongoing insolvency, such companies remain in the market, enabled by loose financial policies or political interventions.

Another important characteristic of a zombie firm is reliance on external financial sources to continue operations, i.e. dependence on long-term subsidies and loans. Instead of financing their activities independently, they survive through external financial injections, usually from state funds. These companies also often receive favourable loans or financial assistance from banks, and in some cases, direct government support. Especially in developed economies, banks tend to refinance zombie companies to avoid losses from bad debts. This practice became particularly pronounced in Japan during the so-called "lost decade" of the 1990s, when many banks continued to support companies with no prospects for recovery, thus keeping zombie companies alive.

Another characteristic of zombie firms is their low productivity, inefficiency, and low or negative return on investment. Such companies often lack the capacity to innovate, adapt to market changes, or increase competitiveness. They are typically burdened with outdated technology, inefficient business processes, and high debt. As a result, they are unable to respond to market challenges or become competitive participants within their industry. Despite having sufficient capital to maintain operations, their ability to generate added value is minimal. Their operating margins are low or even negative, so investors and shareholders do not achieve the expected returns on invested capital. This indicates an unsustainable business model in the long term, which cannot keep up with market competition or invest in development and innovation.

Due to low productivity, zombie companies often reduce the efficiency of the entire economy, as they affect the allocation of resources within it. Zombie companies retain capital and labour that could otherwise be redirected to more profitable and innovative sectors, thus distorting economic dynamics and ultimately reducing overall productivity. Although keeping such companies alive may initially seem like a way to prevent job losses and mitigate economic shocks, the long-term consequences can be severe. Their continued presence in the market can create an illusion of stability by concealing the real structural problems that require more fundamental economic reforms. Zombie companies thus create an additional problem by distorting market conditions. Because these companies survive thanks to external assistance, they keep

the prices of products and services low, making it difficult for healthy companies operating under market conditions to compete. This unfair competition can lead to the failure of other companies that, although efficient, are unable to cope with artificially low prices. In the long run, this can reduce overall competitiveness in the industry and prevent the growth of more productive entities. From a strategic management perspective, zombie companies have poor internal processes, making them vulnerable to changes in the external environment. They are often unable to adapt their operations to new technological trends or market changes. This problem is further exacerbated when zombie companies use outdated technologies, resulting in stagnation and an inability to innovate, which further slows the overall development of the industry in which they operate.

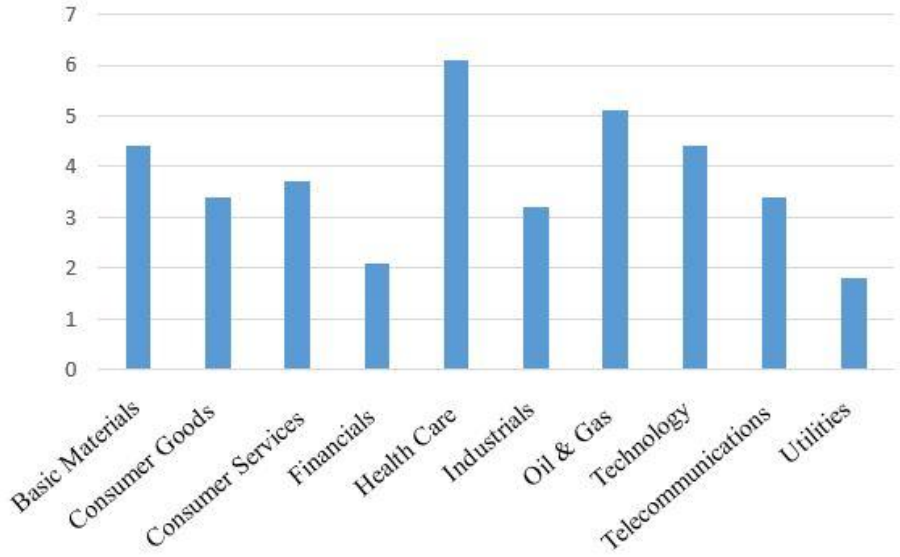


Fig. 1. Share of zombie firms by industry classification (%)

Source: Allianz Global Investors, 2023 [3]

Zombie firms are more common in industrial sectors than in service industries, except for health care. These industries tend to be more financially vulnerable, less productive, and often experience more credit-driven booms while facing limited opportunities for growth.

Zombie firms also make monetary policy and economic planning more challenging. Their presence makes it harder for central banks and governments to assess the actual need for stimulus or restrictive measures. For example, in times of crisis when support is required for businesses, zombie companies may use this assistance without generating long-term benefits for the economy.

Finally, one of the characteristics is their persistence during and after economic crises. During the global financial crisis of 2007-2008, many companies became zombie companies due to significant financial losses and their inability to return to market viability without government support. Similarly, Japan's "lost decade" of the

1990s saw an increase in the number of zombie companies, as banks continued to provide loans even though these companies were unable to repay their debts [4].

Figure 2 shows the evolution of the share of zombie firms over time.

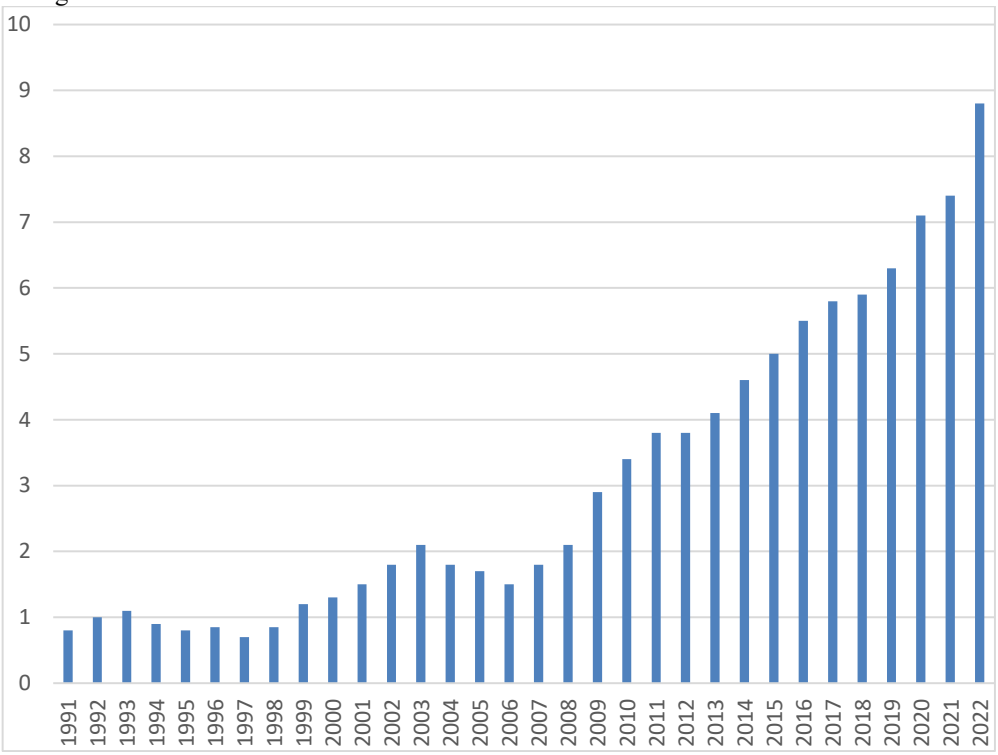


Fig. 2. Share of zombie firms over time (%)

Source: Allianz Global Investors, 2023 [3]

The share of zombie firms has increased significantly since the late 1990s. Addressing the problem of zombie firms and the zombie economy requires a comprehensive macroeconomic and financial approach, including legal system reforms, responsible credit policy, restructuring of state subsidies, and encouragement of innovation. It is important to ensure that resources are allocated to viable and productive enterprises, while zombie companies undergo bankruptcy or restructuring, leading to their recovery or closure.

The proportion of zombie companies varies considerably across countries. Figure 3 shows the presence of zombie firms by country and visualises the share of zombie firms worldwide, while Figure 4 presents zombie firms in Europe.

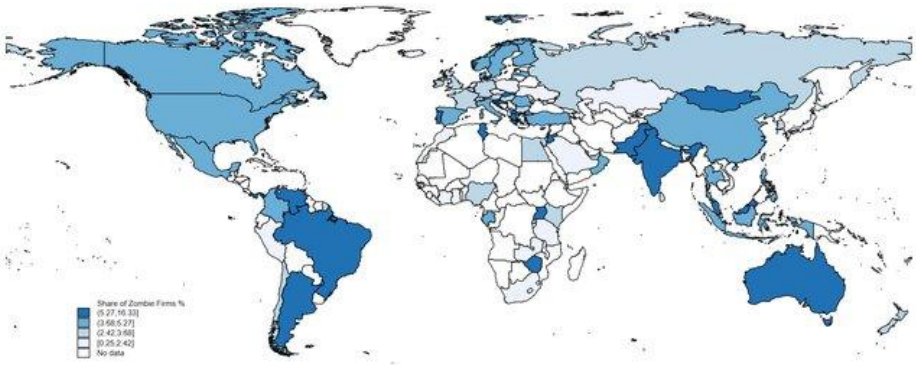


Fig. 3. Share of zombie firms globally (%)

Source: De Martiis, Heil, Peter, 2021 [5]

Zombie firms are those with an interest coverage ratio below one for at least three consecutive years and an age of at least 10 years. The countries with the highest share of zombie firms, shown in dark blue, are Portugal, Greece, Cyprus, Croatia, Macedonia, Slovenia and Slovakia in Europe; Venezuela, Brazil and Argentina in Latin America; Jordan, Pakistan, India, Mongolia, Malaysia and Australia in Asia; and Tunisia, Uganda and Zimbabwe in Africa.

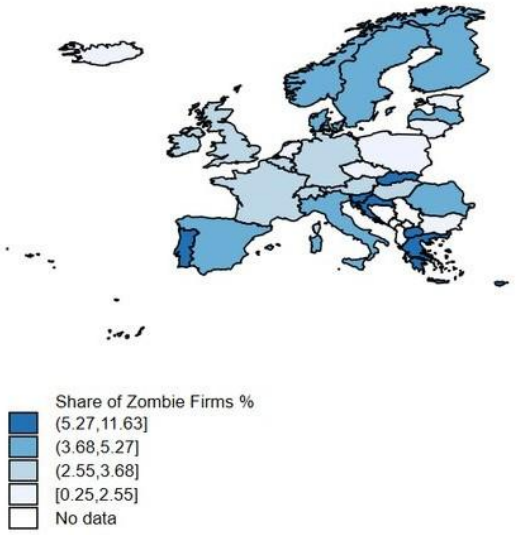


Fig. 4. Share of zombie firms in Europe (%)

Source: De Martiis, Heil, Peter, 2021

The countries for which there are no data are Albania, Serbia, Montenegro, Kosovo, and Bosnia and Herzegovina, shown in white. It can be concluded that zombie firms are more common in countries with weaker economic growth and lower interest rates.

2.3 The emergence of the zombie economy during Japan's "lost decade" of the 1990s

Until its peak, Japan had an exceptionally strong and rapidly growing economy. Japanese industry was among the most advanced in the world, particularly in the automotive and electronics sectors. During this period, when real estate and stock prices were extremely high, people believed their values would only increase and invested ever more in them. After peaking in 1989, Japanese stocks lost up to 60% of their value over the following three years. Commercial real estate also recorded a significant decline, with prices falling by about 50% after their 1992 peak. However, all loans secured by these assets became less valuable than expected at the time. As a result, the Japanese banking system failed to record a net operating profit for more than ten years. The political and legislative response to the financial crisis was to ignore the problem and delay serious economic reforms or the financial consolidation of banks. There has been much debate in society about the main factor contributing to Japan's slow economic growth over the past two decades – a period known as the "lost decade", referring to the prolonged economic stagnation that began in the early 1990s. The biggest problem was created by banks, which had to comply with international capital standards prescribing minimum levels of capital. To meet these requirements, many banks continued to lend to illiquid borrowers. In late 1991, after the collapse of the Japanese bubble and asset prices, Japanese banks continued to support weak and reckless companies rather than allow them to declare bankruptcy. Hoping that these companies would recover or be bailed out by the government, Japanese economy experienced prolonged recessions during the 1990s and 2000s. Until the early 2000s, lending to zombie firms undermined market discipline, resulting in economic stagnation. The recession in Japan was not catastrophic, as most zombie firms eventually recovered during the 2000s [6]. This issue attracted particular attention in Japan because of the so-called zombie firms, i.e. financially weak companies financed at low interest rates.

2.4 The emergence of the zombie economy during the global financial crisis of 2007-2008

The global financial crisis of 2007-2008, also known as the “Great Recession”, was the most severe economic crisis in modern times. It began with problems in the United States housing market and quickly spread worldwide. The financial and economic collapse left a deep mark on consumers and businesses, who faced losses and uncertainty globally, especially in the United States, where millions of people were affected. The crisis was rooted in the collapse of the US housing market, triggered by the massive issuance of high-risk mortgage (subprime) to unqualified borrowers. Falling real estate prices and widespread inability to repay led to mass foreclosures, a stock market crash, and a global financial crisis. As financial institutions began to fail, many were forced to merge with larger corporations, while the US government intervened to offer aid and prevent further collapse, keeping many other institutions afloat. The state began to rescue banks, so-called zombie banks, and other insolvent financial institutions so they could continue operating. The housing market was deeply affected, with evictions and foreclosures occurring, and the stock

market began to fall uncontrollably. Large companies around the world began to fail, losing millions. This culminated in widespread layoffs and prolonged periods of unemployment globally. The decreasing availability of credit and weakening confidence in financial stability led to reduced investment, ultimately resulting in a worldwide recession.

2.5 Zombie firms as a structural constraint on economic growth

Zombie firms present a significant challenge to markets by disrupting natural economic mechanisms such as "creative destruction", in which unsuccessful companies should exit the market to make way for new, innovative entities, thereby enhancing overall productivity. A key characteristic of zombie companies is their inability to generate profits despite low interest rates and favourable financial conditions. Their prolonged survival can lead to reduced innovation and investment, as these companies retain resources that would otherwise be allocated to more productive entities [7]. Additionally, zombie companies often cause stagnation in certain sectors, reducing competitiveness and hindering faster market transformation. Zombie companies remain in the market for years due to government subsidies and support. Despite ongoing financial crises and the unsustainability of their business models, zombie companies survive only because of continual government intervention. This practice is harmful in the long term, as funds invested in unsustainable companies limit opportunities to finance new and innovative businesses. Furthermore, zombie companies retain a workforce that could be redirected to more productive and sustainable enterprises. Although job retention is a key reason for the survival of such companies, long-term unsustainability can lead to even greater problems, including increased national debt and the weakening of the industrial sector. It should be noted that zombie firms exist in almost all developed economies, raising concerns about their impact on growth and innovation. Zombie companies weaken overall economic activity by locking up resources and preventing the growth of more productive firms. After the global financial crisis of 2007-2008, the phenomenon of zombie companies became particularly pronounced in the United States. Many firms, including those in the energy and retail sectors, faced a sharp decline in revenue and a rise in debt. Despite this, many continued to operate due to low interest rates and easy access to credit. A key example is the oil and gas industry, where companies with high debt but low profitability continued to operate thanks to favourable capital market conditions. Although many of these companies remained in operation, they were unable to provide long-term growth or returns on investment, creating an unfavourable situation for the entire industry. Eventually, many were forced to declare bankruptcy or implement comprehensive restructuring to reduce debt and improve operational performance.

3 Empirical studies of zombie firms

The first empirical studies focusing on the phenomenon of zombie firms emerged in Japan at the beginning of the 21st century in response to the economic stagnation in Japan during the 1990s and problems in Japan’s financial system [8]. Since then,

numerous studies have examined zombie firms and the zombie economy in general. Specifically, comparisons across studies by country, sector, age, and size are as follows.

Hallak, Harasztosi and Schich [9] analyse zombie firms across 19 European countries using firm-level data for more than one million companies for 2010 and 2013, indicating that zombie firm shares were higher in Greece and Spain but lower in the Czech Republic and Slovakia, and that larger and older firms, compared to relatively smaller and younger firms, were more likely to be zombie firms.

Gouveia and Osterhold [10] analyse a comprehensive set of firm-level data covering all Portuguese firms for the period 2006 to 2015. They point out that Portugal is one of the OECD countries with the largest drop in exit and restructuring barriers in recent years. As expected from the results in the literature, the high prevalence of zombie firms was associated with significantly lower productivity compared to their healthy counterparts, thereby dragging down aggregate productivity and dampening investment, employment growth, and capital reallocation.

The same authors [11] explore the role of zombies in firm dynamics and the misallocation of labour and capital. Results show that reducing exit and restructuring barriers promotes a more effective exit channel, disproportionately encouraging the exit of non-viable firms and stimulating the restructuring of the most productive zombies.

Another study from Portugal examines the impact of zombie firms on investment, employment growth, and labour productivity in healthy companies within the wholesale and retail trade and hospitality sectors during 2011–2018. The results indicate that the prevalence of zombie companies is higher during economic crises and that most are old and large companies, particularly in the hotel industry. In addition, zombie firms negatively affect how healthy companies are run, because they reduce the latter's investment expenditure and workforce productivity [12].

Cross-country analysis in OECD countries shows that, within industries over the period 2003–2013, a higher share of industry capital sunk in zombie firms is associated with lower investment and employment growth of the typical non-zombie firm, as well as less productivity-enhancing capital reallocation [13].

Considering advanced economies and emerging markets over the past two decades, results show that the share of unproductive and unviable zombie firms has been rising worldwide, especially since the global financial crises and the Covid-19 pandemic. In addition, results reveal that, perhaps surprisingly, the incidence of zombification is lower among private firms, most likely due to the lower survival rates of private firms compared to listed firms. Private zombie firms also tend to exit the market more quickly than listed zombies [14].

A cross-methodology analysis was conducted on a common dataset for 19 euro area countries, i.e. firms from 2004 to 2019, to examine the shares and trends of zombie firms in the economy, comparing the five main methodologies [15].

The relevance and significance of the zombie firms literature is further demonstrated by the fact that this topic was submitted for the degree of Doctor of Philosophy at the Australian National University [16]. The thesis was divided into several chapters: the dynamics of zombie firms in China were examined; the link

between financial conditions and capital invested in zombie firms across 20 OECD countries and 11 industries was explored; the consequences of zombie businesses in Australia before the COVID-19 pandemic were analysed; and the impact of COVID-19 business support measures and mobility restrictions on zombie businesses in Australia from 2020 to 2022 was evaluated.

Blažková and Chmelíková [1] examine the development of zombie firms before and after the 2008 crisis in two different economies, Germany and the Czech Republic, to provide insight into how varying conditions and the overall economic context influence the likelihood of companies becoming zombie firms. The German dataset comprised 1,444,698 observations from 338,923 firms, while the Czech dataset included 2,139,462 observations from 523,542 firms, both covering the years 2000–2016, thus spanning the period before and after the 2008 crisis. The findings confirm that Germany has a more stable economy and a significantly lower risk of zombie firm persistence, while the Czech Republic is at the European average in terms of the proportion of zombie firms.

Using firm-level data on listed companies in 14 advanced economies, Banerjee and Hofmann [17] reveal a ratcheting-up in the prevalence of zombie firms since the late 1980s. The analysis suggests that this increase is linked to reduced financial pressure, which in turn appears to reflect, in part, the effects of lower interest rates. Zombie firms weigh on economic performance because they are less productive and because their presence lowers investment in, and employment at, more productive firms.

The study in Japan [18], which is, to the best of our knowledge, one of the most cited, confirms that zombie-dominated industries experience more depressed job creation and destruction, as well as lower industry productivity. It is noted that the increase in zombies depressed investment and employment growth among non-zombies and widened the productivity gap between zombies and non-zombies.

Another study of zombie firms among small and medium-sized enterprises (SMEs) in Japan [19] finds that many zombie firms exist among SMEs, and that the zombie firm ratio increases as firm size decreases. Some zombie firms eventually recover or emerge from zombie status among SMEs. It is stated that zombie firms are likely problematic from the perspective of the efficiency of the industries to which they belong.

Aside from cross-country comparisons of zombie firms, the literature identifies other research areas within the zombie economy. By combining bibliometric analysis with qualitative content analysis, Antić and Ichev [20] present a comprehensive overview of the zombie firm literature using a dataset of 160 peer-reviewed articles. Furthermore, they offer a clearer overview of the divergent approaches to defining zombie firms.

Another study (from 2000 to 2022) uses a systematic literature review methodology, in which 76 reviewed papers were published in journals ranked on the Australian Business Deans Council list. The study reviews the literature on zombie firms to provide a holistic view by delineating their formation, impact, widespread nature, prevention, and policy implications [21].

Finally, to analyse the 50-year history of zombie firms in Japan, Nakamura [22] reviews long-term changes in zombie firms in Japan over half a century using listed

firm data within a framework in which the concept of zombie firms includes possible efficient bailouts.

4 Zombie firms in a global context: a case study of Toshiba company

One of the most prominent examples of zombie firms is found in Japan, which experienced a "lost decade" during the 1990s. After the real estate and stock market bubble burst, many banks continued to finance insolvent companies to avoid losses on bad debts. This practice enabled numerous zombie companies to survive, despite being unable to achieve positive business results or repay their debts. Toshiba is an example of such a company, having suffered significant losses following several strategic management errors.

As a leader in the production of laptops, televisions, and other consumer electronics, Toshiba became one of the struggling Japanese companies kept afloat by banks. Operating across a diverse range of sectors, including electronics, construction, and nuclear power, the conglomerate generated relatively weak revenues under pressure from competition and technological changes. This led to a sharp decline in investor confidence and serious financial difficulties. Restructuring through the sale of business units and changes in management enabled the company to avoid bankruptcy and become more competitive in the global market. Toshiba managed to redirect resources to profitable areas, demonstrating how effective insolvency management can help a zombie company recover.

Despite its financial problems, the company remained operational due to government intervention and favourable loans. Although Toshiba eventually restructured and improved its performance, its long-term reliance on external support illustrates the challenges faced by zombie companies. Toxic loans and prolonged financing of zombie companies are widely regarded as the main causes of Japan's slow economic growth during the 1990s and early 2000s. These practices reduced the efficiency of resource allocation and slowed the process of "creative destruction", preventing more productive companies from growing and developing.

5 Strategies for overcoming the problem of the zombie economy

Zombie firms survive in the market solely due to financial assistance in the form of grants, subsidies, or preferential loans with subsidised interest rates. In doing so, they disrupt the natural dynamics of the market by occupying resources that could be used more productively. Therefore, effective strategies are required to prevent the long-term damage caused by zombie companies, i.e. strategies that enable restructuring, resource optimisation, and the encouragement of healthy economic entities. The previously mentioned example of the company Toshiba, which has faced financial difficulties for a long time, demonstrates that restructuring through the sale of parts of

its business and changes in management has enabled the company to avoid failure and become more competitive in the global market.

One key measure for addressing the problem of zombie companies is to improve the legal framework for insolvency and bankruptcy. An efficient bankruptcy system allows for the faster identification of companies unable to survive without financial support. State aid and restructuring programmes may allow a zombie company to continue operating, but without genuine reforms and innovation, it will encounter problems again in the long term. This example highlights the importance of resolving insolvency issues quickly and efficiently to prevent zombie companies from using resources to the detriment of more productive entities.

In addition to legal reforms, control over bank credit policies is crucial. Zombie companies often survive due to bank loans they cannot repay, and banks continue to finance them to avoid losses from bad debts. An effective credit policy not only reduces the number of zombie companies but also directs capital towards more productive firms, thereby stimulating economic growth. Conversely, zombie companies often persist because of subsidies or favourable government loans. Without a concrete and clear plan for restructuring and ceasing to finance unsustainable projects, these companies do not contribute to the economy in the long term. Instead, they create an additional burden on the budget and the financial system. Stricter control over the credit policies of banks and state institutions is needed to prevent financial support for companies that cannot ensure sustainable operations.

A key measure in addressing the problem of zombie companies is to encourage innovation and "creative destruction". This concept, well established in economic theory, refers to the process by which unproductive and unsustainable companies exit the market, making room for innovative and productive entities. However, zombie companies hinder this process by occupying resources and market space that could be used by more competitive firms.

Another strategy for overcoming zombie companies is to gradually close or sell unviable parts of the business. During its crisis, Toshiba sold divisions that had become too risky or unprofitable, thereby providing funds to stabilise other parts of the business. Such decisions allow a company to focus on key profitable sectors, while unprofitable segments are transferred to other, potentially more productive market players. Stimulating innovation can also involve technological reforms within industries where zombie companies predominate. Successful companies often redirect resources to research and development to secure a competitive advantage in the market.

Resolving the problem of zombie companies requires a comprehensive approach that includes legal reforms, responsible credit policy, business restructuring, and encouragement of innovation. It is important to ensure that resources are allocated to viable and more productive companies, while zombie companies either undergo bankruptcy or restructuring leading to their recovery or closure. Countries that have implemented bankruptcy reforms and tightened lending conditions have succeeded in reducing the number of zombie businesses and strengthening economic resilience. Innovation and business restructuring are essential for ensuring long-term growth and competitiveness in the global market.

6 Conclusion

The concept of the zombie economy is a significant phenomenon in contemporary economic research, analysis, and policy. Zombie firms are companies that survive in the market largely due to financial support, despite being economically unviable. They are unable to generate sufficient income to cover their operating costs and service their debts without external assistance such as subsidies or soft loans. These companies persist only because of support from banks, the state, or other financial institutions. Zombie firms significantly distort market dynamics and reduce overall economic productivity.

Understanding, monitoring, and addressing the zombie economy is crucial for sustainable economic growth. Tackling the problem of zombie companies requires a combination of measures, including improving the legal framework for insolvency, implementing responsible credit policies, and encouraging innovation. Competition among companies can drive this process, while available resources should be directed towards more productive sectors of the economy. Although removing zombie companies might seem a logical solution, it is important to prepare for the potential challenges that could result from their mass exit from the market. A sudden wave of business closures could lead to a sharp rise in unemployment and financial instability. Furthermore, sectors that depended on these companies for business relationships and transactions could experience disruptions, further destabilising the economy.

Finally, the research findings aim to contribute to a deeper conceptual understanding of zombie firms as entities that often reduce the efficiency of the entire economy by distorting the allocation of resources.

Acknowledgments.

* This paper has been written within the scope of the project „Steering the Energy Rapids: Balancing Energy Sustainability and Security“ (uniri-iz-25-281), funded by European Union – NextGenerationEU via the Croatian National Recovery and Resilience Plan 2021-2026, in conjunction with the University of Rijeka, Faculty of Economics and Business Programme Financing. The views and opinions expressed are solely those of the authors and do not necessarily reflect the official position of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

** This paper has been written within the scope of the project „Regulation and Sustainability in Finance: Challenges and Opportunities for the Stability of the Financial System“ (uniri-iz-25-33), funded by European Union – NextGenerationEU via the Croatian National Recovery and Resilience Plan 2021-2026, in conjunction with the University of Rijeka, Faculty of Economics and Business Programme Financing. The views and opinions expressed are solely those of the authors and do not necessarily reflect the official position of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

*** The paper was written on the basis of the bachelor thesis “*Uzroci i posljedice nastanka zombi ekonomije (Causes and Consequences of the Emergence of the Zombie Economy)”*,

which was written independently by the student Sara Sadrić under the mentorship of Dario Maradin, PhD, Associate Professor.

**** The authors declare that no artificial intelligence tools were used in the preparation, writing or analysis of this manuscript.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Blažková, I., Chmelíková, G.: Zombie firms during and after crisis. Journal of Risk and Financial Management **15**(7), 301 (2022)
2. Albuquerque, B., Iyer, R.: The rise of the walking dead: Zombie firms around the world. Journal of International Economics **152**, 104019 (2024)
3. Allianz Global Investors: Higher for longer: the demise of the “zombies”?. https://hk.allianzgi.com/en-hk/financial-advisor/insights/insights-and-strategies/202311-zombie-companies, last accessed 2026/03/25 (2023)
4. Papava, V.: How zombie economicus is gradually replacing homo economicus. SSRN Electornic Journal (2023)
5. De Martiis, A., Heil, T.L.A., Peter, F.J.: Are you a zombie? A supervised learning method to classify unviable firms and identify the determinants. https://www.efmaefm.org/0EFMAMEETINGS/EFMA%20ANNUAL%20MEETINGS/2021-Leeds/papers/EFMA%202021_stage-2049_question-Full%20Paper_id-243.pdf, last accessed 2026/05/25 (2021)
6. Fukuda, S., Nakamura, J.: What happened to "zombie" firms in Japan? Reexamination for the lost two decades. Global Journal of Economics **2**(2), (2013)
7. Andrews, D., McGowan, M.A., Millot, V.: Insolvenca regimes, zombie firms and capital reallocation. OECD Library, (2017)
8. Blažková, I., Dvouletý, O.: Zombies: Who are they and how do firms become zombies?. Journal of Small Business Management **60**(1), 119–145 (2022)
9. Hallak, I., Harasztosi, P., Schich, S.: Fear the walking dead? Incidence and effects of zombie firms in Europe. Journal of Economic Science Research **1**(1), (2018)
10. Gouveia, A.F., Osterhold, C.: Fear the walking dead: Zombie firms, spillovers and exit barriers. OECD Productivity Working Papers, No. 13. OECD Publishing, Paris (2018)
11. Osterhold, C., Gouveia, A.F.: Productivity, zombie firms and exit barriers in Portugal. International Productivity Monitor **38**, 29–49 (2020)
12. da Silva, A., Gonçalves, A.: How zombie firms affect healthy firms: The case of Portuguese trade sector. Revista Galega De Economía **31**(3), 1–18 (2022)
13. McGowan, M.A., Andrews, D., Millot, V.: The walking dead? Zombie firms and productivity performance in OECD countries. Economic Policy **33**(96), 685–736 (2018)
14. Albuquerque, B., Iyer, R.: The rise of the walking dead: Zombie firms around the world. IMF Working Papers 125 (2023)
15. Mingarelli, L., Ravanetti, B., Shakir, T., Wendelborn, J.: Dawn of the (half) dead: The twisted world of zombie identification. ECB Working Paper, No. 2743. European Central Bank, Frankfurt a. M. (2022)

16. Bowman, J.A.: The working dead: Dynamics, causes and consequences of zombie businesses. A thesis submitted for the degree of Doctor of Philosophy at the Australian National University (2024)
17. Banerjee, R.N., Hofmann, B.: The rise of zombie firms: causes and consequences. BIS Quarterly Review (2018)
18. Caballero, R.J., Hoshi, T., Kashyap, A.K.: Zombie lending and depressed restructuring in Japan. American Economic Review **98**(5), 1943–1977 (2008)
19. Goto, Y., Wilbur, S.: Unfinished business: Zombie firms among SME in Japan's lost decades. Japan and the World Economy **49**, 105–112 (2019)
20. Antić, R., Ichev, R.: Walking dead in the marketplace: A bibliometric review of the zombie firm literature. Journal of Economic Surveys **40**(3), 1400–1424 (2026)
21. Upadhyay, A.P., Baag, P.K.: Surviving beyond expiry: A systematic literature review of zombie firms. China Accounting and Finance Review **26**(1), 31–54 (2024)
22. Nakamura, J.: A 50-year history of "zombie firms" in Japan: How banks and shareholders have been involved in corporate bailouts?. Japan and the World Economy **66**, 101188 (2023)

THEMATIC SECTION

Transformational Potential of Generative AI

Resilient and Intelligent Organizations in Times of Permanent Disruption

SYMORG 2026 • CONFERENCE PROCEEDINGS

Original Research Article *https://doi.org/10.5281/zenodo.23478889*

Transformative Potential of Artificial Intelligence in Ecological and Ecosystem Management: An Analysis of Successful Sustainable Solutions

Davorin Žnidarič[1,2][0009-0000-8522-5767]
and Marjan Senegačnik[3][0000-0002-0013-2762]

[1] University of Ljubljana- Faculty of Law; [2] Institute for Sustainability Development KUM
[3] University Maribor- Faculty of Organisation science Kranj

Abstract. This paper examines the transformative potential of artificial intelligence (AI) in ecological and ecosystem governance, with a particular focus on environmental justice and sustainable transformation in former industrial regions. The paper develops a conceptual and applied methodological framework for AI-supported ecological decision-making systems in degraded industrial environments. Using a qualitative conceptual-methodological approach based on a review of scientific literature, policy analysis, and the analysis of practical AI applications in environmental governance, the paper explores the opportunities and limitations of AI-supported ecosystem governance. The discussion is theoretically applied to the post-industrial region of Zasavje in Slovenia, historically shaped by coal mining, thermoelectric energy production, heavy industry, and long-term environmental degradation. The findings indicate that AI can significantly improve environmental monitoring, ecological restoration planning, transparency, participatory governance, and environmental justice analysis. However, the paper also identifies important limitations related to data quality, algorithmic bias, governance transparency, energy consumption, and unequal digital capacities. The paper concludes that AI can contribute to more sustainable and socially just ecosystem governance only if it is integrated into transparent, participatory, and environmentally responsible governance systems.

Keywords: artificial intelligence, ecosystem management, ecological management, environmental justice, post-industrial regions, Zasavje.

1 Ecological and Ecosystem Management and Artificial Intelligence

Artificial intelligence (AI) is increasingly influencing environmental governance, ecosystem monitoring, and ecological decision-making. Rapid technological development enables the processing of large environmental datasets, predictive modeling, ecological simulations, and real-time environmental monitoring. Consequently, AI systems are becoming an important component of sustainable environmental governance and ecological planning [1-2].

At the same time, increasing environmental degradation, climate change, biodiversity loss, and socio-economic inequalities require new approaches capable of

integrating ecological, social, technological, and governance dimensions [3-4]. Former industrial regions represent particularly vulnerable environments, as they combine long-term environmental burdens, socio-economic inequalities, degraded landscapes, and structural dependence on fossil fuel-based industries [5-6].

One such example is the Zasavje region in Slovenia, historically shaped by coal mining, thermoelectric energy production, cement production, and heavy industry. Previous analyses have shown that long-term industrial development in Zasavje generated significant forms of environmental and political injustice associated with industrial policies, unequal distribution of environmental burdens, and environmentally unsustainable development models.

This paper examines the potential role of artificial intelligence in ecological and ecosystem governance, with a particular focus on environmental justice and sustainable transformation in former industrial regions.

The main objective of this paper is to analyze the potential role of artificial intelligence in ecological and ecosystem governance and to develop a conceptual framework for AI-supported environmental governance and decision-making in degraded industrial environments.

The central research question of the paper is:

How can artificial intelligence contribute to more effective, transparent, and environmentally just ecological and ecosystem governance in former industrial regions?

The contribution of the paper is threefold:

- **Theoretical contribution:** integration of artificial intelligence, ecosystem governance, environmental justice, and sustainable governance concepts.
- **Methodological contribution:** development of a conceptual framework for AI-supported decision-making in degraded industrial environments.
- **Applied contribution:** theoretical application of the framework to the former industrial region of Zasavje in Slovenia.

The paper represents a conceptual and applied methodological study. It does not present a fully empirical implementation of artificial intelligence but rather develops a theoretically grounded framework and a practical application model for future ecosystem governance systems.

2 Metodology

The paper employs a qualitative conceptual-methodological research design that combines literature review, policy analysis, analysis of practical artificial intelligence applications, and conceptual modeling [7-8].

The study is based on an interpretative and interdisciplinary approach integrating environmental sciences, ecosystem governance, sustainability studies, environmental justice theory, governance studies, and artificial intelligence research.

Scientific literature, institutional reports, and international policy documents were selected according to the following criteria:

- relevance to artificial intelligence and ecosystem governance,
- relevance to environmental governance,
- applicability in degraded industrial environments,
- significance for environmental justice and sustainability.

The analysis draws on scientific databases, peer-reviewed literature, and reports from international institutions such as the European Environment Agency (EEA), the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES), and the Organisation for Economic Co-operation and Development (OECD).

The analytical procedure included:

- comparative review of AI applications in ecological governance,
- identification of ecological and governance functions of AI systems,
- analysis of risks and limitations,
- development of a conceptual model of AI-supported ecosystem governance.

The Zasavje case is used as a theoretical example to illustrate potential scenarios for practical implementation in former industrial regions.

The paper does not present an empirical field implementation or quantitative AI modeling. The proposed framework therefore represents a conceptual and methodological foundation for future applied research, pilot projects, and ecosystem governance experiments.

3 Ecological and ecosystem governance and artificial intelligence

According to Elegbede and Zamboni [9], ecological governance refers to the management of natural resources based on scientific analysis of environmental

conditions and ecological processes. Its objective is the preservation of ecosystem functions, biodiversity, and environmental stability.

Ecosystem governance represents a broader and more integrated approach combining ecological, social, economic, and governance dimensions [10-11]. It emphasizes interdisciplinarity, long-term planning, adaptability, stakeholder participation, and sustainability.

Artificial intelligence increasingly contributes to both approaches through:

- real-time environmental monitoring,
- spatial analysis,
- predictive modeling,
- automated classification systems,
- decision-support systems,
- participatory governance tools.

AI systems can process large datasets derived from satellite imagery, environmental sensors, GIS systems, climate models, and ecological databases significantly faster than traditional analytical methods [2].

The growing integration of artificial intelligence into environmental governance also reflects broader transformations associated with digital environmental governance and smart sustainability systems. AI-supported systems increasingly enable adaptive ecosystem governance based on continuous environmental feedback, predictive simulations, and real-time ecological assessment.

4 The role of ai in integrated ecological and ecosystem governance

Artificial intelligence refers to computer systems capable of performing tasks that typically require human intelligence, including pattern recognition, data interpretation, prediction, and adaptive decision-making [1].

The application of AI in ecological and ecosystem governance can be divided into three major functional areas.

4.1 Spatial and Environmental Analysis

- AI-supported spatial analysis enables:
- interpretation of satellite imagery,
- biodiversity mapping,
- land degradation monitoring,

- deforestation analysis,
- flood and drought prediction,
- identification of pollution hotspots,
- ecosystem monitoring.
- Convolutional neural networks (CNNs), machine learning algorithms, GIS integration systems, and remote sensing technologies enable highly accurate environmental analyses.

Hansen et al. [12], for example, demonstrated that AI-supported forest monitoring systems in the Amazon achieved detection accuracy exceeding 90%.

In the case of Zasavje, spatial analysis is particularly important because it enables precise identification of degraded areas and real-time environmental assessment. Such data provide spatial planners with improved analytical support for land-use planning, environmental remediation, and ecosystem restoration strategies.

4.2 Predictive Ecological Governance

AI systems can simulate ecosystem responses to:

- climate change,
- restoration measures,
- industrial pollution,
- hydrological changes,
- biodiversity decline,
- land-use transformation.

Predictive ecological modeling enables scenario analysis prior to environmental interventions and supports adaptive ecosystem governance strategies [13].

In the post-industrial region of Zasavje, such modeling is particularly important because it enables detailed mapping of degraded areas and assessment of potential impacts on living and non-living environmental systems. Simulations of ecosystem and spatial dynamics may also support the identification of areas suitable for increased food self-sufficiency, which, together with energy resilience, represents an important dimension of sustainable regional transformation [14].

In cases of environmental pollution, AI systems can provide real-time environmental assessment and continuous data validation, thereby improving the reliability and transparency of environmental monitoring systems.

4.3 Environmental Justice and Participatory Governance

Artificial intelligence can contribute to environmental justice through:

- transparent access to environmental data,
- identification of vulnerable communities,
- participatory mapping,
- citizen science platforms,
- open-data governance systems.

AI-supported systems can identify unequal distribution of environmental burdens and socio-economic vulnerability more efficiently than conventional analytical approaches [15-16].

In the Zasavje region, such systems may be connected with degraded land monitoring and spatial environmental assessment. AI systems are particularly important because they enable continuous monitoring of potentially toxic elements (PTEs) in contaminated soils and support long-term environmental remediation strategies.

In addition, AI-supported participatory systems can strengthen public environmental transparency and improve communication between institutions, local communities, and environmental stakeholders [17-18]. They may also provide residents of degraded areas with greater insight into the actual environmental quality of their living environment.

5 Theoretical application: zasavje as a former industrial region

Zasavje represents one of Slovenia's historically most industrialized regions, with development based primarily on:

- coal mining,
- thermoelectric energy production,
- cement manufacturing,
- chemical industry,
- heavy industry.

Long-term industrialization has contributed to:

- air pollution,
- particulate matter emissions,
- heavy metal contamination,
- land degradation,
- ecosystem degradation,

- socio-economic vulnerability.

The narrow valley topography further exacerbates pollutant accumulation and limits natural ventilation.

AI-supported ecological monitoring in Zasavje could integrate:

- air quality stations,
- satellite imagery,
- drone surveillance,
- IoT environmental sensors,
- meteorological data,
- hydrological monitoring,
- GIS-based environmental databases.

Such a system could analyze:

- PM_{10} and $PM_{2.5}$ concentrations,
- SO_2 emissions,
- spatial distribution of pollution,
- soil contamination,
- biodiversity changes,
- effectiveness of ecosystem restoration.

A decision-making workflow supported by AI may involve:

Input Data:

- air pollution data,
- soil contamination data,
- satellite imagery,
- hydrological and meteorological data,
- public health indicators,
- demographic indicators.

AI Techniques:

- machine learning models,
- CNNs,
- GIS-based spatial analysis,
- predictive ecological modeling,
- decision-support algorithms [1-2].

Outputs:

- pollution hotspot identification,
- ecological risk mapping,
- priority restoration areas,
- early warning systems,
- climate vulnerability scenarios,
- environmental justice indicators.

Decision Roles:

- municipal remediation planning,
- environmental policy development,
- spatial planning,
- participatory governance,
- public environmental transparency,
- ecosystem restoration prioritization.

Even with AI capabilities, human oversight remains critical. Input selection, scenario simulation, and ethical decision-making require human judgment, expertise, and moral considerations [2, 19].

6 Decision support system (dss) proposal, ai-supported

The paper proposes a conceptual AI-supported decision support system (DSS) for ecological and ecosystem governance in degraded industrial regions. The framework integrates four dimensions:

6.1 Environmental Dimension

- real-time environmental monitoring,
- pollution analysis,
- biodiversity monitoring,
- ecosystem resilience assessment.

6.2 Social Dimension

- environmental justice indicators,
- public participation platforms,
- citizen science systems,
- vulnerability mapping.

6.3 Governance Dimension

- open-data transparency,
- institutional coordination,
- participatory decision-making,
- policy evaluation systems.

6.4 Technological Dimension

- predictive modeling using AI,
- machine learning algorithms,
- sensor networks,
- GIS integration systems,
- cloud-based environmental databases.

The proposed framework supports holistic ecosystem governance and environmentally just transition by combining environmental monitoring, social participation, institutional coordination, and predictive environmental analysis [4, 19].

7 Challenges and limitations of ai

Despite significant potential, AI systems introduce several risks and limitations.

7.1 Data Quality and Bias

AI systems depend on data quality, accessibility, completeness, and representativeness. Incomplete or biased datasets can produce inaccurate environmental assessments and reinforce existing environmental inequalities [1-2].

Historical environmental monitoring in post-industrial regions often suffered from spatial gaps, fragmented datasets, or institutional inconsistency, potentially skewing analyses. Political ecology perspectives indicate that pollution assessment was historically influenced by power relations among stakeholders [5-6].

7.2 Governance and Transparency

AI governance requires transparent management, institutional accountability, public oversight, explainable algorithms, and participatory mechanisms. Lack of transparency can reproduce political or economic power asymmetries and reinforce environmental injustice [4, 20].

7.3 Energy Consumption and Environmental Costs

Large AI systems consume substantial computational energy. Strubell, Ganesh, and McCallum [21] demonstrated that training large AI models generates significant carbon emissions. Sustainable AI implementation requires renewable energy integration, energy-efficient data centers, optimized models, and environmental assessment of digital infrastructure.

7.4 Digital Inequality

Unequal technological capacities across regions can exacerbate socio-economic disparities. Economically weaker regions often lack digital infrastructure, environmental monitoring systems, AI expertise, and access to advanced analytical technologies [2].

7.5 Human Oversight

Despite growing AI autonomy, humans remain central to development, governance, and application. Humans define objectives, select input data, interpret results, ensure validity, oversee ethical compliance, prevent algorithmic bias, protect privacy, and facilitate participatory decision-making [1, 17, 19].

8 Discussion

Analysis indicates that AI can significantly support environmental and ecosystem governance, especially in degraded industrial contexts. AI systems enhance:

- environmental monitoring,
- predictive ecological modeling,
- participatory governance,
- transparency of environmental information,
- environmental justice assessment.

However, AI should not be regarded merely as a technological solution. Environmental challenges in former industrial regions are closely intertwined with political, economic, social, and governance structures [6, 22].

Effective AI-supported ecological governance requires:

- transparent governance systems,
- interdisciplinary collaboration,
- public participation,
- environmental accountability,

- socially just sustainability policies [15-17].

The conceptual framework demonstrates that AI can facilitate integrated ecosystem management and regional sustainable transformation while improving ecological monitoring and public transparency. The transformative potential of AI depends less on technological sophistication and more on institutional arrangements, transparency, accessibility of environmental information, and the capacity of public institutions to integrate digital systems into socially responsible environmental policies [2, 19].

9 Conclusion

Artificial intelligence is becoming an increasingly integral part of ecological and ecosystem governance. AI-supported systems can substantially improve environmental monitoring, ecological restoration planning, predictive ecosystem analysis, and overall environmental management.

Former industrial regions, such as Zasavje, could benefit from AI-supported environmental management systems integrating ecological monitoring, environmental justice indicators, participatory governance, and predictive modeling [2, 15].

At the same time, the analysis highlights critical limitations associated with:

- data quality,
- algorithmic bias,
- governance transparency,
- digital inequality,
- environmental costs of AI infrastructure [21].

AI can contribute to sustainable and environmentally just transformation only if integrated into transparent, participatory, and ecologically responsible governance systems. Future research should focus on pilot implementations, empirical ecosystem monitoring systems, participatory AI-supported governance platforms, and comparative analyses of post-industrial ecological transition models.

Humans remain central to environmental decision-making, as AI cannot independently evaluate ethical, moral, and societal implications of environmental interventions. Human actors determine developmental priorities, interpret environmental problems, and make trade-offs between economic interests, environmental protection, and social justice. Ethical perspectives, social responsibility, intergenerational justice, and societal readiness for sustainable transformation are essential for shaping development models. Environmental degradation often results from political decisions, economic interests, institutional power relations, and prevailing developmental narratives, rather than a lack of technological solutions. Consequently, AI cannot replace human responsibility in environmental governance;

human oversight is critical for ensuring environmental justice and inclusive decision-making [5, 15, 17, 19].

Disclosure of Interests. The authors have no competing interests.

References

1. Russell, S., Norvig, P.: *Artificial Intelligence: A Modern Approach*, 4th edn. Pearson, London (2021).
2. Vinuesa, R., Azizpour, H., Leite, I., et al.: The role of AI in achieving the Sustainable Development Goals. *Nature Communications* 11, 233 (2020). doi:10.1038/s41467-019-14108-y]
3. IPBES: *Global Assessment Report on Biodiversity and Ecosystem Services*. IPBES, Bonn (2019).
4. OECD: *OECD AI Principles*. OECD Publishing, Paris (2021). doi:10.1787/9789264269530-en
5. Bryant, R., Bailey, S.: *Third World Political Ecology*. Routledge, London (1997).
6. Robbins, P.: *Political Ecology: A Critical Introduction*, 3rd edn. Wiley-Blackwell, Hoboken (2020).
7. Denzin, N.K., Lincoln, Y.S.: *The SAGE Handbook of Qualitative Research*, 4th edn. Sage, Thousand Oaks (2011).
8. Yin, R.K.: *Case Study Research and Applications*, 6th edn. Sage, Thousand Oaks (2018).
9. Elegbede, I., Zamboni, R.: Ecological governance: Concepts and approaches. *Environmental Policy Review* 15(2), 45–63 (2022).
10. Christensen, N.L., Bartuska, A.M., Brown, J.H., et al.: The report of the Ecological Society of America on ecological management. *Ecological Applications* 6(3), 665–691 (1996). doi:10.2307/2269387
11. Folke, C., Carpenter, S., Walker, B., et al.: Regime shifts, resilience, and biodiversity in ecosystem management. *Annual Review of Ecology, Evolution, and Systematics* 35, 557–581 (2005). doi:10.1146/annurev.ecolsys.35.021103.105711
12. Hansen, M., et al.: AI-based forest monitoring and deforestation detection. *Remote Sensing of Environment* 284, 113124 (2023). doi:10.1016/j.rse.2023.113124
13. Chapin, F., Kofinas, G., Folke, C.: *Principles of Ecosystem Stewardship: Resilience-Based Natural Resource Management in a Changing World*. Springer, New York (2010).
14. Žnidarič, D.: Rudarsko industrijska regija Zasavje skozi okoljsko in politično (ne)pravičnost (The mining-industrial region of Zasavje through environmental and political (in)justice – in Slovene). *Teorija in praksa* 62(1), 189–204 (2025).
15. Schlosberg, D.: *Defining Environmental Justice: Theories, Movements, and Nature*. Oxford University Press, Oxford (2007).
16. Walker, G.: *Environmental Justice: Concepts, Evidence and Politics*. Routledge, London (2012).
17. Dryzek, J.S.: *The Politics of the Earth: Environmental Discourses*, 3rd edn. Oxford University Press, Oxford (2013).
18. Fung, A., Wright, E.O.: Deepening democracy: Innovations in empowered participatory governance. *Politics & Society* 31(1), 5–41 (2003). doi:10.1177/0032329202250210
19. Floridi, L., Cowls, J., Beltrametti, M., et al.: AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines* 28, 689–707 (2018). doi:10.1007/s11023-018-9482-5

20. European Commission: *Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts*. COM(2021) 206 final, Brussels (2021).
21. Strubell, E., Ganesh, A., McCallum, A.: Energy and policy considerations for deep learning in NLP. *arXiv preprint arXiv:1906.02243* (2019).
22. Bryman, A.: *Social Research Methods,* 5th edn. Oxford University Press, Oxford (2016).

Original Research Article *https://doi.org/10.5281/zenodo.23478925*

Multi-Agent Object Detection Framework Based on Raspberry Pi YOLO Detector and Slack-Ollama Natural Language Interface

Vladimir Kalušev[1(0009-0005-8851-5790)], Branko Brkljač[2(0000-0001-7932-6676)], and Milan Brkljač[3(0000-0002-0617-973X)]

1 The Institute for Artificial Intelligence Research and Development of Serbia
Fruškogorska 1, 21000 Novi Sad, Republic of Serbia
vladimir.kalusev@ivi.ac.rs
2 Department of Power, Electronic and Telecommunication Engineering
Faculty of Technical Sciences, University of Novi Sad
Trg Dositeja Obradovića 6, 21000 Novi Sad, Republic of Serbia
brkljacb@uns.ac.rs
3 Faculty of Finance, Banking and Auditing, Alfa BK University
Bulevar maršala Tolbuhina 8, 11070 Belgrade, Republic of Serbia
milan.brkljac@alfa.edu.rs

Abstract. The paper presents design and prototype implementation of an edge based object detection system within the new paradigm of AI agents orchestration. It goes beyond traditional design approaches by leveraging on LLM based natural language interface for system control and communication and practically demonstrates integration of all system components into a single resource constrained hardware platform. The method is based on the proposed multi-agent object detection framework which tightly integrates different AI agents within the same task of providing object detection and tracking capabilities. The proposed design principles highlight the fast prototyping approach that is characteristic for transformational potential of generative AI systems, which are applied during both development and implementation stages. Instead of specialized communication and control interface, the system is made by using Slack channel chatbot agent and accompanying Ollama LLM reporting agent, which are both run locally on the same Raspberry Pi platform, alongside the dedicated YOLO based computer vision agent performing real time object detection and tracking. Agent orchestration is implemented through a specially designed event based message exchange subsystem, which represents an alternative to completely autonomous agent orchestration and control characteristic for contemporary LLM based frameworks like the recently proposed OpenClaw. Conducted experimental investigation provides valuable insights into limitations of the low cost testbed platforms in the design of completely centralized multi-agent AI systems. The paper also discusses comparative differences between presented approach and the solution that would require additional cloud based external resources. Since the new AI paradigms are also affecting the current labor market, we are also briefly discussing possible implications of the ongoing changes on the way how the existing products and services are being made.

Keywords: Multi-agent systems · Object detection · Ollama LLM · YOLO · Agent orchestration · Raspberry Pi.

1 Introduction

Recent years have witnessed significant advancements in the fields of computer vision and natural language processing. Proliferation of large language models (LLMs) and AI enabled edge devices have brought novel opportunities for integration of different functionalities and improved system performance. In particular, the traditional product development paradigm based on costly and error prone steps with slow incremental improvements is consistently being replaced by fast prototyping characterized by multiple trial and error development cycles and significantly lower cost per iteration. In such fast pace market environment the key to success is often based on efficient exploitation of novel capabilities brought by generative AI systems and their transformational potential [1]. What were previously regarded as separate tasks and required dedicated teams and resources can nowadays be accomplished with fraction of previous costs by specialized software development frameworks and generic low-cost hardware implementation platforms. In the context of such trends, which are contributing to higher proliferation of AI technologies and fostering creativity and innovation in varying application scenarios across different industries, the paper investigates transformational potential of Generative AI in development and operation of multi-agent vision based systems. The study showcases design and implementation of an edge based prototype with real-time processing capabilities and high level integration of several subsystems implementing different AI agent functionalities. The main goal of the study was to demonstrate that system integration of edge based components capable of performing complex vision tasks can be performed without costly development of application specific control and communication subsystems [2], but instead based on contemporary multi-agent AI paradigm and utilization of natural language interface that will play the same role. The hypothesis was that such integration could be performed in fully centralized way, in which all AI agents and their orchestration would be run on standardized, commercially available low-cost hardware, without the need for additional external resources. In order to demonstrate such generic design goals we have selected object detection [3] and tracking [4], as some of the most common vision tasks, and Raspberry Pi [5] as an example of ubiquitously available AI enabled edge platform on which such approach could be fully tested. Developed fast prototyping hardware testbed with custom camera integration and fan equipped device casing is illustrated in Fig. 1.

The proposed multi-agent object detection framework explores design approach in which AI agents are actively collaborating and competing for shared resources on the same platform, but are not fully autonomous nor autonomously built from scratch, which would be characteristic for some of the most recent and alternative design approaches. Namely, in comparison to traditional AI agent orchestration in which their deployment, collaboration and communication

are based on adaptive but in essence still, to the large extent, deterministic programs, some of the most recent trends consider the AI agents orchestration from the more open-ended and non-deterministic design perspectives. The best example is the recently announced OpenClaw [6] orchestration framework, in which (besides the orchestration task) automated LLM inference and host system resources can be also used to build a platform specific multi-agent AI systems from scratch. Main characteristics of such non-traditional design approaches is that the complete system design and agent architectures can be described and implemented by using only natural language, as well as their orchestration.

(a) (b)

Fig. 1. Developed fast prototyping hardware testbed: (a) custom Raspberry Pi camera setup, (b) MIPI CSI camera board with Sony IMX219 color sensor.

Motivated by the recent success of such fully autonomous and high level AI automation tools, we have decided to explore an alternative approach that could be feasible to implement on resource constrained hardware platforms in the described fully centralized way. Although some of the first considerations also included system design based on the described OpenClaw paradigm, after initial tests it was concluded that such approach could not be feasible due to limited performance of locally run LLM and high demand for LLM based inference at both development and deployment phases of such multi-agent systems. Instead, it was decided to pursue a hybrid solution that would incorporate the best of the both worlds, natural language interface for smooth and efficient communication with multi-agent system, and an event based message exchange subsystem for task specific orchestration of locally run AI agents. In the case of practical applications that require complex system components that can allocate a significant part of available development budget and resources, e.g. development of specific vision based analytics subsystems, the advantages of utilizing LLM based natural language interface for system control and communication are that the costs of wider system integration of such subsystems with costly

development can be significantly lowered by leveraging on the existing generative AI technologies.

Thus, the presented object detection framework is proposed as the generic template for fast prototyping of similar fully centralized multi-agent AI systems on resource constrained platforms, enabling complete multi-agent system integration and testing in short development cycles. We note that the use of generative AI can be present at both development and operational level of such systems. In the proposed framework we assume that generative AI in the form of locally run LLM is utilized in automating system operation, while the automated generation of system software components is not pursued due to high computational cost of locally run LLM inference. However, this does not exclude alternative development scenarios in which the development process would be performed externally and independently from the selected deployment platform, e.g. by utilizing externally run LLM inference or complete platform emulation, allowing for full development automation even in the case of the proposed hybrid design solution presented in this study.

The rest of the paper is organized as follows. In Section 2 we provide an overview of the related work, with particular emphasis on computer vision tasks of object detection and tracking, natural language interfaces based on open source LLMs, automated chatbots, and multi-agent AI systems. In Section 3 we describe the design of the proposed multi-agent object detection framework and practical implementation details of the developed AI agents, which are run locally and performing vision, reporting, and communication and control tasks on the single Raspberry Pi device. Results of real-time system operation consisting of YOLO type object detection, passive tracking and interaction with the Slack [7] messaging application chatbot running on the locally deployed Ollama [8] LLM are presented in Section 4. It also includes additional discussion of the low-cost multi-agentic framework design, its comparison to the OpenClaw design principles, and possible market implications of the generative AI based fast prototyping. Finally, in Section 5 we refer to main conclusions and possible future research directions.

2 Related Work

Integration of different AI subsystems into single embedded computing platform has gained significant momentum. From home assistants, smart buildings, conversation systems, home automation, to surveillance systems, traffic monitoring, process control, autonomous driving, medicine, or the augmented reality and the gaming industry, the need for combining different aspects of AI is becoming ever more important across range of different industries. Reasons lie in the need for achieving better functionality of the corresponding products and services, but also in providing higher flexibility in system development and deployment, with both aspects resulting in easier adaptation of generic AI solutions to wider range of the user specific application scenarios. Thus, on the one hand there is a significant progress at the level of hardware miniaturization

and its software abstraction, which is characteristic for numerous currently available commercial development boards or their open source counterparts. While, on the other hand, the ongoing algorithmic improvements in the fields of computer vision (CV) and natural language processing (NLP) are opening numerous opportunities for ideation and prototyping of innovative technical systems that are exploiting high level inference capabilities of complex AI models [9]. In such environment the next logical step is a convergence of these technologies towards a unified and diverse technical ecosystem in which AI systems will be easier to bring to the market and test under real operational conditions and user acceptance levels. In this process the role of natural language interfaces will seem to be twofold. Primarily they will provide natural and easier communication with human users, without the need for specialized training or technical knowledge, lowering the barrier for complex technology acceptance. However, natural language interfaces will be also the best bridge towards automated development of agentic AI systems, as a universal language for providing necessary instructions and constraints for the software development process ideation, as well its automation at various stages of continuous integration, delivery and deployment (CI/CD). The second claim is based on the already existing trend of increased usage of generative AI tools in the software engineering industry, which are reducing the development costs and increasing work productivity of the employees. In that sense, natural language interfaces can be at the core of development process, or have a more traditional role in providing easier command and control of other AI subsystems. In this work, we are exploring the latter one, by demonstrating design of an agentic AI system in which Ollama LLM is combined with Slack channel chatbot application in order to provide functional communication and control in fast prototyping of complex vision system on an embedded device.

Object detection is one of the most explored high level vision tasks with strong research track record. Despite its maturity there are still various types of challenges like occlusions, clutter, adverse operating conditions or complex and dynamic scenes, affecting the overall speed and accuracy of the currently available algorithms. Thus, there are various design improvements that are constantly being made in order to improve algorithm performance and robustness. Notable examples include combining multiple sources of information, addressing multi-scale nature of scene objects at both image and feature levels, applying highly complex hierarchical feature extraction mechanisms based on deep neural networks [10], adaptive attention mechanisms at various levels of information extraction process, or optimizing the postprocessing steps and final classification and regression operations, which are key for successful object detection and its precise localization.

Historically, the object detection paradigm has gone through the gradual transformation from sliding window approaches and the two stage detectors [11, 12] towards the end-to-end predictions of object location and type [13]. In that process the recent change of focus from convolutional neural networks (CNNs) towards attention based mechanisms and transformer based architectures that are

capable of capturing the global context of the object detection task have contributed to high variety of possible architectures and design choices. Although transformer based architectures like DETR [14] have been proven to be on pair with computationally efficient YOLO architectures, the wider engineering community still seems to be preferring the well established CNN architectures and their improved variants. From production side point of view, specialized software development frameworks offering high quality of the shelf implementations of the single stage detectors from the YOLO family [13] have been widely accepted by the software development community due to their constant maintenance, benchmarking and possibly easier porting to specific hardware platforms, which are usually from the start designed to fully support their software stack. As the main algorithmic solution for the demonstration of the proposed multi-agent object detection framework we have decided to rely on the YOLO implementations provided by the REF. and specifically refer to some of the less demanding variants like Yolov8n. We note that there could also be numerous other solutions, depending on available hardware platform and design requirements. However, the selected choice was made in particular to demonstrate the fast prototyping on the available Raspberry Pi platform, which has evolved from general purpose ARM based system on chip (SoC) to high quality AI enabled edge computing device. When it comes to LLMs as one of the key enabling technologies for the design of natural language interfaces, we have relied on Ollama open-source tool for running LLMs locally on the selected Raspberry Pi device. Multi-agent systems can be based on diverse system architectures and agent types, but their design is usually driven by the desired system operation and infrastructure requirements. Thus, we have followed and adopted an approach in which the proposed solution should be easy to adapt to different resource constrained platforms and support diverse agent configurations and their interactions in order to solve complex tasks.

3 Design and Development

Object detection [15] and multi-agent systems are well known concepts that have been systematically explored in different contexts. The key research question of this study was to investigate feasibility of implementing a combination of these complex functional concepts on the edge device like Raspberry Pi platform. It was done in order to test the hypothesis that all AI agents and their orchestration can be run locally on the low-cost hardware, without the need for additional external resources like LLM APIs or other processing off-loading solutions. In addition, in order to demonstrate the current trend of exploiting generative AI tools for cutting costs and improving productivity in the software engineering industry, i.e. providing fast prototyping capabilities, we have decided to demonstrate the use of such tools in development of high fidelity natural language interface, without the need for full development of specialized communication and control for designed multi-agent

system. Namely, an additional advantage of using user familiar environment like the Slack channel chatbot for receiving system generated messages and reports, or giving commands to control agent and system orchestrator is that the potential users or testers of the deployed system do not need to be technically knowledgeable, nor go through time consuming training for system operation. Instead, the feedback can be received almost instantly on a relatively large sample of potential system testers. Development and implementation of the showcased prototype have also been done as a design template for similar application scenarios or multi-agent systems that would run under same conditions on the low-cost hardware. Thus the focus of the proposed object detection framework was on the practical realization of the multi-agent system and investigation of its limitations.1

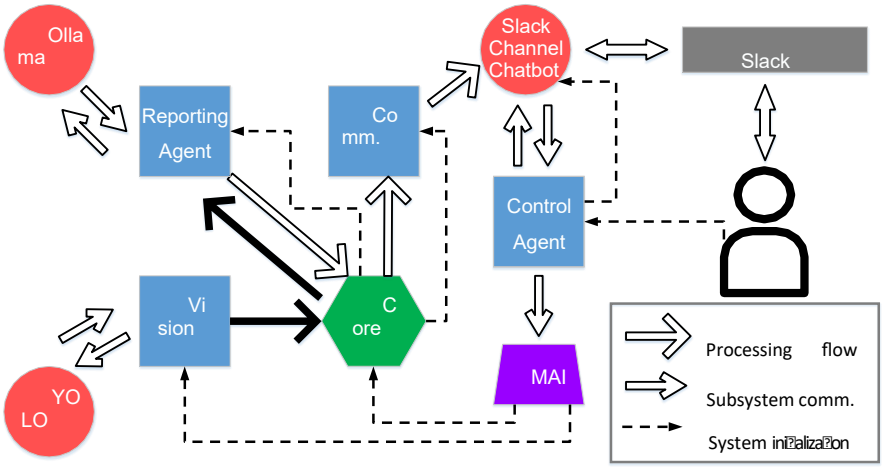


Fig. 2. Multi-agent system architecture: functional elements of the proposed object detection framework and implemented agentic components.

3.1 Multi-Agent System Architecture

The proposed system consists of several agentic components that are loaded into memory and continuously providing different types of functionalities. Their orchestration is based on an event driven pipeline, which goes through specifically designed message router or core message passing interface. The router is facilitating different event callbacks and coordinating execution of designed agentic components. In essence, it is traditional event driven framework, but with the exception of being run on the low-cost hardware and

1Implementation code will be made available upon manuscript publication.

with AI components that in general case can have non-deterministic behavior. Since it was designed for the relatively low resource embedded platform, agent orchestration was not done by a specialized AI agent or AI based middleware, but instead by the developed low complexity router that accepts predefined set of agents and agent generated events. Thus, instead of running an AI based orchestrator, e.g. like OpenClaw, system resources have been mostly utilized for specific AI subsystems of designed agentic components. This is illustrated in Fig. 2, where the relationships between vision, reporting, communication and control agents and their orchestration through the designed core router are depicted, alongside their AI subsystems and user interaction.

It can be seen that the main information or processing flow during system operation goes from the vision agent, as the source of environment perception based on the real-time camera feed analysis. Vision agent asynchronously propagates event messages via system core or router component towards LLM equipped reporting agent, which compiles event triggered natural language reports. Reports are then sent again via router towards communication agent, which is waiting for reporting events and operates preloaded chatbot. After receiving a message, it forwards information in the form of text and images to the corresponding chat channel, which can be also accessed by the subscribed user. Thus, the user receives information through the familiar mobile phone or web client interface, and if necessary sends commands to the control agent that is subscribed to the same chat channel. It is also the main way of system initialization, which is performed by sending a message to the preloaded control agent, which then starts or stops the remaining system components. This process is also illustrated in Fig. 2 by the corresponding dashed arrows, while the long arrows depict the described event triggered information flow during system operation. Short arrows are representing local communication between agents and their AI enabled subsystems.

Multi-agent setup makes clear distinction between responsibilities and scopes of different modules or subsystems, making it highly modular and easy to upgrade and maintain on the long run. It also allows flexibility and modularity in system design by supporting additional agents integration or system independent bridges towards external services or systems that would rely on the information generated by the multi-agent system. In the presented object detection use case these additional services could be email reporting, database filling, processing of multiple camera feeds, GPS locations, map browsing, etc. It is also highly scalable, in the sense that the choice of generic and ubiquitously available hardware could guarantee sustained production of system according to market needs. On the technical side, adding more resources to the selected platform [16] would directly scale with the available agentic capabilities. Real-time processing and coordination of multiple components, i.e. successful integration of heterogeneous AI models from CV and NLP domains makes the proposed system architecture also widely applicable to different tasks, beyond demonstrated object detection.

3.2 Resource Constrained Implementation

Described engineering task is constrained by the relatively low memory and computational resources of the selected Raspberry Pi System on Chip (SoC) platform, Fig. 1, as well as by the imposed requirements of having multiple AI components running at the same time on the local machine. Specific hardware characteristics [5] are reviewed in Table 1 and indicate parallel processing capabilities, but without neural processing unit (NPU) or other neural networks hardware acceleration capabilities. Nevertheless, reported peak performance of 9.96 Gflops [16] combined with hardware accelerated H.264 video encoding makes the platform suitable for adequately designed edge AI workloads.

Table 1. Hardware testbed specifications.

Component	**Details, Raspberry Pi 4 model B**
Processor	Broadcom BCM2711, Quad core Cortex-A72 (ARM v8) 64-bit SoC @ 1.8GHz
Memory	8 GB LPDDR4-3200 SDRAM
Wireless	2.4 GHz and 5.0 GHz IEEE 802.11ac wireless, Bluetooth 5.0
Ethernet	Gigabit Ethernet
USB ports	2 USB 3.0 ports; 2 USB 2.0 ports
GPIO	Raspberry Pi standard 40 pin GPIO header
Display	2 × micro-HDMI ports (up to 4Kp60)
Camera/Display ports	2-lane MIPI DSI display port; 2-lane MIPI CSI camera port
Audio/Video	4-pole stereo audio and composite video port
Hardware codec	H264 (1080p30 encoder), H.264 and H.265 decoder
Graphics	OpenGL ES 3.1, Vulkan 1.0
Storage	Micro-SD card slot for OS and data storage
Power	5V DC via USB-C or GPIO (min 3A); PoE (separate HAT)
Temperature	0–50°C ambient

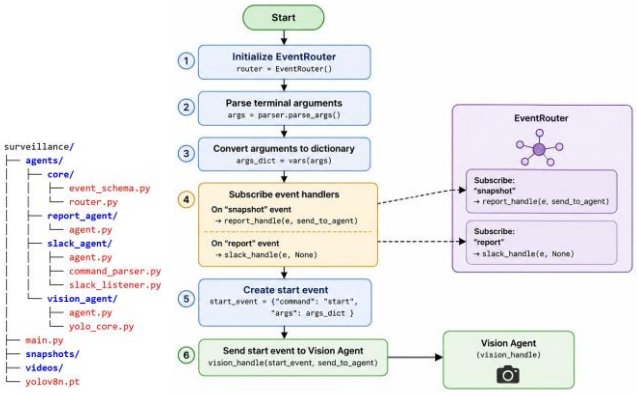
As already mentioned in the introduction, generic design goal was to perform AI subsystems integration in a fully centralized way without the need for additional external resources for their operation. Therefore, software implementation of the architecture depicted in Fig. 2 followed a minimal viable or lean architecture approach, by implementing only the main agentic components and subsystems needed for the proposed object detection

framework. It avoids building dedicated web dashboards or communication protocols, but instead focuses on repurposing of the existing infrastructure like Slack's messaging for both commands and status updates based on natural language interface. This provides more resources for computationally heavy tasks like real-time object detection and tracking, and LLM based reporting. However, it also preserves necessary functionality for fast prototyping and testing of developed solution by leveraging on Slack communication channel authentication, persistence, multiple user notifications or nested conversation chains.

Table 2. Camera board specifications.

Specification	**Details, Raspberry Pi module v2**
Resolution	8 megapixels
Max Video	1080p47
Still Modes	1640 × 1232p41; 640 × 480p206
Sensor	Sony IMX219
Sensor Resolution	3280 × 2464 pixels

From Fig. 1b it can be seen that the utilized camera board with IMX219 sensor, Table 2, is connected through MIPI serial interface, while the CPU cooling of the prototype is enhanced by an additional fan adapter [17], Fig. 1a. Developed testbed is running Debian Linux distribution [18] and is also equipped with HDMI to USB video capture card for recording of system operation.



(a) (b)

Fig. 3. Software components implementing proposed multi-agent architecture in Fig. 2: (a) modules structure, (b) code snippet from "main.py" illustrating system initialization phase with reporting and communication agent subscription to agent specific router events.

Structure of developed software components implementing proposed multi-agent architecture in Fig. 2 is shown in Fig. 3a, where the corresponding folders or modules are depicted. It is a reminiscent of initial decision to implement communication and control agents as single software component named "slack_agent", however the final solution, Fig. 2 clearly differentiates between them.

Fig. 3b illustrates the system initialization process, initiated by the user message via Slack channel to the control agent, Fig. 2. It is implemented by the preloaded system module named "slack_listener", which is responsible for launching the main initialization script. In this process, reporting and communication agent are subscribing to event driven router messages, Fig. 3a, which alongside the launching of the vision agent with parsed command arguments constitutes the main part of the system initialization process.

3.3 Vision Agent

Vision agent implements multi-class object detector and accompanying object tracking, which makes it a versatile tool for different application scenarios. Its main characteristic is small memory footprint (in comparison to other system elements, like LLM) and near to real-time operation. As the model of interest for fast prototyping of described multi-agent architecture were chosen single stage detectors from the YOLO family and their highly optimized implementations provided by [19]. Simultaneous tracking of detected objects was added to enhance system capabilities in possible surveillance and monitoring applications, by providing identification of the same object or higher level inference over multiple frames of the input camera feed. In order to lower the computational cost of the agent operation, tracking was implemented in a passive manner, by utilizing corresponding search algorithm and simple matching criteria based on the intersection between detections' bounding boxes, Algorithm 1. Predefined setup for vision agent's events generation in analyzed experiments was that the object form the chosen category has been detected and that the same object was present in the scene for the predefined period after initial detection. This trigger corresponds to possible surveillance scenarios in which the system is alarmed after the same object stays for too long in the same zone of the scene. Obviously, there are numerous other event scenarios, but the described simple choice was made primarily to provide an efficient event generation for the analysis of the proposed multi-agent architecture performance on the selected low-resource platform.

3.4 Reporting Agent

Events generated by the vision agent are resolved by the reporting agent, Fig. 2, which generates event related reports, but by doing so employs a LLM based subsystem. This process is performed asynchronously, as depicted in Algorithm 2, where it is shown that the system residing LLM is invoked in a separate

thread through the localhost POST HTTP request, which does not require acknowledgment in the sense of a formal handshake protocol. Report generation time and the report quality depend on the specific model size and performance, respectively. However, the reporting results are also influenced by the prompt formulation provided by the format arguments function of the Algorithm 2.

3.5 Communication and Control Agents

As illustrated in Fig. 3a, communication and control agents have been implemented inside the same software module, the so called "slack_agent", which incorporates the Slack channel listener (control agent) and communication agent resolving the raised report generation events. Since the used communication infrastructure for the proposed natural language interface assumed the Slack messaging application, in Fig. 5 are provided some more details regarding its configuration and operation modes. We note that the choice of messaging application is completely independent from the proposed system architecture, and as the result any other messaging provider with chatbot functionality could be applied. The main goal of having this infrastructure was to avoid unnecessary software development of components that would provide graphical interface and type of interaction that users can experience by using mobile or web based messaging clients.

Algorithm 1 Passive Tracker Update

procedure PASSIVETRACKERUPDATE($\mathcal{T}$, $\mathcal{D}$, θ, $L_{\max}$)
  Init: $\mathcal{T}' \leftarrow \emptyset$, $\mathcal{U} \leftarrow \emptyset$, $\mathcal{R} \leftarrow \emptyset$
  #1: Greedy matching
  For each $t_i \in \mathcal{T}$:
    $j^* \leftarrow \arg\max_{j \notin \mathcal{U}} \mathrm{IoU}(b_i, d_j)$
    If $\mathrm{IoU}(b_i, d_{j^*}) > \theta$:
      $\mathcal{T}'[i] \leftarrow \{d_{j^*}, l = 0\}$
      $\mathcal{U} \cup = \{j^*\}$, $\mathcal{R} \cup = \{(i, d_{j^*})\}$
    Else: $t_i.l \leftarrow t_i.l + 1$
      If $t_i.l < L_{\max}$: $\mathcal{T}'[i] \leftarrow t_i$
  #2: New tracks
  For each $d_j \notin \mathcal{U}$:
    $\mathrm{ID}_{\mathrm{new}} \leftarrow \mathrm{ID}_{\mathrm{new}} + 1$
    $\mathcal{T}'[\mathrm{ID}_{\mathrm{new}}] \leftarrow \{d_j, l = 0\}$
  $\mathcal{T} \leftarrow \mathcal{T}'$, ***Return:*** $\mathcal{R}$
end procedure

Algorithm 2 Reporting Flow

Trigger: $event.type$ ="snapshot"

Background daemon thread:
↓
GENERATE_AND_SEND($event$)
  $path, detections, timestamp, args \leftarrow event.payload$
↓
FORMAT_ARGS($args$) $\rightarrow$ $args_str$
↓
GENERATE_CAPTION($path$, $detections$, $timestamp$, $args_str$)
  POST($OLLAMA_localhost$) $\rightarrow$ $caption$
↓
MAKE_EVENT("report", {$path$,$caption$}) $\rightarrow$ $report_event$
↓
SEND_TO_AGENT("router",$report_event$)

Fig. 4. Illustration of some of the vision agent functionalities. Algorithm 1 describes a passive tracker implementation with a greedy IoU-based tracking update. It follows a standard framework and notation described in [20], which associates new detections with the existing tracks, manages track persistence via an age counter, and initializes new tracks for unmatched detections. (a). Algorithm 2 illustrates reporting agent's asynchronous LLM system call, which is triggered by a snapshot events. Detections and metadata are processed in a background thread to generate a structured caption using a local language model, which is then forwarded to the routing system (b).

3.6 Experimental Setup

Developed prototype of the proposed multi-agent object detection framework was tested in several working modes and system configurations corresponding to potential application scenarios. Namely, the setup investigated differences in prototype operation in cases when all agents were active, against computationally less demanding modes when only a subset of agents was utilized.

The main configuration assumed the full version of the architecture in Fig. 2 and the reporting agent's LLM subsystem running locally, resulting in peak system load. Alternative configurations considered possible use cases where LLM operation could be off-loaded via an external API, when the system load would mainly be attributed to vision agent's object detection and tracking. In all analyzed working modes the system performance and resource allocation were evaluated based on the reported performance metrics, visual assessment of object detections and tracking quality, reporting speed and communication channel event responsiveness. Besides varying number of active agents, analyzed working modes mostly differed in the type of utilized object detection

models from the YOLO family, since the model size was directly affecting resource allocation of the vision agent's detection subsystem, while the resulting performance of the chosen object detection model influenced passive tracking quality and the vision event reporting frequency. Similar also holds for LLM subsystem, where models with different number of parameters required different amounts of resources, but also caused different delays and quality of the reporting agent operation. All tests were conducted in real-time without predefined scene dynamics, by placing the prototype in the roles of live public street surveillance from the distance (e.g. by filming cars and pedestrians on the street) or indoor operation (resulting in significantly larger scale of the objects).

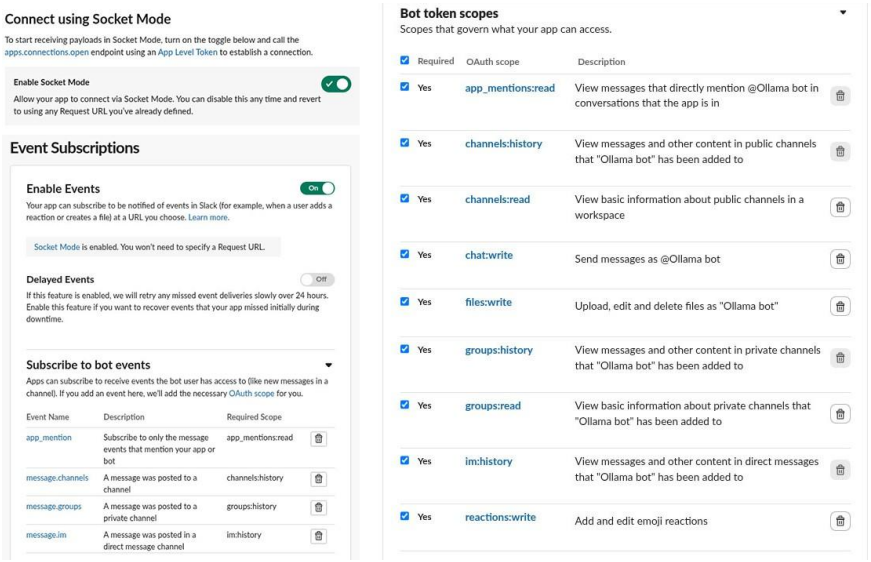


Fig. 5. Characteristics of the messaging application: socket mode communication, message events subscription and the bot token scopes defining what functionality communication and control agents can ask for from the chatbot instance.

4 Results and Discussion

Described design approach based on the Raspberry Pi platform resulted in a fully functional system, Fig. 6, implementing the proposed architecture in Fig. 2.

Presented solution showcases that the proposed multi-agent object detection framework can be an effective design template for similar edge based AI systems running on the low-cost embedded computing platforms. It also demonstrates that the transformative potential of generative AI systems can significantly reduce development costs of such systems, by leveraging on LLM based communication and control interfaces instead of developing specialized subsystems or infrastructure for the same roles. This holds both for development and deployment phase, where natural language interface can ease system testing,

as well as later adoption by the users. Thus, the initial hypothesis that the integration of multiple AI components could be done on the low resource platform in a fully centralized way was confirmed. However, the results of performed experiments also clearly indicate that such integration is still far from being perfect.

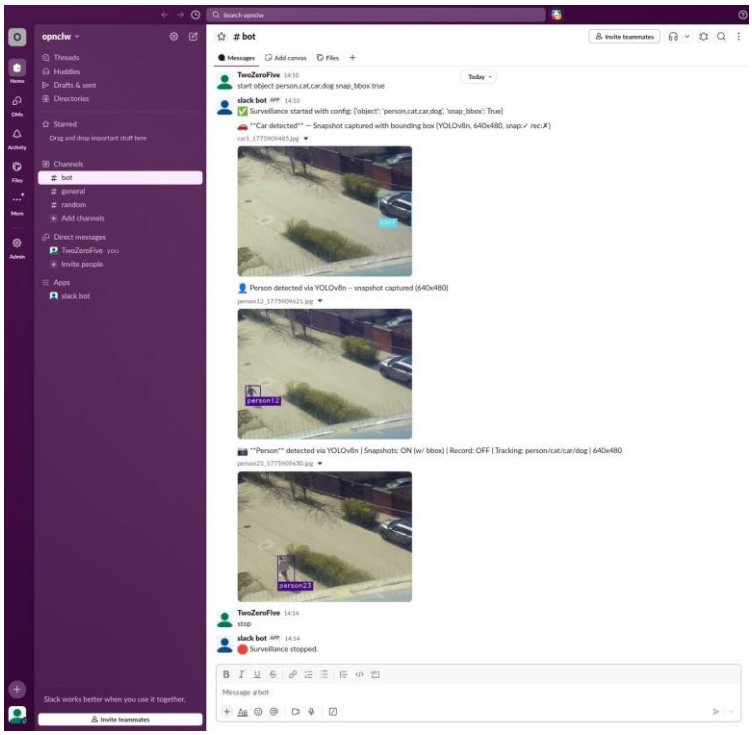


Fig. 6. Multi-agent system operation with user interaction through created natural language interface based on the Slack channel and Ollama event report generation. Both notification and control commands are provided through familiar messaging application, while the object detection and tracking events are triggering combined image and text system responses.

Although all agents were capable of performing defined tasks, in experimental configurations involving locally run LLM models, the computational burden was too high, resulting in delayed report generation for system generated vision events. Namely, although the report generation was performed in asynchronous mode, in case of multiple events it resulted in timeouts since the rate of token processing by the locally run LLM was relatively modest. Even in the working regimes with low number of vision events, the time necessary for report generation was at the order of tens of seconds (despite the carefully and conservatively designed LLM input prompts). It means that this type of system configuration with fully centralized operation of all subsystems would not be applicable to scenarios in which reporting event

handling and fast user notification would be of utmost importance. However, in cases when certain notification delay would be acceptable, this configuration would be of interest.

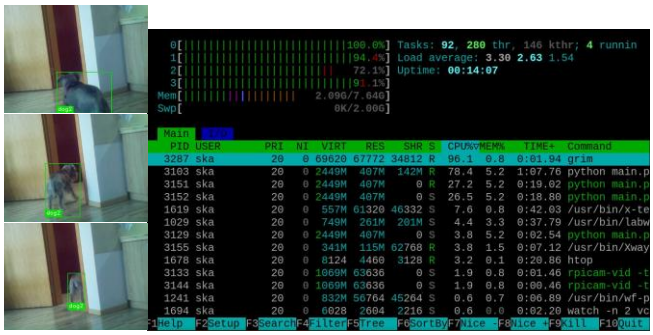
Regarding the working modes, the experiments evaluated system performance for several combinations of YOLO based object detection models and Ollama provided LLMs. We have mainly tested the so called n type of YOLO models provided by [19], which are the smallest and the fastest versions of the corresponding YOLO architectures. Thus, besides the latest yolo26n [21], the following models were also tested: yolo11n [22], yolo12n [23], yolov5nu [24], yolo5su, yolov8n [25]. In overall, the best performing model for the utilized detection and tracking setup was yolov8n, which produced the most reliable tracking output. Namely, the more complex or accurate models usually come with the price of lower frame rate, resulting in frequent tracker reinitializations, e.g. in the case of fast moving objects like cars or bicycles in the considered street surveillance experiments.

Thus, the vision agent equipped with yolov8n object detector was mostly tested with the provided Ollama models. This included tinyllama:latest (637 MB) [26], llama3.2:1b (1.3 GB) [27], gemma3:latest (3.3 GB) [28], and kimi-k2.5:cloud (non-local) [29], which were selected as the most appropriate (smallest in size), but also with varying memory footprints in order to analyze system performance under different loads. The best performing model, in terms of trade-off between generated report quality and the speed of processing was llama3.2:1b, however as already mentioned it was not performing at expected speed. It resulted in delayed system notifications, and in the worst cases lost notifications due to report generation timeouts, besides the low frame rate (less than 3 fps at 640x480 spatial resolution of the input image). Nevertheless, it showed that it was possible to achieve full configuration setup with all agents.

The reason for such performance was probably due to computationally heavy vision tasks that were performed by the vision agent. This can be seen from Fig. 7, where it is shown that even in the configurations when only the vision agent was active (without LLM) the quad core CPU load was still very high, Fig. 7b. Besides different working modes, experiments were also done by varying system configuration parameters like object categories, image resolutions, tracking and detection thresholds, live preview or visualization. It was done through the Slack channel interface, allowing for simple reconfiguration and user interaction.

Regarding the initial considerations of using a fully automated and AI based agent orchestration in the proposed solution. The conducted experiments clearly indicated that the OpenClaw design paradigm would require significantly higher token processing speeds by the locally run LLM. This holds primarily for agents orchestration, which was done by the proposed computationally lightweight event driven pipeline, implemented by the router in Fig. 2. However, it also refers to the potential automated development of multi-agent systems that would be done by the OpenClaw type application running entirely on the local low-cost platform (by using locally hosted LLM). This finding is also in line with the

reported system operation based on the kimi-k2.5:cloud LLM, which has provided the best results in combination with all vision agent models, but required that the reporting agent utilizes LLM subsystem that is not present locally on the Raspberry Pi. Instead, in these alternative system configurations LLM subsystem was running in cloud environment. We note that the integration of the OpenClaw type orchestrator on the selected platform was also tested, but it was not possible to operate such orchestrator without an external cloud based LLM.



(a) (b)

Fig. 7. Illustration of system configuration without LLM based reporting agent: (a) object tracking and video encoding, (b) CPU load and memory usage.

When it comes to the market implications of the fast prototyping based on the generative AI, obtained result confirm the ongoing trend of fundamental transformation of the way how information-communication technologies are designed and implemented. Multi-agent systems open possibilities for relatively simple integration of the existing infrastructure, by providing additional value to the previous investments, and making them the basis for new products and services. Therefore, the proposed object detection framework can be also regarded as a template for engineering solutions that avoid expensive and long development cycles or the need for reimplementation of interfaces in order to ensure their integration into new systems and roles.

5 Conclusion

Advantages of the proposed low-cost multi-agentic framework are primarily reflecting the design goal of achieving integration of complex AI based system components into compact hardware constrained platform. By demonstrating an implementation of the multi-agent system prototype based on the Raspberry Pi device it was possible to investigate and analyze the overall system performance under the different configuration setups and the working modes. This has revealed practical limitations of the current platform, but also emphasized the modular nature of the proposed architecture and software implementation.

The presented design scheme provides straightforward guidelines for further improvements through implementation of more advanced and customized AI models and accompanying algorithms, or the choice of other edge devices.

Manifold application possibilities also come from the design that supports fast prototyping through simple integration and connectivity, at both internal level between the agents, as well as towards the external environment through the natural language interface.

Compact dimensions, small form factor and low power consumption that were directing the framework design from the start, guarantee that the systems based on the provided template would be simple to use and integrate into larger industrial products or services. However, this does not exclude the framework implementation on a larger and more powerful GPU equipped full scale systems. The advantage of the low-cost edge device implementation is also the expected higher scalability and adaptation to budget constraints when the large number of vision agents would be necessary, i.e. enabling reliable device production. Expected higher market availability and simpler logistics of low-cost devices in the case of bulk procurements would be also an advantage.

Presented results confirmed that it is possible to integrate computer vision and LLM components into the same resource constrained physical system, but also that this is still challenging since the current solutions are still not compact enough to provide desired operation (e.g. without notification delays in the case of LLM based report generation). However, this does not mean that the locally run LLM instances, like Ollama models, are not possible to fully operate on the hardware constrained platform. Actually, conducted experiments involving only LLM based agents confirmed that they can be fully functional part of the system and easily accessed through the designed Slack channel interface. Thus, these properties make them a good fit for application scenarios in which real-time responsiveness is not a mandatory requirement. However, when other computational heavy AI agents are also present in the system, advanced NLP functionalities are hard to achieve without an external LLM inference. Similar also holds for the recently proposed OpenClaw AI agent orchestration paradigm, which was possible to implement on the selected platform, but was not possible to utilize without an external LLM instance.

In the future work we plan to further address these challenges and explore more specific use case scenarios involving similar edge devices.

Acknowledgments. The second author would like to acknowledge support by the Science Fund of the Republic of Serbia through the project AI-SPEAK - "Multimodal multilingual human-machine speech communication", grant agreement no. 7449; and the Ministry of Science, Technological Development and Innovation of the Republic of Serbia (Contract No. 451-03-34/2026-03/200156), for the support through the project “Scientific and artistic research work of researchers in teaching and associate positions at the Faculty of Technical Sciences, University of Novi Sad 2026”, No. 01-3609/1.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. S. Sedkaoui and R. Benaichouba, "Generative AI as a transformative force for innovation: a review of opportunities, applications and challenges," European Journal of Innovation Management, vol. 29, no. 3, pp. 677—-701, 08 2024.
2. S. Luu, A. Ravindran, A. D. Pazho, and H. Tabkhi, "VEI: A multicloud edge gateway for computer vision in IoT," in Proceedings of the 1st Workshop on Middleware for the Edge, 2022, pp. 6–11.
3. Z.-Q. Zhao, P. Zheng, S.-t. Xu, and X. Wu, "Object detection with deep learning: A review," IEEE Transactions on Neural Networks and Learning Systems, vol. 30, no. 11, pp. 3212–3232, 2019.
4. L. Wen, D. Du, Z. Cai, Z. Lei, M.-C. Chang, H. Qi, J. Lim, M.-H. Yang, and S. Lyu, "UA-DETRAC: A new benchmark and protocol for multi-object detection and tracking," Computer Vision and Image Understanding, vol. 193, p. 102907, 2020.
5. Raspberry Pi Ltd. (2026) Raspberry Pi 4 Model B Specifications. [Online]. Available: https://www.raspberrypi.com/products/raspberry-pi-4-model-b/ specifications/ (Accessed 2026-04-12).
6. OpenClaw. (2026) OpenClaw: Autonomous AI Agent Framework. [Online]. Available: https://openclaw.ai (Accessed 2026-04-14).
7. Slack Technologies. (2026) Slack. [Online]. Available: https://slack.com (Accessed 2026-04-14).
8. Ollama Inc. (2026) Ollama. [Online]. Available: https://ollama.com (Accessed 2026-04-14).
9. B. Brkljač and M. Janev, "Applications of information geometry driven deep learning," in Pattern Recognition and Computer Vision in the New AI Era, C. H. Chen, Ed. Series in Computer Vision: Volume 9, World Scientific, 2025, pp. 373–397.
10. S. K. Pal, A. Pramanik, J. Maiti, and P. Mitra, "Deep learning in multi-object detection and tracking: State of the art," Applied Intelligence, vol. 51, no. 9, pp. 6400–6429, 2021.
11. R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Region-based convolutional networks for accurate object detection and segmentation," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 38, no. 1, pp. 142–158, 2015.
12. S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137–1149, 2016.
13. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2016.
14. N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, "End-to-end object detection with transformers," in Computer Vision – ECCV 2020, A. Vedaldi, H. Bischof, T. Brox, and J.-M. Frahm, Eds. Cham: Springer International Publishing, 2020, pp. 213–229.
15. Z. Zou, K. Chen, Z. Shi, Y. Guo, and J. Ye, "Object detection in 20 years: A survey," Proceedings of the IEEE, vol. 111, no. 3, pp. 257–276, 2023.
16. D. Papakyriakou and I. S. Barbounakis, "High performance Linpack (HPL) benchmark on Raspberry Pi 4B (8GB) Beowulf cluster," Int. J. Comput. Appl, vol. 185, pp. 11–19, 2023.
17. Waveshare. (2026) Waveshare. Raspberry Pi Fan adapter. [Online]. Available: https://www.waveshare.com/pi-fan-3007.htm (Accessed 2026-04-14).
18. Raspberry Pi Ltd. (2026) Raspberry Pi Operating System. [Online]. Available: https://www.raspberrypi.com/software/operating-systems/ (Accessed 2026-04- 14).
19. Ultralytics. (2026) Ultralytics YOLOv8 Repository. [Online]. Available: https://github.com/ultralytics/ultralytics (Accessed 2026-04-12).

20. E. Bochinski, V. Eiselein, and T. Sikora, “High-speed tracking-by-detection without using image information,” in 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). IEEE, 2017, pp. 1–6.
21. G. Jocher and J. Qiu. (2026) Ultralytics YOLO26. [Online]. Available: https://github.com/ultralytics/ultralytics (Accessed 2026-04-12).
22. G. Jocher and J. Qiu. (2024) Ultralytics YOLO11. [Online]. Available: https://github.com/ultralytics/ultralytics (Accessed 2026-04-12).
23. Y. Tian, Q. Ye, and D. Doermann, “YOLO12: Attention-centric real-time object detectors,” arXiv preprint arXiv:2502.12524, 2025.
24. G. Jocher. (2020) Ultralytics YOLOv5. [Online]. Available: https://github.com/ultralytics/yolov5 (Accessed 2026-04-12).
25. G. Jocher, A. Chaurasia, and J. Qiu. (2023) Ultralytics YOLOv8. [Online]. Available: https://github.com/ultralytics/ultralytics (Accessed 2026-04-12).
26. P. Zhang, G. Zeng, T. Wang, and W. Lu, “TinyLlama: An open-source small language model,” arXiv preprint arXiv:2401.02385, 2024.
27. A. Grattafiori, A. Dubey, A. Jauhri, A. Pandey, A. Kadian, A. Al-Dahle, A. Letman, A. Mathur, A. Schelten, A. Vaughan et al., “The Llama 3 herd of models,” in Neural Information Processing Systems. Curran Associates, 2024.
28. G. Team, A. Kamath, J. Ferret, S. Pathak, and N. V. et al., “Gemma 3 Technical Report,” arXiv preprint arXiv:2503.19786, 2025.
29. K. Team, T. Bai, Y. Bai, Y. Bao, S. Cai, Y. Cao, Y. Charles, H. Che, C. Chen, G. Chen et al., “Kimi K2. 5: Visual Agentic Intelligence,” arXiv preprint arXiv:2602.02276, 2026.

Original Research Article *https://doi.org/10.5281/zenodo.23478955*

AI Agents in Business Process Automation: A Scenario-Based Comparative Analysis of RPA and Hybrid Automation in Financial Processes

Dragan Metikoš[1] [0009-0007-7633-4380], Slobodan Antić[2] [0000-0003-0449-0687], Lena Đorđević Milutinović[2] [0000-0002-3758-0147],

[1] University of Belgrade, Faculty of Organizational Sciences, Belgrade, Serbia
dm20255016@student.fon.bg.ac.rs

Abstract. Modern organizations face increasing demands for agility and efficiency in business processes, driving the adoption of process automation. Traditional approaches based on rule-driven Robotic Process Automation (RPA) are effective in structured environments but often fail to handle unstructured data and dynamic process variations. The emergence of Large Language Models (LLM)-based AI agents enables a shift toward intelligent automation, incorporating contextual understanding and adaptive decision making. This paper presents a comparative analysis of traditional RPA and a hybrid model combining RPA with LLM-based AI agents, using a client request processing workflow in the financial institution as an illustrative case. The results highlight significant differences in automation scope, system adaptability, and reduction of human intervention. The study also addresses key challenges related to data quality, security, and decision transparency. The findings indicate that hybrid intelligent automation can substantially improve process efficiency and flexibility.

Keywords: Artificial Intelligence, AI Agents, RPA, Intelligent Automation, Financial Processes

1 Introduction

The modern business environment is characterized by rapid changes, increased competition, and growing demands for organizational efficiency and agility. In such conditions, business process improvement, particularly through use of Robotic Process Automation (RPA), has become a dominant approach to optimizing routine and repetitive tasks over the past decade [1]. RPA technology enables the automation of structured and clearly defined processes based on predefined rules. However, the main limitation of this technology lies precisely in its reliance on rule-based logic. In cases where processes involve unstructured data, changing user requirements, or the need for contextual decision-making, defining clear and comprehensive rules becomes highly challenging. In practice, not all process complexities are analyzed, as extracting rules for every possible scenario requires significant time and effort, often making such automation economically unjustifiable. Furthermore, such solutions tend to be unstable, as the emergence of new scenarios may cause the software robot to fail or skip certain cases [1].

The development of agents based on Large Language Models (LLMs) opens the possibility of transitioning toward intelligent automation. LLM-based AI agents can understand natural language, interpret unstructured information, and generate contextually relevant responses, making them suitable for process automation scenarios where such capabilities are required [2]. By integrating these agents with RPA systems, it is possible to establish a hybrid automation model in which the AI agent handles interpretation and data extraction tasks, while RPA executes operational steps within information systems [1].

The aim of this paper is to provide a comparative analysis of traditional RPA approaches and a hybrid model combining RPA with LLM-based AI agents, using a client request processing workflow in the financial institution as an illustrative example. Based on this objective, the paper addresses the following research question: To what extent can a hybrid RPA and LLM-based AI agent approach extend automation coverage compared with traditional RPA in financial client request processing under the stated scenario-based assumptions? The analysis focuses on differences in the scope of automation, the level of system adaptability, and the reduction of human intervention. Special attention is given to identifying both the potential benefits and the challenges associated with the use of LLM-based AI agents, including issues related to data security, input data quality, and transparency of system behavior. Methodologically, the paper applies a simplified scenario-based comparative analysis, informed by the general scenario analysis framework proposed by Tourki et al. [3]. The framework distinguishes three broad stages: analysis, scenario definition, and evaluation. In the context of this paper, the analysis stage includes defining the process scope and the key operational parameters; the scenario-definition stage includes specifying the baseline RPA-only scenario and the alternative hybrid RPA and LLM-based AI agent scenario; and the evaluation stage includes comparison of their modeled outcomes in terms of automation coverage and time savings.

The operational parameters are informed by an anonymized automation project in the financial domain. Due to confidentiality constraints and non-disclosure agreements, implementation-specific data are presented in an aggregated form. Therefore, the results should be interpreted as a project-informed scenario comparison rather than as formal empirical validation or a universally generalizable assessment of LLM-based AI agent performance.

The remainder of this paper is structured as follows. Section 2 provides a review of the relevant literature and presents the theoretical framework, focusing on traditional RPA approaches and LLM-based AI agents. Section 3 describes the analyzed process and the baseline RPA automation scenario. Section 4 introduces the proposed hybrid model combining RPA and LLM-based AI agents. Section 5 presents the comparative analysis of the two approaches. Section 6 discusses key challenges related to the implementation of the hybrid model. Finally, Section 7 concludes the paper and outlines directions for future research.

2 Theoretical Background

RPA is a technology that enables the automation of repetitive and clearly defined tasks through software robots that mimic human interactions with information systems [1]. RPA solutions are most applied in processes characterized by stable workflows, limited variability, structured data, and clearly defined rules. The advantages of the RPA approach include rapid implementation, as it does not require changes to existing information technology (IT) infrastructure, reduced operational costs, and increased efficiency, since software robots can operate continuously (24/7) without errors [4].

However, numerous studies highlight the limitations of this type of automation, particularly in situations involving unstructured data, unpredictable variations in inputs, or the need for contextual interpretation where the employee's wider context and expertise is required [5]. In such cases, RPA systems require additional configuration, complex decision logic, or human intervention. These limitations become especially evident in processes initiated through email communication, where client requests may be expressed in various forms.

The concept of Intelligent Automation (IA) has emerged as a response to these limitations, combining RPA with artificial intelligence elements such as natural language processing, machine learning, and data analytics [5]. This approach enables the extension of automation to more complex tasks that go beyond strictly deterministic workflows.

The development of LLMs has enabled significant advancements in intelligent automation. Recent systematic reviews also emphasize that agentic AI is increasingly discussed as a broader technological paradigm for enabling autonomous or semi-autonomous decision-making, task execution, and adaptive interaction across different application domains [6]. Based on these models, LLM-based AI agents have emerged as advanced software systems capable of processing unstructured data, understanding natural language, and performing tasks that require a certain level of interpretation and decision making [1], [7]. LLM-based AI agents can be defined as autonomous or semi-autonomous software systems that use LLMs to interact with users, interpret input information in partial or fully unstructured formats, and execute tasks within business processes. These agents operate based on probabilistic models trained on large datasets, enabling them to process diverse input formats and generate contextually relevant outputs [1]. A key characteristic of these agents is their ability to understand natural language, their flexibility in handling different data formats, and their capability to adapt to various scenarios without the need for strictly predefined rules for each individual case. In the context of business processes, these capabilities are typically reflected in [5], [7]:

- Understanding natural language and interpreting user requests;
- Identifying user intent based on communication content;
- Extracting relevant data from unstructured sources (e.g., emails, PDF documents);
- Generating structured outputs suitable for further automated processing.

These capabilities enable agents to take over tasks that previously required human intervention, particularly in processes initiated through communication with clients or users.

In addition to functional capabilities, understanding the architecture of LLM-based AI agents is essential for analyzing their role in business processes. Although specific implementations may vary, most modern AI agents rely on modular architecture composed of multiple components. At a high level, an LLM-based AI agent consists of the following key elements [1], [7]:

- LLM model – the central component responsible for natural language processing, interpretation of input data, and generation of responses or structured outputs.
- Input layer – enables the reception of data from various sources, such as emails, documents, or API calls.
- Orchestration layer – manages tasks execution flow, defines the sequence of steps, and determines actions based on input data.
- Integration layer (tool/API layer) – enables interaction with external systems such as ERP, CRM, and other business applications, databases, or automation tools such as RPA bots.
- Output layer – generates results in a format suitable for further processing, either as structured data or as user-facing responses.

Importantly, the AI agent does not operate as an isolated system but as a part of broader automation architecture, often playing a role in the initial stages of a process. For example, after performing data extraction and interpretation, the agent can trigger an RPA bot to execute operational steps within information systems.

Fig. 1 illustrates a simplified representation of the key components of an LLM-based AI agent. The agent receives unstructured inputs, uses a language model for interpretation and data extraction, and then, through orchestration and integration layers, triggers task execution and generates the final output.

Unlike traditional RPA systems, LLM-based AI agents rely on probabilistic models trained on large datasets, enabling a higher degree of adaptability and flexibility. By integrating LLM-based AI agents with existing RPA solutions, a hybrid automation model is created in which the AI agent performs cognitive tasks (e.g., understanding requests and structuring data), while RPA executes operational activities within information systems [1], [7]. Comparative studies suggest that this approach can significantly expand the scope of automation to processes that were previously unsuitable for automation, particularly in domains characterized by high variability in input data. However, the adoption of LLM-based AI agents also introduces new challenges, including issues related to reliability of generated outputs, data security, and system transparency.

In financial institutions, the hybrid model combining RPA and LLM-based AI agents represents a promising solution for automating a wide range of processes. On one hand, these processes often involve communication with clients via email, where input structures vary significantly. On the other hand, they include numerous repetitive manual tasks within information systems. This approach forms the basis for the comparative analysis presented in the following sections of this paper [8], [9].

The financial sector is used in this study as an illustrative example, as it represents a domain with a high volume of client-driven processes, a significant share of unstructured inputs, and strict regulatory constraints. These characteristics make it particularly suitable for examining both the limitations of rule-based automation and the potential of hybrid models that combine RPA with LLM-based AI agents. The analysis focuses primarily on automation coverage, process adaptability, and potential time savings. Although data privacy, security, compliance, and governance are essential for implementation in financial institutions, a detailed assessment of these aspects is beyond the scope of this paper and is identified as an important direction for future research. Previous research on AI and RPA usage in finance and accounting processes indicates that automation adoption in this domain depends not only on technological capabilities, but also on process standardization, availability of digital data, and employee competencies [4], [8].

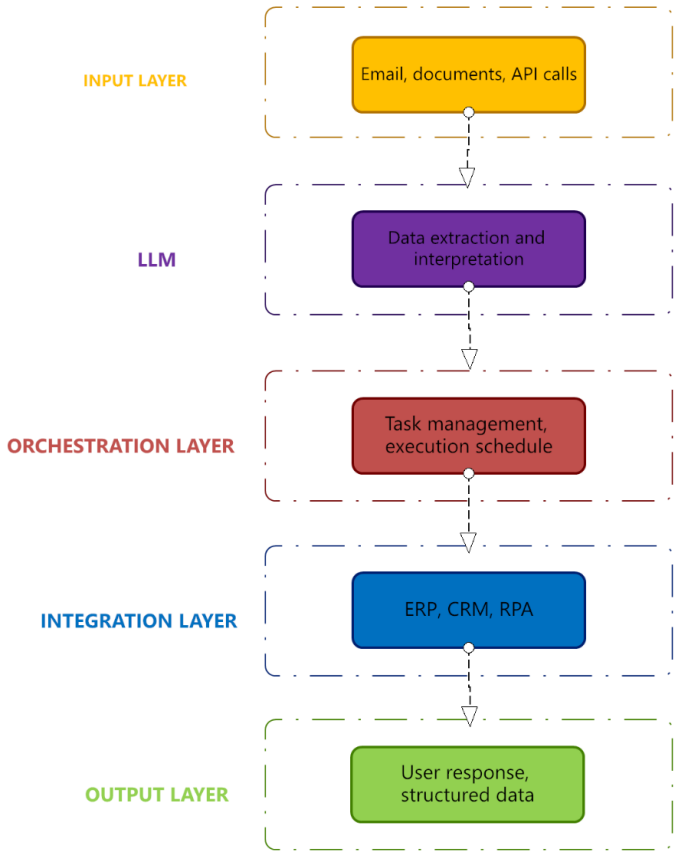


Fig. 1. Simplified overview of key elements of an LLM-based AI agent

3 Process Description and RPA Scenario

For the purpose of the comparative analysis, an anonymized client request processing workflow from a financial institution was examined. The workflow is derived from a real-world automation project and includes receiving client requests via email and entering the required data into the institution's information system. Due to confidentiality constraints and non-disclosure agreements, organization-specific details are omitted, and the process is presented in a simplified and aggregated form. Since this section describes the analyzed project context, rather than a general theoretical model, it is based on project information and does not require external references.

3.1 Description of the Analyzed Process

The process begins with the receipt of a client request via email. The requests typically relate to issuing various types of documents (e.g., account balance confirmations, certificates of no outstanding debt, or account statements for a specific period). The email content may be standardized (using predefined templates) or written in free-text form. After receiving the request, the employee performs the following steps:

1. Identification of the request type based on the email content;
2. Verification of whether all required data are provided (e.g., name, account number, period);
3. Access to the bank's information system;
4. Entry of request parameters;
5. Generation of the appropriate document;
6. Sending the response to the client.

The client request processing workflow and the sequence of activities described above are illustrated in Fig. 2.

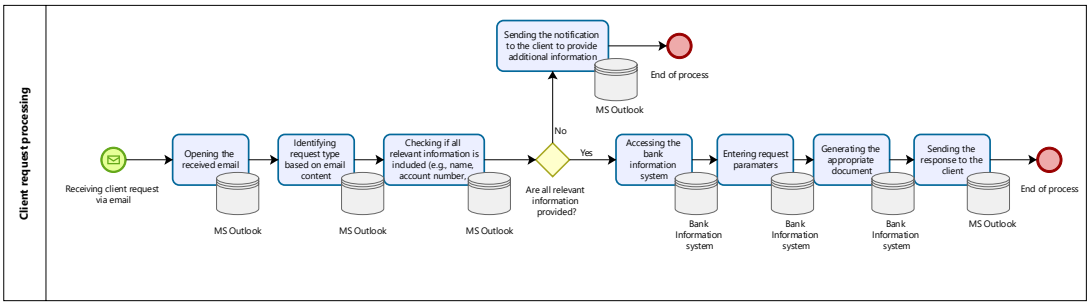


Fig. 2. Client request processing workflow

The process is partially structured. The second part of the workflow is fully defined and executed in the same manner depending on the request type, while the initial part may involve variations in the formulation and structure of input data.

3.2 RPA Automation Scenario

In the scenario where only RPA technology is applied, automation is based on predefined rules and decision logic. The software bot is configured to:

- Monitor incoming emails;
- Filter emails based on predefined subject lines;
- Recognize standardized patterns in the email body;
- Parse data based on known templates;
- Access the bank's information system and enter predefined parameters;
- Send responses directly to the client or forward them to an employee for validation and final delivery.

However, in situations where the email subject deviates from predefined patterns, data is not clearly structured, required information is missing, or the request is implicitly formulated, the RPA system is unable to reliably interpret the content without additional logic or human intervention. In such cases, the process is redirected to manual handling, reducing the overall level of automation.

Therefore, the RPA scenario proves to be effective in environments characterized by a high degree of standardization and stable input data. However, it remains limited in terms of adaptability and the ability to process unstructured information. If we assume that more than 50% of incoming requests in an organization are unstructured, this represents a significant loss of automation potential and time-saving opportunities within the process.

4 Hybrid Model: RPA + LLM-Based AI Agent

In order to overcome the limitations of the RPA-only approach described in the previous section, a hybrid automation model is considered, combining RPA technology with an LLM-based AI agent.

4.1 Role of the LLM-based AI Agent

In the hybrid scenario, the LLM-based AI agent is responsible for processing all unstructured input data received via email. Unlike rule-based parsing, the LLM-based AI agent leverages a language model capable of:

- Understanding natural language;
- Identifying user intent (e.g., request for confirmation, request for account statement);
- Extracting relevant data from unstructured text;
- Detecting missing information;
- Generating structured outputs for further processing;
- Based on the content of the email, the LLM-based AI agent classifies the request.

In cases where the input data is incomplete, the agent can automatically generate a response requesting additional information. If the request is complete, the agent triggers the RPA bot to execute the remaining steps of the process.

4.2 Integration with RPA System

Once the LLM-based AI Agent performs data interpretation and structuring, the RPA system takes over the operational part of the process. The software bot uses the structured output generated by the agent and performs the following activities:

1. Accessing the bank's information system;
2. Entering request parameters;
3. Generating the appropriate confirmation or statement;
4. Sending the response to the client.

In this way, RPA retains its core role in automating well-defined system-based activities, while the LLM-based AI agent extends the scope of automation by enabling the processing of additional, previously no-automation scenarios.

4.3 Changes in the Process Flow

Compared to the RPA-only scenario, the hybrid model introduces several key changes:

- Elimination of dependency on strictly predefined email templates;
- Increased tolerance to variations in request formulation;
- Reduction in the number of cases requiring manual processing;
- Ability to process implicit or more complex requests.

As a result, a higher level of process adaptability is achieved, leading to a significant reduction in the need for human involvement. However, the introduction of LLM-based AI agents also raises new considerations related to the reliability of data interpretation, control over generated outputs, and risk management, which will be further discussed in the following sections [1], [7].

The hybrid model combining RPA and LLM-based AI agents thus represents an evolution from traditional automation toward intelligent automation, where cognitive and operational functions are clearly separated but technically integrated.

5 Scenario-based Comparative Analysis

To illustrate the potential effects of different automation models, a scenario-based comparative analysis was conducted using typical operational parameters of administrative financial processes. The parameters used in this analysis are based on typical operational values observed in real-world process automation projects within the financial sector. Due to confidentiality constraints and non-disclosure agreements, exact organizational details cannot be disclosed. Therefore, the values are presented in an

aggregated and illustrative form and serve as a comparative illustration of the effects of a pure RPA approach and hybrid model, rather than results of a formal experimental study.

Table 1. Distribution of incoming client requests (monthly basis)

Parameter	Value
Average number of emails	120
Number of working days per month	22
Total number of requests per month	2.640
Number of standardized emails	40% (1.056)
Number of unstructured emails	60% (1.584)
Average time of manual request processing	4 minutes
Total time to manually process all requests per month	**176h**

Table 1 presents the basic operational parameters of the analyzed client request processing workflows. On a monthly basis, the organization processes a total of 2.640 requests, of which 60% are unstructured emails and 40% are standardized requests. This distribution indicates that a significant portion of the process involves unstructured inputs requiring interpretation, which represents a key challenge for automation based solely on RPA technology. The total time required for manual processing amounts to 176 hours per month, highlighting the potential for efficiency improvements through automation.

Table 2. RPA Scenario

Parameter	Value
Number of automated requests through RPA technology only	1,056
Manually processed requests	1,584 (remaining 60%)
Total time of manual processing per month	106h
Total time savings per month – RPA technology only	**70.4h**

In the scenario where only RPA technology is applied, automation is possible exclusively for standardized requests, covering 40% of the total volume, while the remaining 60% still require manual processing. Although RPA achieves a considerable time saving of 70.4h per month, a large portion of the process remains manual. This indicates the limited scope of rule-based automation in environments with a high share of unstructured inputs. The remaining manual processing time of 106 hours per month further emphasizes the potential for additional process optimization.

Table 3. Hybrid model: RPA + LLM-based AI agent

Parameter	Value

Number of automated requests through LLM-based AI agent	1,188 (75% of unstructured emails)
Number of automated requests through RPA technology	1,056
Total time of manual processing per month	26.4h
Total time savings per month – LLM-based AI agent only	79.2h
Total time savings per month – RPA technology only	70.4h
Total time savings per month applying the hybrid model	**149.6h**

The scenario-based comparative analysis suggests a potential expansion of automation coverage. In addition to the automation of standardized requests through RPA (1,056 requests), the hybrid scenario assumes that the LLM-based AI agent can process an additional 75% of unstructured requests (1,188 requests). It is important to note that automation of unstructured inputs is not complete (100%), as certain cases still require human intervention. These include more complex or ambiguous requests, missing information in the email content, and situations requiring additional validation when the confidence level of interpretation is lower. Therefore, the 75% automation rate was used as a scenario-based estimate derived from the observed limitations of the analyzed project context. As a result, manual processing time is reduced to only 26.4 hours per month, while total time savings reach 149.6 hours per month, representing more than double the savings achieved by the RPA-only approach.

Under the modeled assumptions, the comparison indicates that while traditional RPA delivers significant operational benefits in stable and structured environments, its impact is limited when a large portion of inputs is unstructured. In contrast, the integration of LLM-based AI agents extends the scope of automation to previously non-automated scenarios. An additional benefit of the hybrid model becomes evident in scalability scenarios. As the volume of requests increases, the adaptability of the AI agent enables a higher level of automation, particularly in environments where the share of unstructured input remains constant or increases over time.

6 Challenges in Implementing the Hybrid Model

Although the hybrid model combining RPA and LLM-based AI agents demonstrates significant potential for improving process efficiency, its implementation introduces a range of technical, organizational, and regulatory challenges. In financial institutions, where processes are highly regulated and subject to strict oversight, these challenges become particularly pronounced [9].

As with all AI-based solutions, data represents a critical factor. The effectiveness of LLM-based AI agents largely depends on the quality of input data. In the observed case, unstructured client requests may include:

- Incomplete information;
- Ambiguous formulations;
- Typographical errors;
- Implicit requests.

If the agent misinterprets the request content, this may lead to incorrect classification or inaccurate data extraction. Therefore, it is essential to introduce additional validation mechanisms, such as:

- Multi-step output validation;
- Confidence thresholds;
- Escalation mechanisms to human operators.

Without adequate data quality control, there is a significant risk of automation errors, which may negate expected operational benefits and potentially introduce additional inefficiencies.

The financial institution involves the processing of sensitive client data. The introduction of LLM-based AI agents raises critical questions related to data storage and processing, whether performed on-premise or in cloud environments, as well as issues concerning data usage for model training, encryption, access control, and the management of systems logs and audit trails. Many organizations still lack clearly defined internal policies for the use of AI technologies, which significantly increase implementation risks. Although this paper focuses primarily on the automation potential of the hybrid RPA and LLM-based AI agent model, data privacy and security represent essential implementation considerations in financial systems. Relevant practices include secure data storage and processing, encryption of data in transit and at rest, role-based access control, audit trails, restrictions on the use of client data for model training, and clearly defined escalation mechanisms for cases requiring human validation. [9], [10]

One of the key challenges in the implementation of LLM-based AI agents is decision transparency. While RPA systems rely on fully deterministic logic, LLM models operate based on probabilistic reasoning and statistical patterns. In highly regulated environments, this raises important questions, how to explain why a particular request was classified in a certain way, how to document the decision-making process, and how to ensure that model outputs can be validated and verified by responsible personnel [11], [12], [13]. The lack of transparency is not only a barrier to developing reliable solutions but also a major obstacle to their adoption within financial institutions. Although the analyzed process does not involve credit-risk approval, similar requirements apply to financial process automation more broadly, as AI-supported financial decision-making requires explainability, accountability, and clearly defined human oversight mechanisms [14].

Beyond technical challenges, the implementation of a hybrid automation model requires a shift in organizational approach. This includes:

- Redefining employee roles;
- Introducing oversight mechanisms for AI systems;
- Developing competencies for managing intelligent automated processes.

The hybrid model does not eliminate the need for human involvement, rather, it shifts the focus from operational data entry toward supervision, control, and continuous process optimization [9], [10].

In addition to the implementation challenges discussed above, the present analysis has several methodological limitations. The scenario-based comparison relies on aggregated project-informed parameters and does not include formal empirical validation, a statistical sensitivity analysis, or a detailed assessment of long-term implementation costs. Furthermore, dimensions such as reliability, explainability, exception handling, and governance requirements are discussed only qualitatively. Future research should address these limitations through empirical case studies in real organizational settings, broader cost-benefit analyses, and the development of methodologies for measuring the reliability and explainability of AI agents in business processes.

7 Conclusion

Software-based automation of business processes has, for many years, represented a significant direction in the development of information systems, with RPA technology enabling efficient execution of repetitive and deterministic tasks. Existing research in the field of intelligent automation highlights the need to overcome the limitations of rule-based approaches, particularly in processes involving unstructured data and variable inputs. The development of Large Language Models and AI agents opens new possibilities for integrating cognitive capabilities into traditional automation workflows.

This paper presented a comparative analysis of two automation models using a client request processing scenario in the financial institution. A model based solely on RPA technology and a hybrid model combining RPA with LLM-based AI agent. The scenario-based comparative analysis demonstrated that while the traditional RPA approach delivers significant operational benefits in environments with a high degree of input standardization, its scope remains limited in processes with a higher proportion of unstructured requests. In contrast, the hybrid model exhibits a higher level of adaptability, broader automation coverage, and significantly greater time savings, particularly in scenarios involving increased process volumes, where system scalability becomes more evident. At the same time, the analysis identified key challenges in the application of LLM-based AI agents, including issues related to data quality, security, and decision transparency. In highly regulated environments such as the financial institutions, the sustainable implementation of intelligent automation requires the establishment of robust mechanisms for control, validation, and oversight. In relation to the research question, the scenario-based comparative analysis suggests that a hybrid RPA and LLM-based AI agent approach can extend the automation potential of traditional RPA by enabling the processing of selected unstructured client requests that would otherwise require manual handling.

Future research may focus on empirical validation of the modeled results through case studies in real organizational settings, analysis of long-term cost implications of hybrid automation models, and the development of methodologies for measuring the reliability and explainability of AI agents in business processes. A particularly important direction for future research is the definition of a standardized framework for integrating AI agents into existing process management systems while ensuring regulatory compliance and operational stability. As a part of the framework, broader

empirical evaluation of hybrid RPA and LLM-based AI agent models by considering not only time savings and automation coverage, but also accuracy, exception handling, human escalation, compliance risk, explainability, implementation complexity, operating costs, and governance requirements should be included. As this technology continues to evolve, it is expected to address some of the current challenges while simultaneously introducing new ones that will require further investigation [10], [11]. Future research should also examine data privacy, security, and governance requirements for the implementation of LLM-based AI agents in financial process automation, particularly in relation to sensitive client data, auditability, human oversight, and regulatory compliance.

Ethical AI Disclosure. The authors used a generative AI-based language model (ChatGPT, Open AI) as an auxiliary tool for language editing, clarity improvement, and stylistic refinement of selected passages and language-level revisions throughout the manuscript where needed. The tool was not used for generating research ideas, defining the research problem, developing the conceptual framework, creating scenario assumptions, conducting the analysis, interpreting the results, or formulating the scientific conclusions. All conceptual framing, methodological decisions, scenario assumptions, analytical interpretation, and final responsibility for the content remain with the authors.

Acknowledgements. This research was supported by the University of Belgrade – Faculty of Organizational Sciences and, in part, by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia through institutional funding (grant number: 200151).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Afrin, S., Roksana, S., Akram, R.: AI-enhanced robotic process automation: A review of intelligent automation innovations. IEEE Access 13, 173–197 (2025). https://doi.org/10.1109/ACCESS.2024.3513279
2. Abou Ali, M., Dornaika, F., Charafeddine, J.: Agentic AI: A comprehensive survey of architectures, applications, and future directions. Artificial Intelligence Review 59, Article 11 (2026). https://doi.org/10.1007/s10462-025-11422-4
3. Tourki, Y., Keisler, J., Linkov, I.: Scenario analysis: A review of methods and applications for engineering and environmental systems. Environment Systems and Decisions 33(1), 3–20 (2013). https://doi.org/10.1007/s10669-013-9437-6
4. Siderska, J., Aunimo, L., Süße, T., von Stamm, J., Kedziora, D., Aini, S.N.B.M.: Towards intelligent automation (IA): Literature review on the evolution of robotic process automation (RPA), its challenges, and future trends. Engineering Management in Production and Services 15(4), 90–103 (2023). https://doi.org/10.2478/emj-2023-0030
5. Cherukuri, R., Yarram, V.K.: From intelligent automation to agentic AI: Engineering the next generation of enterprise systems. International Journal of Emerging Research in

Engineering and Technology 5(4), 142–152 (2024). https://doi.org/10.63282/3050-922X.IJERET-V5I4P114
6. Hosseini, S., Seilani, H.: The role of agentic AI in shaping a smart future: A systematic review. Array 26, Article 100399 (2025). https://doi.org/10.1016/j.array.2025.100399
7. Brohi, S., Mastoi, Q.-u.-A., Jhanjhi, N.Z., Pillai, T.R.: A research landscape of agentic AI and large language models: Applications, challenges and future directions. Algorithms 18(8), Article 499 (2025). https://doi.org/10.3390/a18080499
8. Nechiporenko, A.: Artificial intelligence and RPA usage and competencies in Finnish SMEs' finance and accounting processes. Bachelor's thesis, Haaga-Helia University of Applied Sciences (2022). https://urn.fi/URN:NBN:fi:amk-2022061517987
9. Inala, R., Somu, B.: Building trustworthy agentic AI systems for personalized banking experiences. Metallurgical and Materials Engineering 31(5), 1336–1360 (2025). https://doi.org/10.63278/mme.vi.1714
10. Shavit, Y., Agarwal, S., Brundage, M., Adler, S., O'Keefe, C., Campbell, R., Lee, T., Mishkin, P., Eloundou, T., Hickey, A., Slama, K., Ahmad, L., McMillan, P., Beutel, A., Passos, A., Robinson, D.G.: Practices for governing agentic AI systems. OpenAI, Technical Report (2024). https://cdn.openai.com/papers/practices-for-governing-agentic-ai-systems.pdf
11. Mukherjee, A., Chang, H.H.: Agentic AI: Autonomy, accountability, and algorithmic society. arXiv preprint arXiv:2502.00289 (2025). https://doi.org/10.48550/arXiv.2502.00289
12. Chaduvula, S., Ho, J., Kim, K., Narayanan, A., Alinoori, M., Garg, M., Ramachandram, D., Raza, S.: From features to actions: Explainability in traditional and agentic AI systems. arXiv preprint arXiv:2602.06841 (2026). https://doi.org/10.48550/arXiv.2602.06841
13. Raza, S., Radwan, A.Y., Chaduvula, S., Alinoori, M., Emmanouilidis, C.: Transparency in agentic AI: A survey of interpretability, explainability, and governance. EngrXiv Preprint, No. 6451 (2026). https://engrxiv.org/preprint/view/6451
14. Kubam, C.S.: Agentic AI for autonomous, explainable, and real-time credit risk decision-making. International Journal of Intelligent Systems and Applications in Engineering 12(23s), 3669–3676 (2024). https://www.ijisae.org/index.php/IJISAE/article/view/7827

Original Research Article *https://doi.org/10.5281/zenodo.23478981*

Towards an Adaptive Hierarchical Deep Reinforcement Learning Architecture for Resilient Financial Systems in Non-Stationary Markets

Dragisa Stanojevic[1][0009-0000-8280-6881] and Dragan Vukmirovic[1][0000-0003-0248-016X]

[1] University of Belgrade, Faculty of Organizational Sciences, Jove Ilica 154, 11000 Belgrade, Serbia
ds20255085@student.fon.bg.ac.rs

Abstract. Financial markets are inherently non-stationary, causing monolithic Deep Reinforcement Learning (DRL) agents to fail due to catastrophic forgetting and adaptation latency. To address these limitations, we propose a conceptual four-layer Adaptive Hierarchical DRL (HDRL) architecture with the foundational representation layer (L0) validated empirically. The framework leverages Chronos-T5 for bias-free latent extraction and synthesizes probabilistic models, change-point detection, and Wasserstein distance for rapid regime detection. To mitigate forgetting, the conceptualized framework employs a heterogeneous Mixture of Experts that specializes across regimes. A meta-controller manages capital allocation, optimizing Conditional Value at Risk (CVaR) for global tail-risk protection, while sub-agents maximize the Differential Sharpe Ratio. Empirical validation on 4,876 trading days shows that the Layer-0 pipeline reduces the average composite distribution shift by 64.82%, ensures strict stationarity (Augmented Dickey-Fuller test across all features), and yields highly compressible latent representations (124 principal components retain 90% of variance). These embeddings demonstrate clear unsupervised separation of major shocks (2008 crisis, COVID-19, 2022 inflation), providing empirical evidence of high embedding discriminability, a necessary prerequisite for prospective regime detection.

Keywords: Hierarchical Deep Reinforcement Learning · Algorithmic Trading · Regime Detection · Catastrophic Forgetting · Mixture of Experts · Risk-Sensitive RL

1 Introduction

The contemporary global economy operates in a state of continuous disruption, increasingly defined by organizational theorists as a "permacrisis" characterized by overlapping macroeconomic shocks and geopolitical instability [9]. Within this volatile environment, financial markets exhibit profound stochastic non-stationarity, where the statistical properties of asset returns undergo rapid and unpredictable structural shifts [33]. For financial institutions, organizational resilience has transcended the traditional paradigm of merely surviving isolated crisis events [6]. True resilience now demands the cultivation of intelligent systems capable of continuous, dynamic adaptation, ensuring

that critical capital allocation and risk management functions remain robust regardless of the prevailing market regime [12, 22, 25].

Deep Reinforcement Learning (DRL) has emerged as a transformative paradigm for sequential decision-making in quantitative finance. Unlike traditional supervised learning models that attempt to minimize forecasting errors on future price trajectories, DRL directly optimizes action policies to maximize long-term cumulative utility, naturally encompassing the sequential nature of portfolio management [29]. However, the transition of DRL from academic simulation to enterprise-grade deployment has exposed fundamental algorithmic vulnerabilities. Standard monolithic DRL architectures rely critically on the assumption that the underlying data-generating process is Independent and Identically Distributed (I.I.D.). Financial markets fundamentally violate this assumption due to continuous regime switching. When a monolithic DRL agent, optimized predominantly on low-volatility bullish data, encounters a sudden structural break, it suffers from catastrophic forgetting. In its attempt to rapidly minimize new temporal difference errors, the agent aggressively updates its neural parameters, effectively overwriting previously learned optimal behaviors [21]. Furthermore, monolithic agents are severely compromised by adaptation latency, the simulation-to-reality (Sim2Real) gap, and reward misspecification, where standard metrics like the Sharpe ratio fail to penalize extreme tail losses [16].

To address these systemic flaws, this paper introduces a conceptual Adaptive Hierarchical Deep Reinforcement Learning (HDRL) architecture. The proposed framework conceptualizes algorithmic trading not as a singular, monolithic optimization problem, but as a hierarchical sequence of context evaluation, specialized execution, and risk-sensitive meta-control. The architecture is designed to address the challenges of non-stationarity through four vertically integrated layers: offline latent extraction via foundation models (L0), hybrid regime detector (L1), a heterogeneous Mixture of Experts (MoE) with frozen parameters intended to mitigate forgetting (L2), and a risk-sensitive meta-controller for dynamic capital allocation (L3).

The primary contributions of this research are twofold. First, it proposes an integrated architectural framework that synthesizes Hidden Markov Models, Bayesian change-point detection, optimal transport theory, and hierarchical multi-agent reinforcement learning. Within this architecture, the reward structure is deliberately decoupled: sub-agents are designed to optimize the Differential Sharpe Ratio (DSR) for local gradient stability, while a global meta-controller focuses on Conditional Value at Risk (CVaR) to manage tail-risk. Second, it provides extensive empirical validation of the L0 implementation across 4,876 trading days, demonstrating that foundation model embeddings exhibit high discriminability, providing a necessary prerequisite for prospective regime detection while satisfying strict stationarity requirements. To rigorously evaluate this conceptual architecture in future implementations, the paper also outlines a structured experimental protocol utilizing Generative Adversarial Networks (GANs) to address the Sim2Real gap through out-of-distribution (OOD) stress testing.

The remainder of this paper is structured as follows. Section 2 provides a critical review of the related literature. Section 3 details the methodology and mathematical formalization of the HDRL system. Section 4 details the experimental framework and presents the preliminary empirical results of the foundational layer (L0). Section 5

outlines the proposed evaluation protocol for L1-L3 and the proposed ablation study, while Section 6 concludes the paper.

2 Related Work

The intersection of deep reinforcement learning, financial forecasting, and regime detection represents a rapidly expanding and highly interdisciplinary domain [22]. A critical review of the literature reveals significant advancements in individual sub-fields [12, 13] but exposes lingering gaps regarding the unification of these technologies into cohesive, resilient systems capable of zero-shot adaptation [22, 28].

2.1 Monolithic DRL and Continual Learning Challenges

The application of DRL in quantitative finance has been extensively documented [15], evolving from value-based methods to continuous actor-critic algorithms like PPO, SAC, and TD3 [10–12, 27]. However, the literature consistently demonstrates that monolithic agents exhibit severe policy degradation in non-stationary environments [33]. When a single neural network attempts to generalize across highly divergent market conditions, it converges to a sub-optimal "average" policy. Consequently, monolithic architectures lack the necessary zero-shot adaptation capabilities required for resilient financial operations [22]. Furthermore, the translation of these models into live environments is frequently hindered by the Sim2Real gap [16, 22].

2.2 Regime Detection and Optimal Transport

To mitigate the impact of non-stationarity, market regime detection has been heavily explored [32, 33]. Traditionally, detecting structural breaks relies on Hidden Markov Models (HMMs) [5]. While HMMs provide interpretable probabilistic outputs, they are burdened by an inherent "detection lag", requiring a sequence of consecutive observations to confidently transition between states [23]. Bayesian Online Changepoint Detection (BOCPD) offers a complementary rapid detection mechanism [1]. However, recent extensions emphasize that standard BOCPD assumes I.I.D. data within regimes, a premise often violated by the temporal correlations in financial series [31].

Recently, optimal transport theory, specifically the Wasserstein distance, has emerged as a powerful tool for multivariate regime clustering [14]. By measuring the geometric distance between empirical distributions, Wasserstein-based methods capture subtle deformations in higher statistical moments, such as skewness and kurtosis, before the mean and variance definitively diverge [4, 17]. While the integration of Wasserstein-based detection into offline portfolio construction has been shown to materially increase risk-adjusted performance [24], its direct, real-time integration into a continuous DRL control loop remains a largely unexplored frontier [28].

2.3 Foundation Models in Time Series

The advent of foundation models is revolutionizing natural language processing, and this paradigm is currently shifting toward time-series analysis [2, 26]. Emerging architectures, such as Chronos-T5, follow a minimalist approach by tokenizing continuous numerical values into a fixed vocabulary, allowing standard language model architectures to process temporal data without domain-specific features [2]. These models, alongside other foundational frameworks like Lag-Llama and MOIRAI, have demonstrated exceptional zero-shot forecasting capabilities across diverse domains. In algorithmic trading, they successfully compress complex non-linear temporal dependencies into dense latent embedding vectors [26]. However, the literature currently lacks a standardized hierarchical framework for decoupling foundation model feature extraction from the online RL execution cycle while actively preventing look-ahead bias and inference latency [16, 26].

2.4 Mixture of Experts and Risk-Sensitive Rewards

To combat catastrophic forgetting and improve model capacity, researchers have increasingly adopted Mixture of Experts (MoE) architectures [19]. In an MoE framework, the parameter space is vastly expanded without a proportional increase in computational cost by selectively activating specific expert subnetworks. In RL, this modularity enables agents to learn distinctly different policies for diverse tasks without interference [21].

However, existing MoE implementations in financial DRL predominantly utilize homogeneous experts [28], a practice that induces "parameter contamination" and results in misaligned forecasts when market conditions shift [33]. This approach fails to account for the fundamentally distinct stochastic properties of different regimes, which necessitate heterogeneous optimization mechanics [8, 19]. Furthermore, while standard risk-aware models attempt to bake risk penalties directly into a single objective function, this linear combination often destabilizes the gradient descent process due to conflicting reward signals [12, 22], leading to highly conservative or sub-optimal policies [30].

2.5 Identified Research Gaps

The critical review reveals a lack of a unified architecture capable of addressing financial non-stationarity, with existing methods often operating in isolation [6, 33]. Specifically, while geometric regime detection shows promise, its direct integration into a real-time DRL control loop remains an unexplored frontier [28]. Furthermore, there is a documented lack of heterogeneous MoE deployment matched to regime-specific stochastic properties, leading to "parameter contamination" in standard models [33]. Current literature also highlights insufficient reward bifurcation to separate local gradient stability from global tail-risk protection [8, 22], and limited designs that successfully decouple foundation model embeddings without introducing look-ahead bias [16, 26].

The proposed conceptual HDRL architecture is designed to address these vulnerabilities, thereby achieving robust organizational resilience.

3 Methodology

The proposed solution operates by deconstructing the complex, highly dimensional portfolio optimization problem into a hierarchy of statistically homogeneous subproblems to strengthen the learning process of individual abstractions [28]. The architecture is structured into four vertically integrated modules (L0 to L3). This modularity is designed to ensure statistical robustness and systematically eradicate the phenomenon of catastrophic forgetting by separating strategic goals from execution mechanics.

3.1 Problem Formalization

Standard Markov Decision Process (MDP) assumptions are fundamentally untenable in quantitative finance because the underlying data-generating process is subject to constant changes due to external factors, a state defined as passive non-stationarity [7]. In such environments, an agent possesses only a partial, noisy view of the true macroeconomic environment, where hidden institutional liquidity and impending structural shifts remain unobservable. Consequently, the portfolio optimization problem is rigorously formalized as a Partially Observable Markov Decision Process (POMDP), defined by the tuple $\langle S, A, T, R, \Omega, O, \gamma \rangle$:

- State Space (S): The set of unobservable, latent market regimes $s_t \in R^d$ generating the observed dynamics. This accounts for the fact that market systems have hidden states that an agent can only indirectly infer from historical sequences [5].
- Action Space (A): A continuous vector $a_t = [w_1, w_2, \dots, w_n]$ representing capital allocation across n assets, subject to individual asset bounds $w_i \in [-1,1]$ and the gross exposure constraint $\sum_{i=1}^{n} |w_i| \leq 1.0$.
- Transition Function (T): Models the probability of transitioning between latent states. Assuming the agent's trades have minimal market impact, this is approximated as $T(s_{t+1}|s_t)$.
- Observation Space (Ω): A composite vector $o_t = [p_t^{OHLCV}, \tau_t, \nu_t, \kappa_t, L_t, \psi_t]$ integrating raw prices, technical indicators, latent embeddings (L_t), and sentiment (ψ_t).
- Observation Function (O): $O(o_t|s_{t+1}, a_t)$ models the probability of perceiving observation o_t given the underlying latent state.
- Reward Function (R): A risk-adjusted utility metric, bifurcated into localized gradient signals and global tail-risk constraints.
- Discount Factor (γ): $\gamma \in [0,1]$ balancing short-term execution and long-term capital growth.

3.2 Hierarchical Architecture Overview

The HDRL architecture operates via a unidirectional forward pass with directed feedback through reward evaluation at the meta-level. The deconstruction of the optimization problem is mapped across four distinct layers, each addressing a specific operational vulnerability of monolithic DRL systems.

Raw observations are first processed through the offline L0 module to map into high-dimensional latent spaces. These enriched representation vectors are routed in parallel to L1 (which evaluates the stochastic context and emits an updated probability distribution vector p_t) and L2 (where narrow-domain experts formulate optimal local allocation vectors a_i). The final L3 node, operating as an active meta-agent, aggregates the probabilities and proposals to emit an optimal, risk-sensitive weight vector w_t (see Fig. 1).

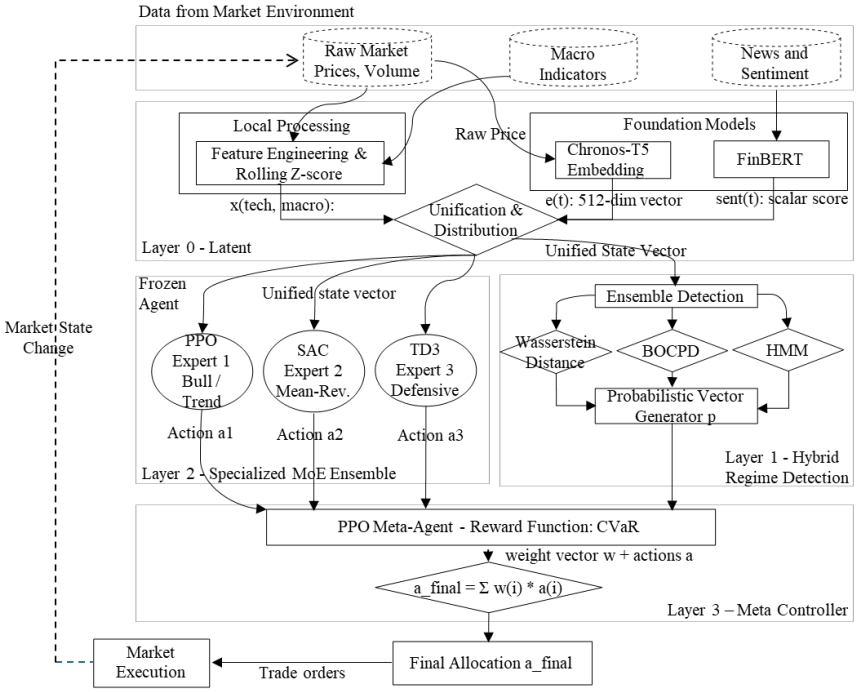


Fig. 1. Schematic representation of the information flow in the adaptive HDRL architecture.

The direct online integration of massive foundation models into a high-frequency DRL training loop introduces unacceptable inference latency and exponentially increases computational resource requirements [2, 22]. The HDRL architecture circumvents this bottleneck by isolating feature extraction into an asynchronous, offline Layer 0, a design pattern necessitated by the practical constraints of real-time financial decision-making [12].

Layer 0 utilizes the Amazon Chronos-T5-Small model to process sequential price data. By tokenizing continuous numerical values into a discrete vocabulary via scaling and quantization, Chronos processes temporal data through a standard T5 transformer. The architecture intercepts the network's final hidden states to extract dense embeddings that capture deep, non-linear temporal dependencies and complex seasonality without domain-specific fine-tuning.

To rigorously prevent data leakage (look-ahead bias), the extraction relies strictly on causal sliding windows. For any time step t, the embedding vector e_t is generated using only the sequence of observations from the preceding *N* periods:

$$e_t = \text{ChronosEncoder}(o_{t-N+1:t}) \tag{1}$$

where o represents the OHLCV sequence. Prior to state aggregation, this raw data is expanded into 348 technical and macroeconomic features, which undergo a strictly causal 252-day rolling Z-score normalization to ensure stationarity and mitigate distribution shifts. Optionally, NLP models such as FinBERT transform textual news into classification probabilities, forming a scalar sentiment indicator $sent_t$ [3]. The ultimate output of Layer 0 is a unified multi-modal state vector $z_t = [x_t, e_t, sent_t]$, which integrates stationary technical and macroeconomic indicators, latent embeddings extracted from foundation models, and sentiment probabilities derived from financial news.

3.3 Layer 1: Hybrid Regime Detector

Recognizing that no single algorithm is universally optimal for regime detection, Layer 1 implements a hybrid ensemble detector that integrates three complementary methodologies to form a continuous probabilistic belief state.

- **Hidden Markov Model (HMM):** Utilizing Gaussian Mixture Models (GMM) for emission probabilities, the HMM is optimized iteratively to provide a stable state probability vector $h_t = [h_{t,bull}, h_{t,bear}, h_{t,sideways}]$. However, HMMs are structurally inert when confronted with sudden market shocks, as they often lag the reality of a regime switch [23].
- **Bayesian Online Changepoint Detection (BOCPD):** To compensate for HMM latency, BOCPD recursively updates the probability distribution over the "runlength" (time since the last structural break). Transitions are governed by a hazard function $H(\tau)$, where the posterior probability that the runlength has collapsed signals an immediate structural fracture.
- **Wasserstein Distance (Optimal Transport):** The most critical innovation in Layer 1 is the application of the Wasserstein distance (Earth Mover's Distance). Unlike standard divergence metrics, the 1-Wasserstein distance measures the geometric cost required to transform one probability distribution into another [17]. It computes the distance between the empirical return distribution in the current window Q_t and historical reference distributions R_i:

$$W_1(Q, R) = (1/n)\Sigma_{k=1}^{n}\left|x_{(k)} - y_{(k)}\right| \tag{2}$$

This metric is uniquely sensitive to changes in higher moments (skewness and kurtosis) that precede market crashes. Distances are inverted and normalized to produce a geometry-based probability vector w_t.

The final output is a unified probabilistic belief state p_t, generated through a convex linear combination:

$$p_t = \lambda_t \cdot h_t + (1 - \lambda_t) \cdot w_t \tag{3}$$

The dynamic weighting hyperparameter $\lambda_t \in [0,1]$ balances historical stability and geometric sensitivity. Crucially, the weighting hyperparameter λ_t dynamically modulates the reliance between historical HMM stability and instantaneous geometric deformation. Let $c_t \in [0,1]$ denote the posterior probability emitted by BOCPD indicating a structural break. The coupling parameter is calculated operationally as $\lambda_t = \lambda_0 \cdot (1 - c_t)$, where λ_0 is the baseline steady-state weight empirically set to $\lambda_0 = 0.7$. Consequently, when BOCPD detects a severe break ($c_t \to 1$), λ_t is mathematically suppressed ($\lambda_t \to 0$), forcing the system to rely strictly on the ultra-fast Wasserstein signal.

3.4 Layer 2: Specialized Mixture of Experts (MoE)

Layer 2 houses the core trading intelligence, structured as a heterogeneous ensemble of k=3 parallel, mutually independent DRL agents. The fundamental reinforcement learning algorithm of each expert is mathematically paired to the stochastic nature of its assigned regime:

- Trend-Following (Bull) Expert: Utilizes Proximal Policy Optimization (PPO). Its clipped surrogate objective ensures stable, monotonic convergence, making it exceptionally suited for trending environments with clear directional signals.
- Mean-Reversion (Sideways) Expert: Utilizes Soft Actor-Critic (SAC). Sideways regimes abound with market noise. SAC integrates a maximum entropy framework, directly incentivizing policy stochasticity to maintain continuous exploration and extract alpha from oscillatory ranges.
- Defensive (Crisis) Expert: Utilizes Twin Delayed DDPG (TD3). Designed to minimize value overestimation bias through clipped double-Q learning, TD3 ensures robust defensive positioning during sudden structural breakdowns.

To systematically mitigate catastrophic forgetting, training is strictly isolated. The architecture utilizes Regime-Specific Experience Replay (RSER) buffers, ensuring that each agent trains exclusively on data mathematically verified as belonging to its target regime. This setup configures the L1 layer not as a standalone system, but as a rich signal vector for the broader HDRL mosaic. To resolve the circular dependency between regime labeling and calibration during initialization, an explicit two-phase procedure is implemented offline prior to expert training (detailed in Section 3.7). Once convergence is achieved, the neural network parameters are permanently frozen, serving as stationary mathematical functions mapping states to regime-specific action proposals (a_i).

3.5 Layer 3: Meta-Controller (Risk-Sensitive Gating Network)

The apex of the hierarchy is the Layer 3 Meta-Controller, an active PPO agent dedicated to the dynamic orchestration of the underlying frozen experts [8]. It does not output asset allocations directly; instead, it generates a gating vector $w_t = [w_1, w_2, w_3]$ dictating proportional capital distribution.

The Meta-Controller ingests a composite state concatenating the latent embeddings (z_t), the probabilistic belief state (p_t), and the expert action proposals (a_i). The final executable portfolio allocation a_{final} is computed via a continuous soft-switching mechanism:

$$a_{final} = \sum_{i=1}^{k} w_i \cdot a_i \tag{4}$$

This weighted aggregation smooths transitions, drastically reducing portfolio turnover and slippage. Furthermore, it has the potential to facilitate zero-shot adaptation. If Layer 1 identifies a novel hybrid market state, the Meta-Controller interpolates the weights to synthesize a calibrated strategy without retraining base models, a critical requirement for enterprise-grade deployment.

3.6 Hierarchical Reward Design

A profound structural vulnerability in existing DRL research is the reliance on unified reward functions that attempt to balance profit-seeking with risk management, leading to unstable gradients and reactive risk treatment [13]. The HDRL architecture resolves this through a strict hierarchical bifurcation of reward topologies.

The frozen experts in Layer 2 optimize the Differential Sharpe Ratio (DSR). The DSR isolates the marginal contribution of the current return to the overall Sharpe ratio, enabling efficient stochastic gradient ascent and addressing reward sparsity challenges [18]. Utilizing exponential moving averages of returns (A_{t-1}) and squared returns (B_{t-1}), the DSR update is computed recursively to weight recent returns more heavily [20]:

$$DSR_t = \frac{B_{t-1}\Delta A_t - \frac{1}{2}A_{t-1}\Delta B_t}{\left(B_{t-1} - A_{t-1}^2\right)^{3/2}} \tag{5}$$

Conversely, the Layer 3 Meta-Controller operates as an enterprise risk manager optimizing the Conditional Value at Risk (CVaR). While DSR controls variance symmetrically, CVaR explicitly targets severe tail risks by calculating the expected loss exceeding a predefined Value at Risk (VaR) threshold α:

$$CVaR_\alpha(X) = E[\, X \mid X \leq VaR_\alpha(X) \,] \tag{6}$$

By penalizing the Meta-Controller based on the aggregate portfolio's CVaR, the system is mathematically coerced into a globally defensive posture. This bifurcated design ensures that local experts aggressively pursue alpha (L2 DSR) while the meta-controller guarantees global capital preservation during systemic shocks (L3 CVaR).

3.7 Reproducibility and Implementation Details

To support independent reproducibility and formalize the intended operational parameters of the proposed architecture, we explicitly outline the core methodological constraints and protocols:

- Portfolio Constraints and Action Space: The portfolio action space is strictly bounded by the projection constraints $w_i \in [-1,1]$ and $\sum_{i=1}^{n}|w_i| \leq 1.0$. This architecture allows for controlled directional exposure while mathematically preventing uncollateralized leverage.
- Two-Phase RSER Initialization: To resolve the circular dependency inherent in L2 initialization, the RSER buffers (capacity of 10^6 transitions) are populated using an explicit unsupervised procedure. First, seed regime labels are generated utilizing the k-means algorithm on the 504-day training window. Second, these seed labels initialize the parameters of a three-state Gaussian HMM, which are iteratively refined using the Expectation-Maximization (Baum-Welch) algorithm, augmented by a deterministic risk-return heuristic to prevent label switching. Only after this stabilization are the finalized labels used to construct the optimal transport reference distributions (R_i) and segregate the RSER buffers.
- Expert Training Protocols: The heterogeneous expert neural networks are configured utilizing established continuous-control hyperparameters to ensure stable convergence. The Bull Expert (PPO) utilizes a learning rate of 3×10^{-4} and a batch size of 64. The Sideways Expert (SAC) operates with a learning rate of 10^{-4}. The Crisis Expert (TD3) utilizes an actor learning rate of 10^{-4}, a critic learning rate of 10^{-3}, and target policy noise of 0.2. All other non-specified hyperparameters (such as discount factor γ and soft-update coefficient τ) strictly follow the default configurations of the original algorithms [10, 11, 27]
- Algorithmic Approximations: To ensure computational tractability during high-frequency execution, the L1 Wasserstein distance geometric computations rely on Sinkhorn approximations. Furthermore, the L3 Meta-Controller computes the global CVaR penalty utilizing a 95% confidence interval ($\alpha = 0.05$).

4 Experimental Framework and Preliminary Results

To empirically validate the theoretical innovations, an experimental framework is designed to quantify resilience, zero-shot adaptation, and Sim2Real gap bridging.

4.1 Research Hypotheses

- H0 (Main Hypothesis): The HDRL architecture will yield statistically significant improvements in out-of-sample risk-adjusted returns (Sharpe and Sortino ratios) compared to monolithic DRL baselines.
- H1 (Detection Efficiency): The Wasserstein-augmented L1 hybrid detector will identify structural breaks with significantly reduced latency compared to isolated HMM or BOCPD mechanisms.

- H2 (Prevention of Forgetting): Frozen L2 experts, utilizing Regime-Specific Experience Replay (RSER), will demonstrate zero competency degradation across alternating regimes, effectively mitigating "parameter contamination" found in regime-agnostic models.
- H3 (Adaptation Speed): The Layer 3 continuous soft-switching mechanism will facilitate rapid zero-shot adaptation, drastically reducing stabilization time compared to online retraining models.
- H4 (Tail-Risk Robustness): The integration of CVaR optimization at the meta-level will produce a statistically significant reduction in both Maximum Drawdown (MDD) and realized tail losses during periods of high volatility.

4.2 Data and Simulation Setup

The framework utilizes daily historical data from a 15-ticker universe, encompassing eight specific Sector ETFs (Technology, Healthcare, Energy, Financials, Industrials, Utilities, Consumer Discretionary, and Materials), Diversification and Hedge assets spanning Real Estate, Emerging Markets, Treasury Bonds, and Commodities (VNQ, EEM, TLT, GLD, DBC), and Benchmark instruments (SPY, QQQ). This universe is evaluated alongside key macroeconomic indicators including the CBOE Volatility Index (VIX), US Dollar Index (DXY), and Federal Funds Rate. The 4,876 trading days spanning 2006–2026 are partitioned into Training (2006–2019), Validation (2020–2021), and Testing (2022–2026) sets. To prevent information leakage due to serial correlation, a 92-day embargo period is enforced between non-contiguous data splits [16].

To address the Sim2Real gap, the simulation enforces a Volume Share Slippage Model, 10bps transaction costs, and parametric market impact models to realistically reflect liquidity constraints. Out-of-Distribution (OOD) stress testing utilizes cGANs to synthesize extreme market trajectories.

4.3 Empirical Results: Layer 0 Implementation

The Layer 0 (L0) foundation model extraction pipeline was fully implemented and rigorously validated on the defined dataset. To prevent the out-of-distribution (OOD) problem commonly observed in Reinforcement Learning (RL), each window represents a strict 90-day historical lookback period across the 15 instruments, structurally prohibiting any look-ahead bias and ensuring absolute zero data leakage.

Stationarity and Distribution Shift: To address the inherent non-stationarity of financial time series and prevent the out-of-distribution (OOD) problem in RL, a rolling Z-score normalization utilizing a 252-day causal window was applied to all 348 engineered features. Rigorous statistical validation was performed using the Augmented Dickey-Fuller (ADF) test across the entire feature space. The tests confirmed that 100% of the evaluated features (348 out of 348) strongly reject the null hypothesis of unit root existence ($p < 0.001$), ensuring strict I(0) stationarity. Furthermore, the stability of the state space was evaluated by measuring the distribution shift between the training (3,371 days) and out-of-sample sets (validation and test, totaling 1,505 days). Using a

composite distribution shift index—which aggregates both mean deviation (Cohen's d approximation) and variance ratio shifts—the average shift decreased substantially from 0.6787 (raw features) to 0.2388 (scaled features). This 64.82% reduction effectively neutralizes long-term market drift and isolates true topological market changes, providing the subsequent HDRL layers with a highly stable and robust observation space.

Chronos-T5 Embedding Quality: The pre-trained Amazon Chronos-T5-Small model processed 69,825 causal sliding windows purely offline, ensuring absolute zero data leakage. The extraction successfully yielded high-density 512-dimensional latent vectors per instrument with zero missing or infinite values. Principal Component Analysis (PCA) demonstrated the high compressibility of the latent space: only 124 principal components are required to retain 90% of the total global variance, confirming a well-structured representation space that effectively mitigates the curse of dimensionality (as shown in Fig. 2).

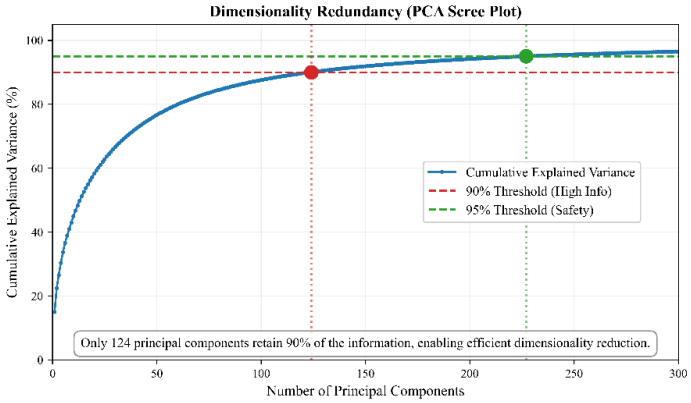


Fig. 2. Dimensionality Redundancy. Principal Component Analysis (PCA) global variance decomposition demonstrating high compressibility of the extracted embeddings.

Regime Separation (Latent Clustering): To validate whether the extracted embeddings capture underlying macroeconomic environments without prior supervision, dimensionality reduction and clustering were applied. A UMAP projection of the global market latent structure (Fig. 3) clearly segregates distinct market regimes, topologically isolating high-volatility periods—such as the 2008 Financial Crisis, the COVID-19 pandemic, and the 2022 Inflation Shock—from low-volatility bull markets.

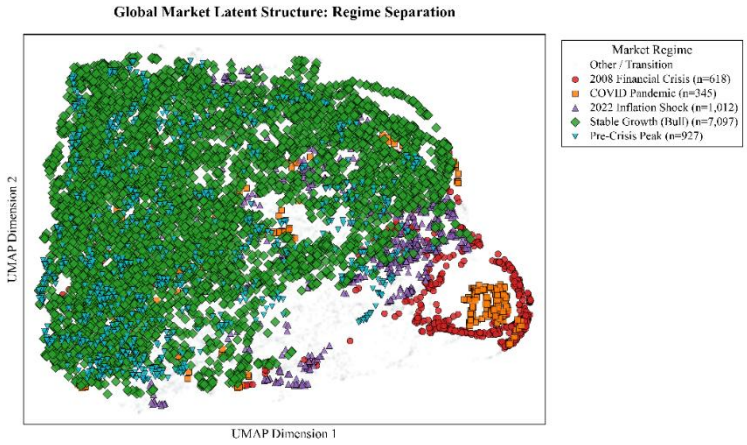


Fig. 3. Global Market Latent Structure (UMAP). Projection of embeddings clearly separates distinct macroeconomic regimes without any prior supervision, demonstrating robust embedding discriminability.

However, we explicitly acknowledge that retrospective, unsupervised clustering of historical data is not equivalent to the prospective, autonomous identification of unlabeled future regimes. Actual prospective detection performance requires the active implementation of the L1 hybrid detector. Ultimately, these quantitative and visual results confirm that L0 extracts robust, regime-discriminative representations despite the inherently noisy nature of financial time series, establishing a necessary, though not sufficient, precondition for detection, thereby providing an optimal observation space for the subsequent HDRL layers.

4.4 Baseline Models

The fully developed HDRL architecture will be benchmarked against three distinct control models:

- Passive Benchmark (Buy-and-Hold): A capitalization-weighted allocation to the SPY ETF, serving as the market baseline.
- Monolithic PPO Agent: A standard DRL agent trained across the entire unified dataset utilizing a continuous action space and a standard unified reward function, serving as the primary control to highlight catastrophic forgetting.
- HMM-RL Hybrid: A contemporary approach utilizing a standard HMM to feed a discrete state signal directly into a singular RL policy network, providing a baseline to assess the latency improvements of the L1 Wasserstein integration.

5 Discussion and Future Work

Building upon the robust, stationary latent representations successfully extracted and validated in the foundational Layer 0, the subsequent deployment of the complete Adaptive HDRL architecture is projected to yield substantial improvements in organizational resilience. This section outlines the anticipated outcomes for the core research hypotheses and details the ablation study designed to rigorously validate the necessity of each architectural component.

5.1 Proposed Evaluation Protocol for L1-L3

To rigorously validate the proposed architecture in future implementations, the comprehensive evaluation protocol will employ out-of-sample risk-adjusted performance metrics (including annualized Sharpe ratio, Sortino ratio, and Maximum Drawdown) alongside formal statistical testing. Regarding the established hypotheses, the integration of the Wasserstein distance within the L1 Hybrid Detector (H1) is anticipated to fundamentally alter system reaction times. By calculating the geometric deformation of the entire distribution, it is designed to flag the fattening of the left tail that reliably precedes major sell-offs, triggering the Meta-Controller hours or days before monolithic baselines recognize the systemic danger. Furthermore, the architectural decision to freeze the neural weights of the L2 MoE post-training (H2) is theoretically designed to mitigate catastrophic forgetting. While monolithic agents are expected to exhibit severe model decay (estimated 20–30% degradation) when transitioning between regimes, the specialized experts are projected to maintain their maximum theoretical efficiency within isolated domains.

Driven by the continuous soft-switching mechanism, the L3 Meta-Controller aims to facilitate zero-shot adaptation (H3), with a target of stabilizing portfolio performance within an estimated 5 to 10 trading days following a regime detection, compared to 20 to 30 days for models requiring online retraining. Finally, during simulated cGAN OOD stress tests (H4), the intense penalties of the CVaR objective are expected to preemptively reroute capital to defensive experts or cash equivalents. This dynamic routing is anticipated to restrict Maximum Drawdown to within acceptable institutional limits (projected under 15%), contrasting sharply with monolithic agents that naively extrapolate training data into unseen territory.

5.2 Proposed Ablation Study and Robustness

As part of our planned future research, an extensive ablation study will be conducted to rigorously isolate and evaluate the contribution of individual architectural components once the full system is empirically tested. The performance of the complete architecture will be compared against the following structural variations using paired Wilcoxon signed-rank tests to ensure statistical significance:

- HDRL w/o L1 (No Explicit Regime Detection): In this variant, the L1 Hybrid Detector is removed, forcing the L3 Meta-Controller to infer market regimes implicitly.

It is anticipated that this implicit inference will suffer from high dimensionality noise, as regime-agnostic models are known to "pollute" learned parameters by mixing heterogeneous market states. This is expected to severely degrade the zero-shot adaptation speed and increase portfolio turnover costs.

- HDRL w/o L2 MoE (Unified Executing Agent): The heterogeneous experts are replaced with a single, unfrozen policy network. This ablation is expected to immediately re-introduce catastrophic forgetting. By removing modular algorithmic heterogeneity and weight-freezing, the agent will likely lose its competency in previously learned regimes when exposed to new market dynamics, proving that task decomposition is essential for stable learning.
- HDRL w/o CVaR (DSR-only Meta-Controller): The L3 reward function is altered to optimize the Differential Sharpe Ratio (DSR) instead of CVaR. While DSR is efficient for alpha generation, it does not distinguish between upside potential and downside risk. CVaR, as a coherent risk measure, is necessary to explicitly penalize rare but severe tail losses. This configuration is expected to result in catastrophic organizational failure during GAN-generated stress tests, proving that proactive tail-risk optimization must occur at the meta-level to ensure resilience against "Black Swan" events.

5.3 Methodological Limitations

Despite the encouraging preliminary validation of the foundational L0 pipeline, the proposed HDRL architecture possesses several methodological constraints that must be navigated prior to live deployment. First, while causal sliding windows are strictly enforced, the offline feature extraction process and the dynamic scaling of reference distributions remain susceptible to subtle forms of look-ahead bias if macroeconomic indicators and price sequences are not perfectly time-aligned. Second, unsupervised dimensionality reduction techniques like UMAP are inherently stochastic and highly sensitive to hyperparameter tuning; consequently, they provide retrospective validation of embedding discriminability rather than functioning as prospective, real-time online detectors. Third, the downstream performance of the hierarchical control loop is highly sensitive to the a priori hyperparameters defining the market regimes—such as the predetermined number of specialized experts (k=3) and the fixed length of the optimal transport reference windows—which require continuous calibration to prevent cluster degeneration. Furthermore, financial markets exhibit complex, residual non-stationarities, including sudden breakdowns in cross-asset correlations and multi-modal regime drift, which may not be fully encapsulated by the current feature space. Finally, the continuous real-time computation of the 1-Wasserstein distance introduces substantial computational latency, currently rendering the architecture unsuitable for ultra-high-frequency trading (UHFT) environments without further algorithmic approximations (e.g., Sinkhorn distances). Addressing these computational and structural bottlenecks constitutes a primary focus for future iterations of this work.

6 Conclusion

This paper presents a mathematically rigorous Adaptive HDRL architecture designed to resolve the vulnerabilities of monolithic trading systems in non-stationary environments. By systematically deconstructing sequential decision-making, the framework synthesizes time-series foundation models for bias-free latent extraction, optimal transport for rapid geometric regime detection, and a heterogeneous Mixture of Experts with frozen parameters to mitigate catastrophic forgetting. The introduction of a bifurcated reward topology—separating local DSR alpha generation from global CVaR tail-risk protection— proposes a theoretical, modular blueprint aimed at metabolizing uncertainty, with the potential to facilitate zero-shot adaptation.

Empirical validation of the foundational layer across 4,876 trading days demonstrates highly coherent, stationary regime clustering that reduces data distribution shifts by 64.82%. While constrained by the computational overhead of real-time geometric calculations (rendering it unsuitable for ultra-high-frequency trading), future research will focus on algorithmic optimization via Sinkhorn approximations. Additionally, subsequent iterations will explore the expansion of the L2 MoE to alternative asset classes and the integration of Large Language Models (LLMs) to provide autonomous, natural-language explainability (XAI) for regulatory compliance.

Acknowledgments. AI-Use Disclosure Statement: During the preparation of this paper, the authors utilized a Large Language Model (LLM) exclusively for AI-assisted copy editing, which included language polishing, stylistic refinement, and grammatical correction of human-generated text. The authors have thoroughly reviewed and edited the paper, and they assume full and exclusive responsibility for the scientific integrity, accuracy, and content of this publication.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Adams, R.P., MacKay, D.J.C.: Bayesian online changepoint detection. arXiv Preprint arXiv:0710.3742. 1–7 (2007). https://doi.org/10.48550/arXiv.0710.3742.
2. Ansari, A.F. et al.: Chronos: learning the language of time series. arXiv Preprint arXiv:2403.07815. 1–43 (2024). https://doi.org/10.48550/arXiv.2403.07815.
3. Araci, D.: FinBERT: financial sentiment analysis with pre-trained language models. arXiv Preprint arXiv:1908.10063. 1–11 (2019). https://doi.org/10.48550/arXiv.1908.10063.
4. Arjovsky, M. et al.: Wasserstein generative adversarial networks. In: Proceedings of the 34th International Conference on Machine Learning. pp. 214–223 PMLR (2017).
5. Baum, L.E., Petrie, T.: Statistical inference for probabilistic functions of finite state Markov chains. Annals of Mathematical Statistics. 37, 6, 1554–1563 (1966). https://doi.org/10.1214/aoms/1177699147.
6. Cetrulo, A.L.S. et al.: Machine learning for financial prediction under regime change using technical analysis: a systematic review. International Journal of Interactive Multimedia and Artificial Intelligence. 9, 1, 137–148 (2024). https://doi.org/10.9781/ijimai.2023.06.003.

7. Chandak, Y.: Reinforcement learning for non-stationary problems. Ph.D. Thesis, University of Massachusetts Amherst (2022).
8. Chen, J. et al.: MARS: A Meta-Adaptive Reinforcement Learning Framework for Risk-Aware Multi-Agent Portfolio Management. arXiv Preprint arXiv:2508.01173. (2025). https://doi.org/10.48550/arXiv.2508.01173.
9. Dakalbab, F. et al.: Artificial intelligence techniques in financial trading: A systematic literature review. Journal of King Saud University - Computer and Information Sciences. 36, 3, 102015 (2024). https://doi.org/10.1016/j.jksuci.2024.102015.
10. Fujimoto, S. et al.: Addressing function approximation error in actor-critic methods. In: Proceedings of the 35th International Conference on Machine Learning. pp. 1587–1596 PMLR (2018).
11. Haarnoja, T. et al.: Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor. In: Proceedings of the 35th International Conference on Machine Learning. pp. 1861–1870 PMLR (2018).
12. Hambly, B. et al.: Recent advances in reinforcement learning in finance. Mathematical Finance. 33, 3, 437–503 (2023). https://doi.org/10.1111/mafi.12382.
13. Hoque, M.R. et al.: Reinforcement Learning in Financial Decision Making: A Systematic Review of Performance, Challenges, and Implementation Strategies. arXiv Preprint arXiv:2512.10913. (2025). https://doi.org/10.48550/arXiv.2512.10913.
14. Horvath, B. et al.: Clustering market regimes using the Wasserstein distance. Journal of Computational Finance. 1–37 (2024). https://doi.org/10.21314/JCF.2024.005.
15. Liu, X.-Y. et al.: FinRL: a deep reinforcement learning library for automated stock trading in quantitative finance. arXiv Preprint arXiv:2011.09607. 1–12 (2020). https://doi.org/10.48550/arXiv.2011.09607.
16. López de Prado, M.M.: Advances in financial machine learning. Wiley, Hoboken, New Jersey (2018).
17. Mc Greevy, J. et al.: Detecting multivariate market regimes via clustering algorithms. SSRN Electronic Journal. 1–97 (2024). https://doi.org/10.2139/ssrn.4758243.
18. Moody, J. et al.: Performance functions and reinforcement learning for trading systems and portfolios. Journal of Forecasting. 17, 5–6, 441–470 (1998). https://doi.org/10.1002/(SICI)1099-131X(1998090)17:5/6%3C441::AID-FOR707%3E3.0.CO;2-%23.
19. Mu, S., Lin, S.: A Comprehensive Survey of Mixture-of-Experts: Algorithms, Theory, and Applications. arXiv Preprint arXiv:2503.07137. (2026). https://doi.org/10.48550/arXiv.2503.07137.
20. Niu, H. et al.: MetaTrader: An Reinforcement Learning Approach Integrating Diverse Policies for Portfolio Optimization. In: Proceedings of the 31st ACM International Conference on Information & Knowledge Management. pp. 1573–1583 Association for Computing Machinery, New York, NY, USA (2022). https://doi.org/10.1145/3511808.3557363.
21. Pan, C. et al.: A Survey of Continual Reinforcement Learning. arXiv Preprint arXiv:2506.21872. (2025). https://doi.org/10.48550/arXiv.2506.21872.
22. Pippas, N. et al.: The evolution of reinforcement learning in quantitative finance: a survey. ACM Computing Surveys. 57, 11, 1–51 (2025). https://doi.org/10.1145/3733714.
23. Pomorski, P.: Construction of effective regime-switching portfolios using a combination of machine learning and traditional approaches. Ph.D. Thesis, University College London (2024).
24. Pun, C.S. et al.: Data-Driven Distributionally Robust CVaR Portfolio Optimization Under A Regime-Switching Ambiguity Set. Manufacturing & Service Operations Management. 25, 5, 1779–1795 (2023). https://doi.org/10.1287/msom.2023.1229.

25. Rocchi, A.: Streamlined Hierarchical Reinforcement Learning for Algorithmic Trading: Architecture Simplification and Empirical Validation. SSRN Electronic Journal. (2025). https://doi.org/10.2139/ssrn.5458097.
26. Saqur, R.: Dynamic adaptation to financial market regimes through filtering and reward engineering of large language models. Ph.D. Thesis, University of Toronto (2025).
27. Schulman, J. et al.: Proximal policy optimization algorithms. arXiv Preprint arXiv:1707.06347. 1–12 (2017). https://doi.org/10.48550/arXiv.1707.06347.
28. Sun, R. et al.: A novel multi-agent dynamic portfolio optimization learning system based on hierarchical deep reinforcement learning. Complex & Intelligent Systems. 11, 7, 311 (2025). https://doi.org/10.1007/s40747-025-01884-y.
29. Sutton, R.S., Barto, A.G.: Reinforcement learning: an introduction. The MIT Press, Cambridge, Massachusetts (2018).
30. Tamar, A. et al.: Optimizing the CVaR via Sampling. Proceedings of the AAAI Conference on Artificial Intelligence. 29, 1, (2015). https://doi.org/10.1609/aaai.v29i1.9561.
31. Tsaknaki, I.-Y. et al.: Online learning of order flow and market impact with Bayesian change-point detection methods. Quantitative Finance. 25, 2, 307–322 (2025). https://doi.org/10.1080/14697688.2024.2337300.
32. Verma, A.K. et al.: Regime-based portfolio allocation using Hidden Markov Models and reinforcement learning. SSRN Electronic Journal. 1–9 (2025). https://doi.org/10.2139/ssrn.5785443.
33. Zhang, Y. et al.: RegimeFolio: A Regime Aware ML System for Sectoral Portfolio Optimization in Dynamic Markets. IEEE Access. 13, 184722–184744 (2025). https://doi.org/10.1109/Access.2025.3624822.

Original Research Article *https://doi.org/10.5281/zenodo.23479014*

Exploring Differences in Generative AI Adoption Across European Countries Using Structural Equation Modeling

Miljana Luković[1[0009-0002-8880-6182]], Milica Maričić[2[0000-0003-0441-9899]] and Miloš Danilović[3[0000-0003-3712-9863]]

[1,2,3] University of Belgrade, Faculty of Organizational Sciences
[1]miljana.lukovic@fon.bg.ac.rs, [2]milica.maricic@fon.bg.ac.rs,
[3]milos.danilovic@fon.bg.ac.rs

Abstract. The rapid expansion of generative artificial intelligence (AI) models and the proliferation of generative AI tools have underscored the need to understand adoption patterns among younger generations better. This study employs recent Eurostat data on the usage of AI and statistical methods to identify key variables associated with generative AI use among Generation Z in Europe and Balkans, focusing on barriers and usage contexts. Structural equation modelling (SEM) is used to determine which barriers significantly affect generative AI adoption in this cohort. The analysis reveals that lack of need and lack of awareness are the primary barriers exerting a statistically significant negative effect on generative AI usage. In contrast, lack of skills and privacy concerns do not demonstrate a significant influence. These results indicate that, for Generation Z, adoption is primarily driven by perceptions of relevance and familiarity rather than by technical or security-related constraints. Descriptive statistics further show that generative AI is predominantly used for private and educational purposes, with professional use being less common. This paper contributes empirical evidence from a Generation Z perspective to the literature on generative AI adoption. It underscores the importance of raising awareness and demonstrating practical value as strategies for enhancing adoption.

Keywords: Generative AI, Generation Z, SEM analysis, trend analysis

1 Introduction

In the last decade, exceptional advances in the domain of machine learning resulted in the rapid evolution of artificial intelligence (AI). Today, there are distinct types of AI available on the market; the most well-known are generative AIs. Generative AIs are created with complex algorithms that can produce new data and syntaxes based on the user's prompt. The simplest definition describes generative AI as [11] "a group of AI algorithms and models that are capable of producing new content, including texts, images, videos and problem-solving strategies, with human-like creativity and adaptability". The evolution of generative AI gives a better perspective on how much every innovation in this domain rapidly enhances the capabilities of generative AI. The researchers Ran He and his team have created a historical classification of the evolution

of generative AI and defined four main stages [12]: rule-based generative systems, model-based generative algorithms, deep generative methodologies, and foundation models. All available generative AI models are in the foundation model stage; in simple terms, these models are all called generative pretrained transformers, or GPTs for short. The main generative AI models used across various platforms as of February 2026 include [15]: *Gemini* 3.1 Pro by Google DeepMind; *Claude Opus* 4.6 and *Claude Sonnet* 4.6 by Anthropic; *GPT-5.3 Codex* by OpenAI; *Grok* 4.20 by xAI; and *Qwen 3.5* by Alibaba.

The traffic data collected in November of 2025 shows that the most visited generative AI platforms were [19]: *ChatGPT* (5.6 billion visits), *Gemini* (650 million monthly users), *DeepSeek* (328.2 million visits), *Perplexity* (239.97 million visits), *Claude* (185.93 million visits), *Character.AI* (141.1 million visits), *Microsoft Copilot* (110.32 million visits), and *QuillBot* (59.06 million visits). While the most downloaded generative AI apps in March 2026 were [19, 20]: *ChatGPT* (40.52% of total downloads), *DeepSeek* (17.59%), *Google Gemini* (9.6%), *Doubao* (8.89%), *another version of DeepSeek* (7.76%), *PixVerse* (6.19%), *Talkie* (4.68%), *Nova* (4.35%), *Microsoft Copilot* (2.83%), and *Character.AI* (2.81%).

New generative AI platforms are being created rapidly, often built on existing foundational models such as GPT-4 or Gemini. Their differentiation typically lies in data sources, system architecture, UI/UX design and functionality, and target users. Within the field of research focused on generative AI platforms, several tools are widely recognised and utilised by researchers and students for academic purposes. Notable examples include [9]: *Anara AI*, *Jenni AI*, *Research Rabbit*, *Elict, Semantic Scholar, Google Labs, Inciteful, Julius AI, Thesify,* and many others.

The International AI Safety Report, issued in February of 2026 [4], concluded that generative AI models are becoming more precise, reliable, and safer to use. A common question is: Is generative AI making our day-to-day lives easier? A research study published in January 2026 [8] has shown that adopting generative AI can reduce mental load and help individuals be more productive. Generative AI has transformed how individuals use, understand, and transform data, as well as how they express their creativity and logic. It can be said that the rise of generative AI has a similar effect to that of the rise of the Internet, in that it has drastically changed not only the everyday lives of individuals but also various industries and the entire job market in a short period of time. The McKinsey & Company 2025 survey [16] provides supporting evidence that, in the view of major companies, generative AI has shifted from an experimental technology to a widely adopted enterprise resource. Based on the survey report, 62% of organisations are testing and experimenting with the use of AI agents, while 23% of organisations are reporting that they are scaling the use of AI in their organisation's procedures.

As new technologies emerge in the global market, they face different challenges and barriers to adoption and implementation. Although it is common for sceptics and traditionalists to boycott new technology out of fear or a lack of understanding, with generative AI, the fear of being replaced by an algorithm has sparked global concern. This global concern has risen due to many factors, such as mass layoffs, academic and

scientific dishonesty, ethical concerns, private companies' data leaks, the general fear of dependency and other similar reasons [3, 12].

Recent studies have shown an increase in barriers to using generative AI, while multinational companies promote its use and boast of positive results. Senior employees express concern that the younger generations' skills will experience atrophy and unnecessary co-dependency on AI tools [12]. The job market data [10] shows a distinct pattern of massive layoffs in the IT sector, while the companies report an increase in profit. The leading CEOs have shifted the blame for mass layoffs and employee displeasure to the recent adoption and implementation of generative AI in the workplace [10]. The industry reports do not support these narratives, due to the fact that most organisations using generative AI agents are still in the early stages of implementation and that only 39% of them reported measurable financial impact [16]. Universities and other educational institutions have been cautious about the use of generative AI in academic work. A study done in 2025 [4] has shown that over 87% of educational institutions use AI detection software. AI detection software was created to stop academic and research dishonest behaviours, but studies show that the number of false positive results is starting to rise [1]. Namely, individuals who are native speakers and have the highest level of writing have been targeted for having "too clean a writing style."

The main research question of this paper is: How and for what reason are members of Generation Z using generative AI across European countries? To answer this question, this study uses publicly available data provided by Eurostat [6, 7] collected from two datasets. The first set focuses on the purpose for which Generation Z uses generative AI, while the second observes why Generation Z is not using generative AI.

After the introduction, the paper contains five sections. Section 2 provides a brief literature review of recent papers on factors that influence and hinder generative AI adoption. Section 3 showcases the SEM methodology and the data used in the analysis. Section 4 focuses on the analysis results. Section 5 is reserved for the discussion, followed by Section 6, which outlines the conclusion.

2 Factors that affect the use of generative AI

In this section, the authors have conducted a brief literature review of papers published in 2026 which explored the drivers and/or barriers of generative AI adoption. Particular emphasis was placed on studies published in 2026 mainly due to the rapidly evolving nature of generative AI technologies and platforms, as well as usage patterns. Earlier studies were excluded, even though they could be considered primarily for theoretical and conceptual grounding, in favour of more recent empirical studies to reflect better the contemporary adoption environment associated with the large-scale diffusion of generative AI tools. The selection criteria for a paper to be used in this review were: the paper must have been published in 2026, in a Scopus journal, must be a full paper written in English, and must contain quantitative results.

To provide a structured overview of recent empirical findings related to generative AI adoption, the selected studies were compared across several dimensions, including the level of analysis, theoretical approach, methodology, adoption drivers, and reported

barriers. The purpose of this comparison is to identify recurring patterns, contradictions, and research gaps within the rapidly evolving field of generative AI adoption research.

The eight papers identified are shown in Table 1 through six different criteria: *Level of analysis*: individual, organisational or cross-country; *The scientific approach that the authors used*; *How was the data gathered and what methodology was employed*; *What was the focus of the analysis; What key drivers in adoption of generative AI did the authors observe;* and *What key barriers in adoption of generative AI did the authors observe*.

Table 1. Overview of the eight papers identified in the bibliometric query

Ref	Level of Analysis	Approach	Data & Methodology	Focus of Analysis	Key Drivers	Key Barriers
[21]	Organizational	Multi dimensional taxonomy (strategy, risk, autonomy, governance)	Conceptual synthesis of academic and industry literature	Structuring GenAI use cases by complexity, risk, and implementation sequencing	Strategic alignment, value creation, internal operational efficiency	Governance gaps, risk escalation, lack of structured adoption pathways
[13]	Organizational	TOE framework	Mixed methods (qualitative and survey), SEM (n=304)	Organizational adoption drivers and barriers	Innovation potential, operational efficiency, organizational readiness, trust	Skill gaps, integration complexity, regulatory uncertainty
[15]	Individual	UGT + PMT (dual model)	Longitudinal survey, SEM (n=1271)	Psychological drivers and resistance mechanisms	Gratification (curiosity, enjoyment), AI self-efficacy	Fear of losing power, perceived threats, technostress
[24]	Individual (students)	TAM + TPB + SDT (integrated model)	Survey, CFA and SEM (n=517)	Cognitive, social, and motivational determinants of adoption	Perceived usefulness, ease of use, autonomy, competence, relatedness	Ethical concerns, lack of control, variability of social influence
[8]	Individual and Organizational	JD-R model and social cognitive theory	Survey, PLS-SEM (n≈130)	Impact of GenAI adoption on employee well-being	Positive attitudes, performance gains, team cohesion	Anxiety, social isolation, lack of training, role uncertainty

Ref	Level of Analysis	Approach	Data & Methodology	Focus of Analysis	Key Drivers	Key Barriers
[18]	Individual (faculty)	Exploratory empirical (no strict theoretical model)	Survey, descriptive statistics and non-parametric tests (n=77)	Perceptions and academic use practices	Usefulness in writing, idea generation, efficiency gains	Trust issues, plagiarism, lack of institutional policy
[17]	Cross-country	Exploratory comparative approach	Large-scale international survey	Cross-country and cross-disciplinary adoption patterns	Productivity, accessibility (e.g., translation), writing support	Inaccuracy, bias, lack of transparency, privacy risks
[22]	Organisational (System-level)	Utilitarian theory	Panel data (2010–2023), econometric modelling and text mining	Impact of GenAI adoption on ESG behaviour in supply chains	Efficiency gains, innovation, resource optimization	ESG opportunism, agency costs, weakened internal controls

**Note for the abbreviations*. TOE = Technology–Organization–Environment framework; UGT = Uses and Gratifications Theory; PMT = Protection Motivation Theory; TAM = Technology Acceptance Model; TPB = Theory of Planned Behaviour; SDT = Self-Determination Theory; JD-R = Job Demands–Resources model; SEM = Structural Equation Modelling; CFA = Confirmatory Factor Analysis; PLS-SEM = Partial Least Squares Structural Equation Modelling; ESG = Environmental, Social, and Governance.

The selected studies have consistently highlighted that generative AI adoption and implementation are new, complex, and multi-factor phenomena that reach beyond traditional rational decision-making models. The reviewed studies suggest that generative AI adoption cannot be explained solely by traditional technology acceptance perspectives but also by broader psychological, organisational, and contextual factors. Today, generative AI models are, in fact, a new evolutionary step in how people perceive data. While early research focused on emphasising the perceived usefulness and ease of use of AI models, the most recent findings highlight the growing importance of psychological and emotional factors, such as curiosity, enjoyment and even fear [15, 24]. A quite interesting correlation was observed: fear-related constructs, such as fear of missing out, have a dual effect on people. The fear of missing out acts as both a barrier and driver of generative AI adoption [15].

Even though there is a trend in the mass media to highlight concerns regarding the use of generative AI, the empirical evidence suggests otherwise. The results have shown that users frequently disregard concerns regarding their reliability, accuracy, and ethical implications and wholeheartedly adopt generative AI tools in their everyday, school, and work life [17, 18]. The data have shown that users of generative AI tools most often used them for writing and content generation. The same study also found that the least-selected reason for using generative AI was to support advanced analytical or critical-thinking activities [17]. Looking at adoption at the organisational level, companies have shown increasing recognition of generative AI, as it has proven to be a strategic investment that increases overall value and competitive advantage. Unfortunately, companies often struggle with insufficient readiness, lack of governance frameworks, and skill shortages, which impact the integration process, resulting in highly insufficient adoption of generative AI into companies' procedures and data [13, 21].

While most research papers focus on the adoption and implementation of generative AI, there is a growing trend of research exclusively examining its broader consequences. These research papers are putting the spotlight on employee well-being, unintended risks such as opportunistic behaviour, and governance challenges [8, 22]. The main collective conclusion is that there is a distinction when it comes to adopting and researching the effects of generative AI. Researchers are using more holistic approaches for adopting generative AI while also taking into consideration the social and psychological factors that using generative AI has on users[13, 24].

The analysed recent studies provide a valuable insight into generative AI adoption, but several important gaps remain. The selected existing empirical studies are conducted with individual-level survey data that the authors collected within a single organisation, university, or country [14, 16, 19, 25]. Looking at cross-country analyses based strictly on publicly available statistical datasets, the results remain quite limited. This paper addresses this gap by combining Eurostat datasets with SEM analysis to explore the factors influencing generative AI usage and reported barriers among Generation Z users.

3 Methodology

This paper uses aggregated data published by Eurostat on the topic of the adoption of generative AI, focusing on the opinions and stance of individuals between 16 and 24 years of age known as Generation Z. For this paper, two Eurostat datasets were combined to explore both how Generation Z uses generative AI, as well as the reasons why some individuals from the same cohort do not use generative AI. In order to examine the relationships between these opposite aggregated data, the study applies the Structural Equation Modelling or SEM analysis. The section is structured into two main parts: first, database creation and analysis are explained, followed by the methodology for constructing the proposed SEM model.

The available Eurostat datasets do not contain publicly available individual-level microdata; instead, they report only aggregated data presented as country-level percentages. These percentages represent the share of respondents selecting a given response category within each country, and, because individual-level microdata are unavailable, this study uses country-level data as the unit of analysis. The analysis therefore examines associations among aggregated indicators and does not make direct individual-level behavioural or causal claims about Generation Z users.

3.1 Data origin and database creation logic

The paper's main research question: How do distinct factors shape both the usage and non-usage of generative AI among Generation Z across European countries? has guided the process of data exploration and selection. The newest data published by Eurostat as a result of their extensive and comprehensive survey provides the needed datasets. The dataset represents the aggregated results of their survey [5] conducted by National Statistical Institutes across European countries focused on the use of Information and

Communication Technologies (ICT) in households and by individuals. The publicly available dataset and metadata do not contain any individual microdata, just the aggregated country data. The percentage value represents the proportion of individuals for each country that have expressed the same opinion or have the same attitude related to the use of ICT, and in our case, generative AI.

The two datasets chosen [7, 8] include Generation Zs' opinions and stances on the use of generative AI, as well as the reasons they choose not to use generative AI. These datasets received the last update on the 25th of February 2026. In total, the data was available for thirty-four countries that participated in the survey.

3.2 The proposed conceptual model

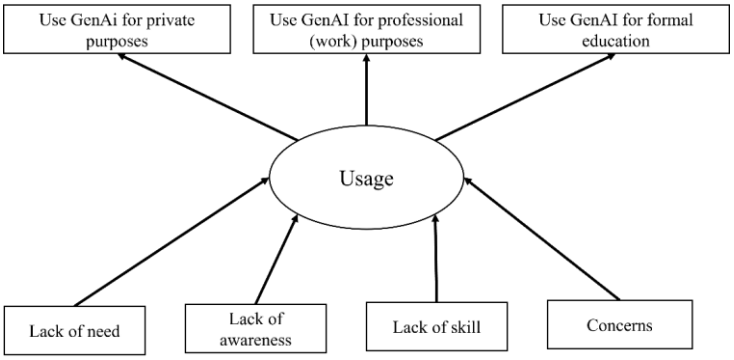
After the available data from Eurostat was collected, based on the nature of the data and the research hypothesis, the following model elements were considered:

Latent Factor – Usage, grouped all three *reasons for using generative AI;* - measured by those who have used generative AI
Predictor 1 - Lack of need, is the variable for *no need*;
Predictor 2 - Lack of awareness, is the variable *did not know generative AI existed*;
Predictor 3 - Lack of skills, is the variable for *did not know how to use generative AI*;
Predictor 4 - Concerns, is the variable for *concerns related to privacy, security, and/or safety of using generative AI*.
Predictors 1-4 – are measured by those who have not used generative AI.

The five defined model elements constitute the proposed SEM model (Figure 1). In the proposed model, the endogenous or dependent variable is *Usage*, while the exogenous or independent variables are the remaining four. So, the latent variable is the general usage of generative AI, and it is measured by the variables: *private use, professional or work use, and use in formal education*. As the latent variable is represented by these three variables, the model acknowledges that the adoption of generative AI manifests itself throughout different domains of users' lives. This means that the proposed model should provide a well-rounded understanding of how Generation Z incorporates generative AI into their everyday and professional life.

Having independent variables be the actual barrier to adopting generative AI, the proposed model could explain both how the younger generations use generative AI and the core reason they choose not to use generative AI. Constructing a model using one latent variable and four distinct independent variables will provide great understanding of how adoption of generative AI is perceived, by not only looking at the usefulness of it but also the social factors such as awareness, competencies, and risk perceptions of young people.

Fig. 1. The proposed conceptual model



Based on the created SEM model, a hypothesis set was formed regarding the impact of each independent variable on the latent construct:

H1: Country-level prevalence of perceived lack of need among non-users is negatively associated with the use of generative AI.
H2: Country-level prevalence of perceived lack of awareness among non-users is negatively associated with the use of generative AI.
H3: Country-level prevalence of perceived lack of skills among non-users is negatively associated with the use of generative AI.
H4: Country-level prevalence of perceived concerns among non-users is negatively associated with the use of generative AI.

The SEM approach was selected because it enables the simultaneous examination of multiple relationships between observed indicators and the latent construct representing generative AI usage. The data set created may be considered smaller in size as it does contain a relatively small number of country-level observations. But this was not an issue because the SEM analysis is considered appropriate for exploratory modelling due to its structure and ability to evaluate both the structural relationships between variables and its overall model fit within a single analytical framework [14]. The analysis will be interpreted with caution and reasonableness, given the aggregated structure of the data and the limited sample size. The model assessment was performed using the AMOS 22 and SPSS 25.

4 Results

This section highlights results and, as such, is divided into three main subsections to present the data analysis comprehensively. Given the nature of the aggregated data, the analysis focused on each country, allowing results to be presented as cross-country comparisons within the European Union (EU) and to provide regional insights for the Balkan region. The final subsection highlights the results of the proposed SEM model

and which of the four predictors have a statistically significant effect on the usage of generative AI.

4.1 Results for the European Union-level analysis

At the outset, the data were analysed using standard descriptive statistics metrics: max value, min value, mean, standard deviation, and coefficient of variation. The goal was to see how generative AI users are dispersed amongst EU countries, and what the main reason is for the use of generative AI. There are 27 EU countries in the database, and there is no missing data. The findings show a notable difference in the preferred reason for using generative AI. The results of the EU-level descriptive statistics are presented in Table 2.

Table 2. Descriptive statistics on the use of generative AI in EU countries

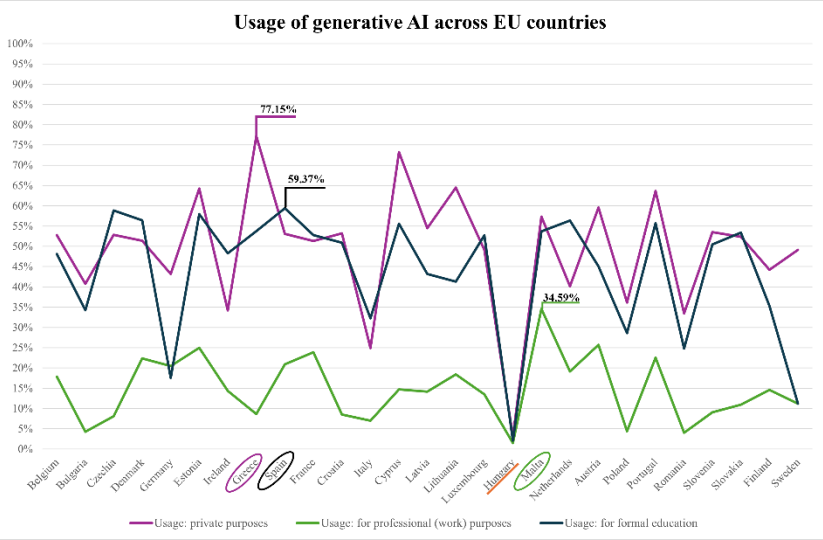
Descriptive statistics	Usage			Barriers			
	Private purposes	Professional purposes	Educational purposes	Lack of need	Lack of awareness	Lack of skill	Concerns
Max value	77.15%	34.59%	59.37%	46.39%	12.35%	9.72%	5.09%
Min value	3.06%	1.46%	2.11%	12.08%	0.00%	0.00%	0.00%
Mean	49.36%	14.81%	43.69%	22.16%	1.67%	2.29%	1.48%
Standard deviation	0.1506	0.0805	0.1548	0.0917	0.0257	0.0222	0.0123
Coefficient of variation	0.3052	0.5435	0.3543	0.4139	1.5405	0.9684	0.8316

Based on these results, the overall reason for using generative AI for Generation Z users is stated to be for private purposes (mean=49.36%). Bearing in mind that Generation Z is just starting to be part of the workforce, it comes as no surprise that the least stated usage of generative AI is for professional purposes (mean=14.81%), seeing that most of them are still in the educational system. The results of the values of the coefficient of variation give an interesting perspective on the dispersion of the reasons why generative AI is used among the countries. The lowest value of this coefficient is for using generative AI for private purposes. This shows that the usage of generative AI among Generation Z is relatively consistent across EU countries. The opposite can be said for the reason of using generative AI for professional purposes, as this reason has the highest value of the coefficient of variation, indicating disparity between countries. These findings suggest that, based on country-level data on reported generative AI usage for professional purposes, this type of usage shows the greatest variation across EU countries compared to private usage.

The reported data show that some countries have higher rates of generative AI use, while others have considerably lower rates. The main reason (mean= 22.16%) that Generation Z from EU countries is not using generative AI is, by far, their stance that they do not need generative AI in their life. Most of these non-AI users are GenZs from Poland. Non-generative AI users from Hungary reported the highest response rate for their lack of skills as the main reason they are not using generative AI. The relatively lowest reported reason for not using generative AI was privacy and safety concerns

(mean= 1.48%). Figure 2 displays the analysed data and the countries with the highest usage and the lowest generative AI adoption percentages.

Fig. 2. Generative AI usage in EU countries – private, professional, and educational usage

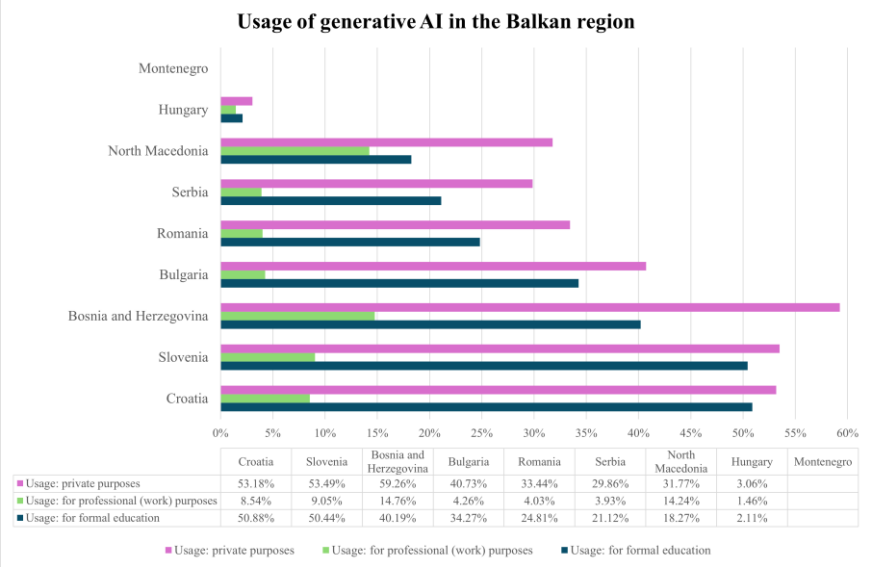


The leading countries when it comes to using generative AI amongst Generation Z are: Greece 77.15% for private use, Spain 59.37% for use in education and Malta 34.59% for professional use. On the contrary, Hungary has the lowest measured response for each type of usage, as all the usage types measured were between 3% and 1%. These findings are further supported by the analysis of reasons why users do not use generative AI, where the most reported reason for Hungarian non-use was the stance that they do not need generative AI. These findings are interpreted at the country level, as the analysis is based on aggregated indicators rather than individual-level behavioural data.

4.2 Results for the Balkan region-level analysis

The Balkan region consists of countries that are part of the EU, such as Bulgaria, Croatia, Hungary, Slovenia, and Romania, as well as non-EU countries Serbia, North Macedonia, Bosnia and Herzegovina, and Montenegro. Having said that, differences in the level of usage are expected. Figure 3 presents an overview of how the Balkan countries compare with each other in how their Generation Z individuals use generative AI. As can be noted, there were no recorded data for Montenegro on the usage of generative AI.

Fig. 3. Generative AI usage in the Balkan region– private, professional, and educational usage

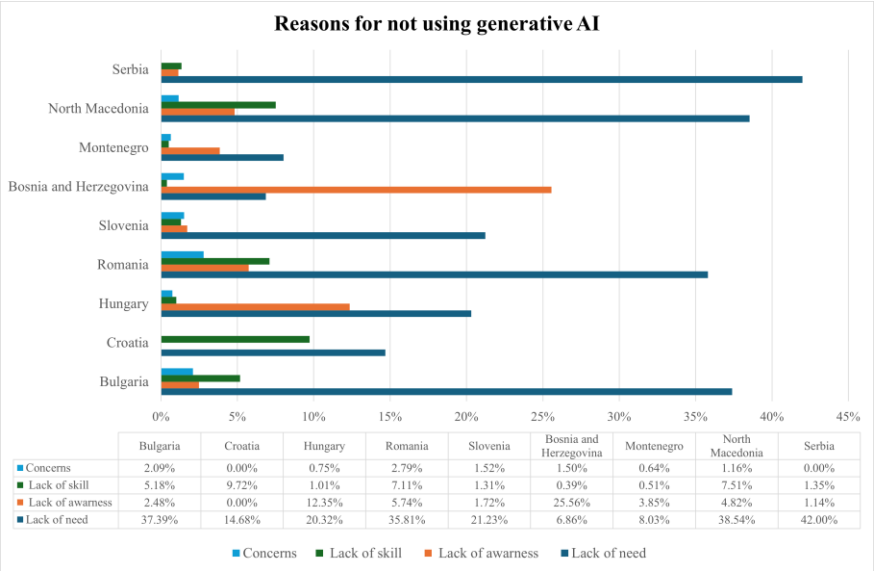


	Croatia	Slovenia	Bosnia and Herzegovina	Bulgaria	Romania	Serbia	North Macedonia	Hungary	Montenegro
Usage: private purposes	53.18%	53.49%	59.26%	40.73%	33.44%	29.86%	31.77%	3.06%	
Usage: for professional (work) purposes	8.54%	9.05%	14.76%	4.26%	4.03%	3.93%	14.24%	1.46%	
Usage: for formal education	50.88%	50.44%	40.19%	34.27%	24.81%	21.12%	18.27%	2.11%	

Bosnia and Herzegovina reported the highest reason for using generative AI was for private purposes (59.26%), followed by Slovenia with the response rate of 53.49% and Croatia with 53.13%. In Bosnia and Herzegovina, the professional usage of generative AI was notably lower (14.76%) in comparison with the reported private use. Interestingly, the two countries where generative AI is highly used for educational purposes are Croatia (50.88%) and Slovenia (50.44%). Serbia reported the use of generative AI mostly for private (29.86%) or educational (21.12%) purposes, while just 3.93% was reported for the use of generative AI for work. Romania and North Macedonia also follow the pattern of AI users using generative AI for private purposes (33.44% and 31.77%) and using generative AI the least for work. Seeing that there is no data for Montenegro, Hungary has the lowest use of generative AI in the Balkan region. Focusing on the use of generative AI for educational purposes, the smallest distribution gap amongst the usages is in Croatia and Slovenia, while the biggest is North Macedonia.

When these findings are compared with the reported barriers presented in Figure 4, additional insights into country-level non-usage patterns can be observed. The most frequently reported barrier across the observed Balkan countries was the reported lack of need for generative AI usage. Reported concerns related to privacy and safety remained relatively low across the observed countries. Bosnia and Herzegovina reported the highest prevalence of lack of awareness as a reported barrier to generative AI usage.

Fig. 4. Reasons for not using generative AI in the Balkan region



	Bulgaria	Croatia	Hungary	Romania	Slovenia	Bosnia and Herzegovina	Montenegro	North Macedonia	Serbia
Concerns	2.09%	0.00%	0.75%	2.79%	1.52%	1.50%	0.64%	1.16%	0.00%
Lack of skill	5.18%	9.72%	1.01%	7.11%	1.31%	0.39%	0.51%	7.51%	1.35%
Lack of awarness	2.48%	0.00%	12.35%	5.74%	1.72%	25.56%	3.85%	4.82%	1.14%
Lack of need	37.39%	14.68%	20.32%	35.81%	21.23%	6.86%	8.03%	38.54%	42.00%

By analysing the aggregated data, insights are provided on how Generation Z living in the Balkan region uses generative AI, as well as the main reasons they do not use generative AI.

4.3 Results of the SEM analysis

Before conducting the SEM analysis, we tested the internal reliability of Factor 1 using Cronbach's alpha coefficient. The measured Cronbach's alpha coefficient for this latent factor is 0.815, indicating satisfactory internal consistency.

The initial proposed model has a good fit with the collected data: Chi-square = 15.549, df = 14, p = 0.342; the comparative fit index (CFI) has the value of 0.973; Tucker–Lewis's index (TLI) is 0.945; and the root mean square error of approximation (RMSEA) is 0.058. Comparing the obtained values of indicators with the threshold, CFI and TLI should be above 0.9, while the RMSE should be less than 0.08; we can conclude that the model fits the data [23].

Based on the SEM analysis, the proposed model explains 72.9% of the variance in how the four predictors/barriers affect the usage of generative AI for Generation Z users. As shown in Table 3, two out of four barriers have statistically significant effects on the usage of generative AI at the county level. The SEM results indicate that the two statistically significant reasons users aged 16 to 24 do not use generative AI are a lack of need and a lack of awareness. In contrast, the lack of skills and concerns about privacy, security, and safety were not statistically associated with generative AI usage.

Table 3. Evaluation results of the proposed SEM model

Outcome	Predictor	Std. coeff	UnStd. coeff	C.R.	P value	R^2
Use of GenAI	Lack of need	-0.762	-0.905	-5.508	<0.001	0.729
	Lack of awareness	-0.374	-1.045	-3.091	0.002	
	Lack of skill	-0.086	-0.488	-0.741	0.459	
	Concerns	-0.012	-0.134	-0.107	0.915	

**Note for the abbreviations*: Std. coeff = standardized coefficient; UnStd. coeff = unstandardized coefficient; C.R. = critical ratio; p-value= indicates statistical significance; R^2 = coefficient of determination.

On the other hand, it is interesting to see that the other two hypothesised barriers have no statistically significant effect on the use of generative AI. The results show that, at a country-level analysis, the users belonging to Generation Z do not lack the skill to use generative AI tools; rather, based on the aggregated data, they choose not to use them. Even though we live in a digital world and data is the new currency, this aggregated country-level data reports that Generation Z does not express concern or prioritise the risk of data leaks when it comes to making the decision on using generative AI.

5 Discussion

Using the SEM, the question of what impacts the adoption and use of generative AI for Generation Z has been analysed. The SEM analysis results suggest that non-AI users do not lack the skills nor express strong ethical concerns regarding generative AI. At the country-level analysis, the aggregated indicators did not show statistically significant associations between reported lack of skills or concerns and generative AI usage.

At the European Union level, recent studies indicate that awareness and experimentation with generative AI tools are rapidly increasing, particularly among the younger population. However, the data on the actual usage remains uneven and often limited to specific use cases such as content generation and academic writing support. This suggests that high exposure does not necessarily lead to consistent or meaningful use [16, 17]. If we compare our statistical results with the data from recent industry-based studies, there is a bit of a paradox. The data analysis has shown that Generation Z is not using generative AI for educational purposes. That leads to the question: Why are educational institutions then investing resources into buying and creating generative AI detection software? This can be explained in two ways [2, 26]: some respondents may have been reluctant or uncomfortable admitting to using generative AI tools. In contrast, others may not have fully distinguished between the term "generative AI" and well-known generative AI tools, leading them to give partially accurate answers.

Examining the results, for the Balkan region, this paradox can be partially explained by differences in digital maturity, institutional readiness, and the lack of structured guidance for using generative AI tools [25]. Research suggests that users in such environments often experiment with new technologies but do not fully integrate them into everyday practices [18, 24]. On the other hand, rather than assuming inaccurate responses, this discrepancy can be interpreted as a gap between perceived and actual use, as well as differences in awareness and understanding of generative AI tools [17, 18].

This is an interpretation based on the observed mismatch between results and broader industry trends. These results appear inconsistent with the increasing number of generative AI tools being introduced and promoted for educational purposes. If Generation Z is not using them extensively, the question remains who the primary users are. When the focus shifts to Serbia, previous research highlights a noticeable imbalance between rapid student adoption of generative AI tools and a slower institutional response. While students tend to adopt tools such as ChatGPT quickly and often without a full understanding of their implications, educators and institutions approach these technologies with caution and uncertainty. These patterns become even more pronounced. The lack of clear institutional policies, formal training, and integration of generative AI into educational systems contributes to inconsistent and selective usage among students [25].

Even though the Eurostat data can be further discussed, the result of the SEM analysis can be used by different interested parties depending on their focus. Marketing research specialists can use this result to eliminate the two main barriers to adopting new generative AI solutions. The aggregated country-level findings suggest that reported lack of skills and reported privacy concerns were not statistically significant barriers within the proposed SEM model. This can be interpreted to mean that Generation Z, which has grown up with the Internet and social media, reports no lack of skill set to use any new platform or application, as well as reporting the lack of any major concerns for data privacy or safety. The SEM results confirm that the main barriers are not technical or risk-related, but rather a lack of perceived need and lack of awareness of generative AI tools. This aligns with previous studies showing that perceived usefulness and awareness are key drivers of technology adoption.

The relatively low level of reported concerns may reflect differences in awareness, digital habits, or perceptions regarding data privacy and cybersecurity. Sales specialists can use this information to create an online training platform or course targeting countries with the highest value due to a lack of awareness of generative AI tools and how they can be used. Other researchers can use the generated results and create a new conceptual model for adopting generative AI or create a new system for surveying Generation Z on the use of generative AI.

6 Conclusion

This paper combined SEM analysis with descriptive statistical analysis to answer the question of how generative AI is used across several countries and the most common reasons Generation Z is not using generative AI. To gain a broader understanding of the topic, the first step was to conduct a brief bibliometric study of the use and adoption of AI.

When it comes to how Generation Z uses generative AI, results across countries vary significantly. The most common reason Generation Z uses generative AI was stated to be private use, followed by usage for educational purposes. The created SEM model was tested and proven to have a quite good fit with the Eurostat data, which led to the model explaining 72.9% of the variance in how the four main reasons for not using AI affect the general usage and adoption of generative AI for Generation Z. The findings

suggest that the lack of need and lack of awareness are the two main barriers that are most strongly associated with the lower generative AI usage. In contrast, the lack of skills and privacy concerns did not demonstrate statistically significant associations within the proposed SEM model.

The main limitation of this study is the use of third-party aggregated data, without the possibility of conducting secondary metadata testing, and only having data for the age group between 16 and 24 years old. As discussed in Section 5, because the questionnaire used to collect data was created by Eurostat, the authors lack the means to test or verify whether the participants were well informed or honest during the survey. Additionally, the aggregated structure of the data limits the interpretation of the findings to only country-level associations rather than individual-level behavioural conclusions. This includes data limitations associated with using covariance-based structural equation modelling (CB-SEM) with a relatively small sample of countries. Even though the model produced acceptable model estimates, CB-SEM generally performs better with larger sample sizes. Therefore, the findings should be interpreted as exploratory and should be validated in future studies using larger datasets or alternative analytical approaches.

We suggest that further research in this area could explore differences not between countries but at a macro level within a single country. Another research path could focus on how education on the use of generative AI affects the adoption process among Generation Z. Additionally, the list of factors that influence an individual's decision to use AI can be extended.

References

1. Ali, A. et al.: Different Time, Different Language: Revisiting the Bias Against Non-Native Speakers in GPT Detectors. Association for Computational Linguistics, Rabat, Morocco (2026). https://doi.org/10.18653/v1/2026.eacl-srw.20.
2. Andersen, J.P. et al.: Generative Artificial Intelligence (GenAI) in the research process – A survey of researchers' practices and perceptions. Technology in Society. 81, 102813 (2025). https://doi.org/10.1016/j.techsoc.2025.102813.
3. Aykens, P. et al.: 9 Trends Shaping Work in 2026 and Beyond, https://hbr.org/2026/02/9-trends-shaping-work-in-2026-and-beyond, (2026).
4. Elad, B.: AI in Education Statistics 2025: Funding, Privacy, and Performance, https://sqmagazine.co.uk/ai-in-education-statistics/, last accessed 2026/04/06.
5. Eurostat: ICT usage in households and by individuals (isoc_i), https://ec.europa.eu/eurostat/cache/metadata/en/isoc_i_esms.htm#shortdata_descrDisseminated, last accessed 2026/04/08.
6. Eurostat: Individuals - reasons for not using generative AI tools, https://ec.europa.eu/eurostat/databrowser/view/isoc_ai_iaiuxr/default/table?lang=en&category=isoc.isoc_i.isoc_ai, last accessed 2026/04/06.
7. Eurostat: Individuals - use of generative AI tools, https://ec.europa.eu/eurostat/databrowser/view/isoc_ai_iaiu/default/table?lang=en&category=isoc.isoc_i.isoc_ai, last accessed 2026/04/06.

8. Filippelli, S. et al.: Generative AI and employee well-being: Exploring the emotional, social, and cognitive impacts of adoption. Journal of Innovation & Knowledge. 11, 100844 (2026). https://doi.org/10.1016/j.jik.2025.100844.
9. Giugliano, A.: Best AI Tools for Academic Research in 2026: A Step-by-Step Workflow Guide, https://www.thesify.ai/blog/best-ai-tools-academic-research, last accessed 2026/04/06.
10. Hays, K.: Why tech CEOs suddenly love blaming AI for mass layoffs, https://www.bbc.com/news/articles/cde5y2x51y8o, last accessed 2026/04/06.
11. He, R. et al.: Generative artificial intelligence: a historical perspective. Natl Sci Rev. 12, 5, nwaf050 (2025). https://doi.org/10.1093/nsr/nwaf050.
12. Hermann, E. et al.: Why Gen AI Feels So Threatening to Workers, https://hbr.org/2026/03/why-gen-ai-feels-so-threatening-to-workers, (2026).
13. Hughes, L. et al.: Beyond the hype: Organisational adoption of Generative AI through the lens of the TOE framework–A mixed methods perspective. International Journal of Information Management. 86, 102982 (2026). https://doi.org/10.1016/j.ijinfomgt.2025.102982.
14. Kline, R.: Principles and Practice of Structural Equation Modeling. Guilford Press (2023).
15. Liao, C.-H. et al.: Out of gratification or fears? A dual model to explore drivers of generative artificial intelligence adoption. Technology in Society. 85, 103177 (2026). https://doi.org/10.1016/j.techsoc.2025.103177.
16. McKinsey: The State of AI: Global Survey 2025 | McKinsey, https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai#/, last accessed 2026/04/20.
17. Mohammadi, E. et al.: Is generative AI reshaping academic practices worldwide? A survey of adoption, benefits, and concerns. Information Processing & Management. 63, 1, 104350 (2026). https://doi.org/10.1016/j.ipm.2025.104350.
18. Pereira, C.S. et al.: Adoption of Generative AI in Portuguese Higher Education: An Empirical Study of Perceptions and Academic Practices. Procedia Computer Science. 278, 620–627 (2026). https://doi.org/10.1016/j.procs.2026.03.032.
19. Pohrebniyak, I.: 350+ Generative AI Statistics [March 2026], https://masterofcode.com/blog/generative-ai-statistics, last accessed 2026/04/06.
20. Statista: Most downloaded generative AI apps 2026, https://www.statista.com/statistics/1554189/top-gen-ai-apps-by-downloads/, last accessed 2026/04/06.
21. Stein, H.: Toward a taxonomy of generative AI use cases in business contexts: Integrating complexity, risk, and strategy. Journal of Computational Science. 96, 102826 (2026). https://doi.org/10.1016/j.jocs.2026.102826.
22. Sun, Z. et al.: Generative AI and ESG opportunism in supply chains: A utilitarian perspective on unintended consequences for sustainability. Technological Forecasting and Social Change. 224, 124498 (2026). https://doi.org/10.1016/j.techfore.2025.124498.
23. Taasoobshirazi, G., Wang, S.: The Performance of the Srmr, Rmsea, Cfi, and Tli: An Examination of Sample Size, Path Size, and Degrees of Freedom. Journal of Applied Quantitative Methods. 11 (3), 31–39 (2016).
24. Tbaishat, D. et al.: Modeling generative AI adoption in higher education: An integrated TAM–TPB–SDT framework with SEM validation. Computers and Education: Artificial Intelligence. 10, 100541 (2026). https://doi.org/10.1016/j.caeai.2026.100541.

25. Vukmirović, D. et al.: Challenges of Integration and Implementation of Generative Artificial Intelligence in the Process of Higher Education. Conference: International Multidisciplinary Conference “Challenges of Contemporary Higher Education” - CCHE 2024. 556–595 (2024).
26. Wu, C. et al.: Exploring PhD students’ utilization of generative AI in academic writing for publication purposes: Insights for EAP. Journal of English for Academic Purposes. 79, 101612 (2026). https://doi.org/10.1016/j.jeap.2025.101612.

Original Research Article *https://doi.org/10.5281/zenodo.23479044*

Chatbots in Integrated Marketing Communications in Higher Education: Strategic Implications and Implementation Challenges

Dragana Todovic Milanovic[2][0009-0006-0005-7869], Velimir Stavljanin[1][0000-0001-5450-806X], Srdjan Masic[2][0000-0001-8852-2054] and Milica Kostic-Stankovic[1][0000-0002-8334-7063]

[1]University of Belgrade, Faculty of Organizational Science
[2] University of East Sarajevo, Faculty of Medicine
todovicd@yahoo.com
velimir.stavljanin@fon.bg.ac.rs
smasic@gmail.com
milica.kostic-stankovic@fon.bg.ac.rs

Abstract. This paper analyzes the role of chatbots as tools of integrated marketing communications (IMC) with potential applications in higher education. In the context of intensive digitalization and increasing competition among higher education institutions, chatbot systems are becoming a crucial tool for enhancing communication with students and other stakeholders, as well as for refining marketing strategies. At the same time, the effective integration of chatbot technologies requires alignment between technological capabilities, user needs, regulatory requirements, and ethical principles.
A review of relevant literature indicates that chatbots enable personalized and timely communication, enhance user experience, and contribute to more efficient management of communication processes. However, their implementation also raises challenges related to information reliability, data protection, ethical concerns, and the need for continuous system monitoring.
The analysis suggests that chatbots represent an important strategic tool within integrated marketing communications in higher education. Their successful implementation depends on the ability of institutions to balance technological innovation with user needs, while respecting ethical and regulatory frameworks. Particular attention is given to countries such as Serbia, where the implementation of these technologies is not yet systematically developed but represents significant potential for improving communication practices through a strategic and responsible approach.

Keywords:AI, IMC, ChatGPT, higher education institutions

1 Introduction

The growing use of artificial intelligence has significantly influenced marketing activities and research, particularly in dynamic and competitive environment [1]. The integration of artificial intelligence into the higher education has created new

opportunities for improving teaching, research and communication with students [2]**.** The development of AI technologies, especially chatbot systems, represents an important step in the digital transformation of higher education, offering new possibilities for enhancing learning and communication processes, while also introducing certain challenges [3].One of the most well-known examples of modern chatbot systems is ChatGPT, based on the GPT architecture, which enables the generation of different types of content, including text, images, and audio [4]. The development of these systems relies on advances in generative artificial intelligence (GAI), large language models (LLMs) and natural language processing (NLP) technologies [3].

AI and chatbots have already transformed learning processes in higher education [5, 6]. The COVID-19 pandemic further accelerated the digitalization of education and reduced the importance of traditional learning model, increasing the need for modern digital communication solutions between institutions and users [7]. Chatbots are now widely used in educational systems worldwide [8], enabling more personalized and timely communication across different aspects of the educational process [9]. Their use contributes to improved organizational processes and strengthens the competitive position of educational institutions [1].

Marketing in higher education institutions is increasingly evolving toward digitalization, personalized communication, and strategic communication management. In a competitive environment, institutions adopt innovative promotional strategies and digital tools to attract and retain students. Integrated marketing communications play a key role in ensuring consistent communication across channels, while AI technologies, particularly chatbots, further enhances the effectiveness of these activities [1].

Despite the large number of studies on chatbots in education, most previous research has focused on their educational and administrative functions. Their role in the integrated marketing communications (IMC) of higher education institutions has received much less attention. This paper also emphasizes their importance in today's digital environment.

The aim of this paper is to identify the main benefits, challenges, application models, and strategic implications of using chatbots in higher education. To achieve this aim, the following research questions were defined:

- What are the main communication-related applications of AI-based chatbots in higher education?
- What benefits do chatbots provide in integrated marketing communications?
- What are the main challenges and limitations of their use?
- What strategic implications does the use of chatbots have for higher education institutions, especially in developing countries?

The paper examines the role of chatbots as tools of integrated marketing communications in higher education, focusing on their advantages, limitations, and implementation challenges.

2 Methodology

A structured literature review was conducted to examine the role of AI-based chatbots as tools of integrated marketing communications in higher education. The review was motivated by the increasing use of artificial intelligence technologies in educational institutions and the growing importance of chatbots for student communication, study program promotion, and user experience improvement.

The literature search was conducted in Google Scholar, Scopus, and Web of Science using the keywords: "AI chatbots," "artificial intelligence," "higher education," and "integrated marketing communications." Journal articles and conference papers published in English between 2018 and 2025 were included.

The relevance of the studies was assessed based on their titles, abstracts, and full texts. Studies related to the use of chatbots in the integrated marketing communications of higher education institutions were included, while studies not directly related to the research topic were excluded.

A total of 45 full-text studies were reviewed and grouped into two thematic categories: studies on the application of artificial intelligence in higher education (n = 21) and studies on chatbots, marketing, and integrated marketing communications (n = 24). After assessing their relevance and consistency with the research objectives, 38 studies were included in the final analysis. The search and selection process is presented in Figure 1.

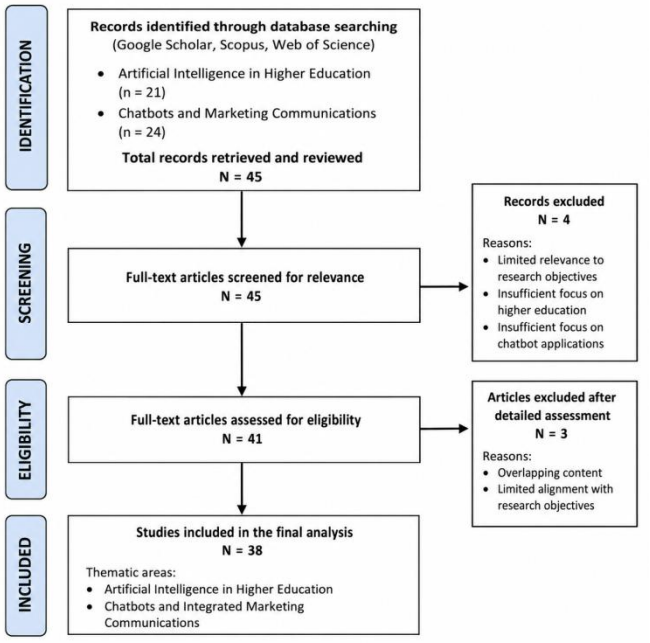


Fig. 1. Literature search and selection

The literature was analyzed using a thematic classification framework consisting of four dimensions: (1) chatbot applications in higher education, (2) benefits of their implementation, (3) challenges and limitations of their use, and (4) strategic implications for integrated marketing communications. This approach enabled a systematic overview of existing knowledge on the use of chatbots in integrated marketing communications in higher education and helped identify their main benefits, challenges, and strategic implications.

Table 1. presents a comparative analysis of selected studies included in the review, focusing on their main findings, challenges, and implications for integrated marketing communications in higher education. The studies presented in Table 1 indicate that AI chatbots can significantly improve communication, personalization, and user engagement in higher education. Most studies highlight their potential to provide timely information, support students, and improve communication efficiency.

At the same time, several challenges are identified, including privacy concerns, ethical issues, response reliability, and the need for human supervision. Some studies also emphasize the risks of inaccurate information and excessive dependence on AI systems.

The findings suggest that the successful use of chatbots depends not only on technological capabilities but also on institutional support, digital literacy, and responsible implementation. Overall, AI chatbots have considerable potential to strengthen integrated marketing communications in higher education when applied in a strategic and ethical manner.

Table 1. Comparative Analysis of Selected Studies on AI Chatbots in Higher Education and Integrated Marketing Communications

Author (Year)	Study Context	Key Findings	Challenges and Limitations	Strategic Implications for IMC
Yoldaş & Aycı (2024)	AI and integrated marketing communications	ChatGPT improves engagement, personalization, and communication outcomes.	Ethical concerns, output reliability issues, and the need for human supervision.	ChatGPT should complement human efforts through personalization, automation, and ethical AI use.
Wu & Chen (2021)	AI and HEI branding	AI improves communication, brand visibility, and institutional competitiveness.	Limited communication reach and challenges in information management.	Strategic AI use can enhance personalization, channel integration, and brand management.
Lytvynchuk (2025)	AI personalization in marketing communications	AI chatbots enable real-time personalization and audience segmentation.	Privacy concerns, personalization errors, reputational risks, and overreliance on AI.	Strategic AI use should combine personalization, human oversight, and ethical governance.
McGrath et al. (2025)	Generative AI in higher education	Studies report both benefits and risks of generative AI chatbots.	Small samples, exploratory designs, and limited theoretical grounding.	AI integration requires critical evaluation and responsible implementation.
Okonkwo & Ade-Ibijola (2021)	Chatbots in higher education	Chatbots improve student support,	Limited ability to handle complex queries and	Chatbots can enhance student engagement and institutional responsiveness.

		communication, and information access.	human-like interaction.	
Abdallah et al. (2025)	ChatGPT in higher education	ChatGPT supports learning, academic assistance, and communication efficiency.	Accuracy concerns, AI hallucinations, and academic integrity issues.	Institutions should combine AI tools with human oversight and clear policies.
Ilieva et al. (2023)	Generative AI in higher education	AI chatbots enhance learning, engagement, and personalized support.	Ethical concerns, data privacy issues, and response reliability.	AI can improve personalized communication and student experience.
Labadze et al. (2023)	AI chatbots in education	AI chatbots improve access to information and educational support.	Privacy concerns, implementation barriers, and uneven adoption.	Successful implementation requires institutional support and clear governance.
Stöhr et al. (2024)	Student perceptions of AI chatbots in higher education	Students generally perceive AI chatbots positively across disciplines.	Unequal adoption, varying digital skills, and concerns about future use.	Institutions should promote AI literacy and inclusive implementation strategies.
Dragojević & Turudić (2024)	Student perceptions of ChatGPT in higher education	Students recognize ChatGPT's potential to improve learning quality.	Concerns about responsible use, accuracy, and academic integrity.	Effective implementation requires institutional guidance and responsible use policies.

3 Integrated Marketing Communications in Higher Education

Communication activities play a crucial role in the functioning of higher education institutions but require active participation from both institutions and users [10]. The concept of integrated marketing communications emphasizes the alignment of communication content with users' informational needs [11]. Students gather information about study programs, admission requirements, and academic activities through various IMC tools and sources, including official university channels, competing institutions, and online communities. Inconsistent or inaccurate information can lead to misperceptions and reduce communication effectiveness [12].

The development of digital media has provided numerous advantages due to its accessibility and interactivity, but it has also increased the complexity of communication. Marketing managers must develop strategies to manage communication effectively in such an environment [13].According to Schultz (1992), as sited in [14], integrated marketing communications represent a strategic business process that includes the planning, development, implementation, and evaluation of coordinated and measurable brand communication programs, integrating both internal and external audiences.

For higher education institutions, the importance of integrated marketing communications lies in the integration of communication content,channels, and the overall communication process [14]. One of the key advantages of the IMC approach is the ability to target clearly defined audiences, supported by technologies that enable the collection and analysis of user behavior data [13]. Unlike the traditional communication model, which focuses on the message sent by the organization, the IMC approach places the user and their needs at the center [8]. The goal of these activities is to build

a strong brand, while the effects include attracting students and academic staff, improving graduate employability, strengthening financial capacity, and enhancing the international influence of the institution [14].

4 Chatbots as an IMC Tool in Higher Education

In a digitally transformed environment, artificial intelligence–based tools play an important role in integrated marketing communications, with chatbots as one of their key applications [1]. The development of AI technologies has transformed marketing communications, including communication in higher education [13]. AI enables institutions to communicate with users through integrated marketing strategies supported by AI technologies, while maintaining professional and ethical standards[1].

The use of artificial intelligence allows educational institutions to better understand user needs and develop personalized communication strategies. These tools support the analysis of large volumes of data, optimization of marketing campaigns, and more efficient management of user acquisition and retention[1].

Marketing communication management is a strategic activity supported by information technologies, aimed at improving connections between organizations and users. In this context, chatbots are increasingly integrated into digital strategies as tools for continuous communication with users [15, 16]. In higher education institutions, they enable the analysis of user behavior and preferences, forming the basis for targeted communication activities [17]. AI algorithms analyze data from social media to support more precise targeting and personalization of marketing messages. However, the digital environment is also characterized by information overload, which may reduce the communication effectiveness, while the digital divide limits access to these technologies for some user groups [18]. Data collected through AI systems supports the development of communication strategies tailored to their needs.

Chatbots are an important tool of integrated marketing communications, enabling and interactive two-way communication between organizations and users [8]. In higher education, they are used to improve communication and provide timely information to students during study program selection, enrollment, and further academic development [8, 19]. Research shows that their role goes beyond communication, influencing decision-making processes and user interaction with institutions. Users are often motivated by time efficiency and the availability of a non-judgmental communication environment. Chatbots are still mostly used as support tools rather than a fully integrated communication components, which limits their potential within IMC. User engagement depends not only on technological features but also on design elements, such as gamification. Their effectiveness also depends on factors such as digital literacy levels and institutional support [20], which is particularly relevant in Serbia. Furthermore, AI-driven systems not only improve access to information but also actively influence decision-making processes by reducing information overload and increasing confidence in choices. In this context, the so-called personalization–

privacy paradox is particularly relevant, whereby increased personalization enhances communication relevance while simultaneously raising concerns regarding privacy protection. Additionally, the use of predictive analytics enables a shift from reactive to proactive communication approaches [21].

The increasing globalization of higher education has intensified competition among higher education institutions. As a result, universities are increasingly using innovative communication and marketing strategies to attract and retain students. In this context, AI-based chatbots have become an important tool of integrated marketing communications because they enable fast, personalized, and continuously available communication with current and prospective students. Their use improves information delivery, enhances the user experience, and strengthens communication between higher education institutions and their target audiences [1].

Recent studies indicate that AI chatbots play a significant role in addressing structural challenges in higher education, such as the imbalance between the number of teachers and students, delays in providing feedback, and the burden of repetitive inquiries on teaching staff [22]. Empirical findings confirm high levels of efficiency and user satisfaction in administrative and communication tasks, but also highlight variability in response quality, with some studies indicating that only a small proportion of responses are fully accurate. A particular issue is the generation of inaccurate or fabricated information and references (AI hallucinations), which raises concerns about the reliability of these systems in academic contexts [23]. Furthermore, the use of AI chatbots varies significantly depending on digital literacy, field of study, and students' socioeconomic status, indicating different usage patterns and potential digital inequalities. Therefore, the implementation of these technologies requires a carefully planned and differentiated approach [24].

In the case of Serbia, the implementation of AI chatbots represents both a challenge and an opportunity for the adoption of AI solutions based on the experiences of developed educational systems. Research among university teachers in Serbia shows that chatbots are mainly used for research, teaching preparation, and administrative tasks, while their use in teaching and assessment remains limited. Although AI chatbots offer many benefits, concerns remain regarding the reliability of their responses, academic integrity, and their potential impact on students' independent and critical thinking. Their successful implementation requires adequate technological infrastructure, clear institutional guidelines, and strategic support [25]. Existing findings indicate that students recognize the potential of ChatGPT to improve learning and the quality of studying, while also emphasizing the need for responsible use and institutional support [26]. Given the limited number of studies conducted in Serbia, further research is needed to examine the long-term effects of AI chatbot adoption in higher education and support the development of evidence-based institutional policies.

4.1 Application of AI Chatbots in Higher Education

One of the most important AI technologies used to support teaching and learning activities is chatbot systems [9]. Their use in education supports the development of

personalized learning experiences [2, 27]. Over time, chatbots have evolved from simple rule-based systems, capable of responding to a limited set of commands, to advanced solutions that use artificial intelligence and natural language processing (NLP), enabling more complex interactions [4, 28]. Contemporary systems, based on large language models (LLMs) and NLP technologies, allow the understanding of user intent and the generation of relevant responses in real time [28].

In the educational context, chatbot systems support more dynamic and interactive learning, overcoming the limitations of traditional communication methods such as email and conventional teacher–student interaction [9]. They encourage students to take an active role in learning by supporting the development of digital competencies, creativity, and collaboration [1, 2].Chatbot systems provide access to various educational resources, including textbooks, academic papers, video content, and interactive simulations, supporting self-directed learning and the development of students' digital competencies [2]. As communication tools within higher education institutions, they can understand student queries, support multiple communication channels, offer multilingual capabilities, and integrate with learning platforms and databases [29]. Their use enables the personalization of the learning process through the analysis of learning styles and the provision of individualized recommendations, allowing students to access educational content according to their own pace and needs [2, 4]. In addition, chatbot systems support teachers by enabling the monitoring of student progress, assessing levels of understanding, and providing guidance on academic issues and study program selection [9, 13]. They are also used to optimize administrative activities within higher education institutions, including student enrollment, information on scholarships, tuition fees, courses, and grading methods, thereby reducing administrative workload and increasing the efficiency of organizational processes [5, 6]. The use of chatbots can improve learning outcomes and student satisfaction, which are important indicators of the educational quality. Their ability to support a large number of users simultaneously contributes to more efficient learning, time savings [9]. They also support the development of students' skills and competencies, relevant to the modern labor market [5]. In developed educational systems, AI chatbots are already integrated into learning processes and widely accepted by students. However, their effectiveness depends on the accuracy of information and the system's ability to interpret user queries. In contrast, in Serbia, the implementation of these systems is still not systematically developed, indicating a significant gap between current practice in developed countries and the potential for their application in higher education [25].

5 Strategic Implications and Challenges of chatbot Implementation in IMC

The contemporary market environment, characterized by rapid changes in user behavior and intense competition, further complicates traditional marketing communication management, which is why artificial intelligence is gaining an increasingly important role in optimizing communication and marketing processes

[19]. Within integrated marketing communications, the application of AI brings significant strategic advantages, but at the same time introduces a range of operational, ethical, and social challenges. Although chatbots and other AI tools improve the efficiency of communication processes, they cannot fully replace human expertise, as they require continuous monitoring to prevent errors and deviations [13]. In this context, Jarrahi (2018) emphasizes that managers must be prepared for continuous change and adaptation in order to fully exploit the potential of AI in marketing communications [19].

The effectiveness of chatbot implementation in educational and marketing contexts largely depends on user perception. Research indicates that their use has a significant emotional impact on students. While their use may reduce anxiety and increase a sense of security, negative effects are also observed, such as frustration due to inaccurate information, feelings of isolation, and the emergence of technological stress. Although these systems may reduce cognitive load, there is a risk that they may encourage superficial learning and weaken the development of critical thinking. Excessive use may also lead to reduced learning satisfaction and social withdrawal, indicating the need for controlled and responsible implementation of these technologies [30]. User satisfaction represents a key factor in implementation success, as it influences students' willingness to continue using these systems. Service quality and alignment with user expectations directly contribute to their effectiveness [31,32]. Therefore, higher education institutions should systematically monitor user attitudes and integrate them into the development of communication strategies [9, 28].

The automation of marketing activities based on artificial intelligence may lead to transformations in the labor market, particularly in traditional segments of marketing. These changes require continuous employee training and the development of digital competencies, representing an important strategic challenge for organizations [33]. At the same time, although there are concerns about job displacement, some evidence suggests an increasing demand for professionals in digital marketing and AI system management [19].

The application of AI technologies within IMC is accompanied by numerous ethical and regulatory challenges. Key issues include data protection, user privacy, algorithm transparency, and potential bias in decision-making. Chatbot systems may generate inaccurate or incomplete information, making additional human supervision and verification necessary [13]. Their effectiveness largely depends on the quality of available data, and a lack of reliable and up-to-date information may lead to incorrect conclusions and ineffective communication strategies. Therefore, ensuring access to relevant data and their continuous updating is essential for the optimal functioning of AI systems [28,32]. In this context, the development of appropriate ethical standards and regulatory frameworks is a prerequisite for the responsible use of artificial intelligence. Particular importance is given to the EU Artificial Intelligence Act, which defines principles of transparency, accountability, and data protection, as well as the obligation of additional evaluation for high-risk AI systems, including personalized chatbots. Non-compliance with these regulations may result in significant financial penalties [15]. At the same time, different approaches to AI implementation globally, including certain restrictions on the use of ChatGPT due to

concerns related to privacy and academic integrity, indicate uneven development of these systems [34]. In the context of Serbia, establishing clear institutional frameworks implies defining rules for the use of AI tools in academic work, adapting assessment methods, training teaching staff, and developing guidelines for ethical and responsible use. Such an approach enables risk reduction and more efficient integration of AI technologies into the educational system.

The use of AI technologies in integrated marketing communications enables more efficient data management and more precise targeting of prospective students. However, research also indicates initial resistance among teaching staff, primarily due to insufficient digital literacy and the complexity of the technology. Education and training significantly contribute to increasing trust and acceptance of these systems, confirming that successful implementation depends not only on technical capabilities but also on the development of digital competencies and the balance between automation and the human factor [35]. Studies show that more than 85% of students in developed educational systems use AI tools in the learning process, with intrinsic motivation being the key driver of their use. At the same time, a certain level of distrust is present among teachers due to potential ethical concerns. These findings indicate that successful integration of AI technologies requires not only technological development but also user education, which in the context of Serbia represents both an additional challenge and a significant development opportunity [3].

From a strategic perspective, the successful integration of AI into IMC requires alignment of communication channels and consistency of information delivered to end users. Since users gather information from multiple sources, the integration of chatbots with other communication tools represents a key factor in building trust and long-term relationships with target groups[8]. Transparency in data usage and the protection of user privacy represent fundamental prerequisites for the effective application of artificial intelligence in communication activities [13].

The implementation of AI technologies requires significant investments in infrastructure and human resources. The need for highly qualified professionals may represent a limiting factor, particularly in less developed economic environments, potentially leading to digital inequalities between institutions and regions [33]. Research indicates a complex relationship between students and AI chatbots, characterized by both positive attitudes and a significant level of concern regarding their future use. Differences in usage and attitudes toward AI chatbots have been observed depending on the field of study and level of digital literacy, with students in technical and engineering disciplines demonstrating higher usage and more positive attitudes, while students in humanities and health-related fields tend to express greater skepticism. The use of AI tools may further deepen existing educational inequalities, as students with better access to technology and higher levels of digital competence gain advantages over others. These findings highlight the need for the development of AI literacy and tailored approaches to the implementation of these technologies in education [36]. Consequently, organizations are increasingly investing in the development of employee competencies and interdisciplinary knowledge that connects marketing and technology [19]. This issue is particularly relevant in the

context of Serbia, where disparities in access to technology and levels of digital literacy may further deepen existing educational inequalities.

Despite these challenges, artificial intelligence has significant potential to enhance integrated marketing communications. It enables the analysis of large volumes of data, a better understanding of user behavior, and more precise targeting of communication activities, thereby improving communication efficiency in digital environments [37]. The successful application of artificial intelligence in integrated marketing communications depends on the ability of organizations to align technological innovation with ethical principles, regulatory requirements, and actual user needs. Achieving such a balance represents a key factor for the sustainable and responsible integration of AI technologies into modern communication strategies.The application of AI in IMC provides significant strategic advantages but also introduces operational, ethical, and social challenges. Chatbots improve efficiency but cannot fully replace human expertise and require continuous monitoring.

User perception plays a key role in implementation success. While chatbots can reduce anxiety and improve accessibility, they may also cause frustration and technological stress. Ethical issues, including data privacy, transparency, and algorithmic bias, remain critical challenges.

The implementation of AI technologies requires investment in infrastructure and human resources, as well as the development of digital competencies. Differences in digital literacy and access to technology may lead to inequalities in education.

6 Comparative Analysis of Studies

The reviewed studies consistently highlight the positive impact of AI chatbots on communication efficiency, user experience, and personalization in higher education. Most studies agree that chatbot systems contribute to improved access to information, reduced administrative workload, and enhanced student engagement.

However, differences can be observed in the reported effectiveness of chatbot systems. While some studies emphasize high levels of user satisfaction and accuracy, others point to challenges related to response reliability, ethical concerns, and limitations in handling complex queries.

Additionally, variations exist depending on the context of implementation, including differences in technological infrastructure, digital literacy, and institutional support. Studies conducted in developed countries report a higher level of integration and acceptance of AI chatbot systems, whereas research in developing contexts highlights implementation barriers and untapped potential.

7 Limitations and Future Research

This study has several limitations that should be considered when interpreting the results. First, the paper is based on a review of existing literature and does not include empirical research conducted among students, faculty members, or higher education institutions. Therefore, the conclusions are based on the findings of previous studies.

In addition, only studies published in English and available in the selected databases were included in the analysis, which may have led to the exclusion of relevant research published in other languages or indexed in other sources. Furthermore, the rapid development of artificial intelligence technologies may result in new findings and trends that are not covered in this review.
Another important limitation is related to the geographical distribution of the analyzed studies. Most of the available research has been conducted in technologically developed countries, while studies from Serbia and other Western Balkan countries are still limited.
Future research should focus on the empirical examination of chatbot use in higher education institutions, especially in developing countries. Particular attention should be given to the attitudes of students and faculty members toward chatbot use, as well as their impact on student engagement, communication quality, and the effectiveness of higher education marketing activities. In addition, comparative studies across different educational systems and countries could contribute to a better understanding of the opportunities and challenges of implementing AI-powered chatbots in higher education.

8 Conclusion

In the contemporary environment of digital transformation of business processes, the effectiveness of marketing communications largely depends on the ability of organizations to establish timely, flexible, and personalized interactions with users. AI-powered chatbots have evolved beyond simple automated systems and have become a significant component of integrated marketing communications, enabling continuous and interactive communication with users [15]. As widely applied AI tools, chatbots contribute to improving communication with students, enhancing information personalization, and enabling more efficient management of communication processes, thereby becoming an essential element of integrated marketing communications in higher education [13]. At the same time, their implementation requires continuous monitoring, ensuring data quality, and compliance with ethical principles and regulatory requirements, particularly regarding privacy protection and transparency.
Organizations that invest in digital technologies have greater opportunities to achieve competitive advantage, improve business performance, and develop innovative communication strategies [38]. The successful integration of artificial intelligence into IMC requires a strategic approach, appropriate technology selection, and continuous monitoring of system performance [33].
The application of chatbots as AI tools in integrated marketing communications represents significant potential for improving communication and marketing activities in higher education. However, their effectiveness depends on aligning technological innovations with user needs, ethical principles, and contemporary regulatory frameworks. Compared to developed educational systems, where chatbots represent a standard component of digital communication infrastructure, their implementation in

Serbia is still not systematically developed. In this context, the strategic and responsible integration of AI chatbots represents an important development opportunity for modernizing communication practices and strengthening the competitiveness of higher education institutions in the contemporary digital environment.

Acknowledgements. The authors acknowledge the use of Grammarly Premium and ChatGPT 5.2 to improve English in this manuscript. The AI tools were used to assist with language enhancement but were not involved in developing the scientific content. The authors take full responsibility for the originality, validity, and integrity of the manuscript.

Disclosure of Interests. The authors has no competing interests to declare that are relevant to the content of this article.

References

1. Trusca, A.: Implementation models of artificial intelligence in higher education marketing. *Journal of Emerging Trends in Marketing and Management* 1(4), 29-39, (2025)
2. Mukherjee, A., Singh, S.: Leveraging AI Chatbots for Self-Directed Learning: A Study of Practices and Drivers to Adoption among College Students. *Turkish Online Journal of Distance Education* 27(1), 70–81 (2026). DOI: 10.17718/tojde.1632775
3. McGrath, C., Farazouli, A., Cerratto-Pargman, T.: Generative AI chatbots in higher education: A review of an emerging research area. *Higher Education* 89(6), 1533–1549 (2025)
4. Kooli, C.: Chatbots in education and research: A critical examination of ethical implications and solutions. *Sustainability* 15(7), 5614 (2023)
5. Jiménez-García, E., Ruiz-Lázaro, J., Martínez-Requejo, S., Redondo-Duarte, S.: Artificial intelligence and chatbots for sustainable higher education: A systematic review. *RIED-Revista Iberoamericana de Educación a Distancia* 28(2), 81–104 (2025)
6. Williams, R.T.: The Ethical Implications of Using Generative Chatbots in Higher Education. *Frontiers in Education* 8, 1331607 (2024)
7. Williams, R.: An exploration into the pedagogical benefits of using social media: Can educators incorporate social media into pedagogy successfully? PhD thesis, Teesside University (2022)
8. Khoa, B.T.: The impact of chatbots on the relationship between integrated marketing communication and online purchasing behavior in the frontier market. *Jurnal The Messenger* 13(1), 19–32 (2021)
9. Okonkwo, C.W., Ade-Ibijola, A.: *C*hatbots Applications in Education: A Systematic Review. *Computers and Education: Artificial Intelligence 2*, 100033 (2021). DOI: 10.1016/j.caeai.2021.100033
10. Finne, Å., Grönroos, C.: Rethinking marketing communication: From integrated marketing communication to relationship communication. In: *The Evolution of Integrated Marketing Communications*, pp. 105–121. Routledge, London (2013)

11. Khoa, B.T., Huynh, L.T., Nguyen, M.H.: The relationship between perceived value and peer engagement in sharing economy: A case study of ridesharing services. *Journal of System and Management Sciences* 10(4), 149–172 (2020)
12. Lee, D.H., Park, C.W.: Conceptualization and measurement of multidimensionality of integrated marketing communications. *Journal of Advertising Research* 47(3), 222–236 (2007)
13. Yoldaş, E.N., Aycı, A.: The role of artificial intelligence in integrated marketing communication: An evaluation of ChatGPT. Pamukkale Üniversitesi İşletme Araştırmaları Dergisi 11(2), 611–637 (2024)
14. Lingling, W., Fuli, C.: Role of AI technology in brand building of Chinese higher education institution – Thought based on integrated marketing communication. *Маркетинг і цифрові технології* 5(2), 7–13 (2021)
15. Lytvynchuk, O.: The Impact of Artificial Intelligence on Chatbot-Driven Personalization in Marketing Communications. *Global Prosperity* **5**(1), 73–87 (2025). DOI: 10.5281/zenodo.15720486
16. Werenowska, A., Jaska, E.: Using Chatbots in Marketing Communication. In: *Proceedings of the International Scientific Conference Economic Sciences for Agribusiness and Rural Economy*, no. 3, Warsaw, 5–7 June 2019, pp. 153–158 (2019). DOI: 10.22630/ESARE.2019.3.20
17. Labib, E.: Artificial intelligence in marketing: Exploring current and future trends. *Cogent Business & Management* 11(1), 2348728 (2024)
18. Danish, F., Akhtar, N., Imran, M.: AI-driven personalization in educational marketing: A framework for enhancing student recruitment and retention. *Journal of Political Stability Archive* 3(2), 559–590 (2025)
19. Brobbey, E.E., Ankrah, E., Kankam, P.K.: The role of artificial intelligence in integrated marketing communications: A case study of Jumia Online Ghana. *Inkanyiso* 13(1), 120–136 (2021)
20. Fageeh, A.: The rise of chatbots in higher education: Exploring user profiles, motivations, and integration strategies. *Social Sciences & Humanities Open* 12, 101996 (2025)
21. Al-Dmour, H., Al-Dmour, R., Al-Dmour, Y., Al-Dmour, A.: Transforming international student recruitment: The role of AI, personalization, and trust in Jordanian higher education. *Journal of International Students* 15(8), 25–52 (2025)
22. Gonda, D.E., Chu, B.: Chatbot as a Learning Resource? Creating Conversational Bots as a Supplement for Teaching Assistant Training Course. In: *2019 IEEE International Conference on Engineering, Technology and Education (TALE)*, pp. 1–5. IEEE (2019). DOI: 10.1109/TALE48000.2019.9225974
23. Skrebeca, J., Kalniete, P., Goldbergs, J., Pitkevica, L., Tihomirova, D., Romanovs, A.: *Modern Development Trends of Chatbots Using Artificial Intelligence (AI)*. In: *2021 62nd International Scientific Conference on Information Technology and Management Science of Riga Technical University (ITMS)*, pp. 1–6. IEEE (2021). DOI: 10.1109/ITMS52826.2021.9615334
24. Davar, N.F., Dewan, M.A.A., Zhang, X.: AI chatbots in education: Challenges and opportunities. *Information* 16(3), 235 (2025)
25. Cairovic Pavlovic, N., Pesic, Z.: University Faculty's Beliefs about the AI Use in Education. *Andragoške studije* 11, 85–113 (2025). DOI: 10.5937/AndStud2502085C

26. Dragojevic, T., Turudic, J.: The Use of ChatGPT in Higher Education: Students' Perceptions. *Nastava i vaspitanje* 73(3), 315–325 (2024). DOI: 10.5937/nasvas2403315D
27. Cunningham-Nelson, S., Boles, W., Trouton, L., Margerison, E.: A review of chatbots in education: Practical steps forward. In: *30th Annual Conference of the Australasian Association for Engineering Education (AAEE 2019)*, pp. 299–306. Engineers Australia (2019)
28. Ilieva, G., Yankova, T., Klisarova-Belcheva, S., Dimitrov, A., Bratkov, M., Angelov, D.: Effects of generative chatbots in higher education. *Information* 14(9), 492 (2023)
29. Baidoo-Anu, D., Ansah, L.O.: Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI* 7(1), 52–62 (2023)
30. Abdallah, N., Katmah, R., Khalaf, K., Jelinek, H.F.: Systematic review of ChatGPT in higher education: Navigating impact on learning, wellbeing, and collaboration. *Social Sciences & Humanities Open* 12, 101866 (2025)
31. Kanwar, A., Sanjeeva, M.: Student satisfaction survey: A key for quality improvement in the higher education institution. *Journal of Innovation and Entrepreneurship* 11(1), 27 (2022)
32. Tian, W., Ge, J., Zhao, Y., Zheng, X.: AI chatbots in Chinese higher education: Adoption, perception, and influence among graduate students – An integrated analysis utilizing UTAUT and ECM models. *Frontiers in Psychology* 15, 1268549 (2024)
33. Lyndyuk, A., Havrylyuk, I., Tomashevskii, Yu., Khirivskyi, R., Kohut, M.: *The Impact of Artificial Intelligence on Marketing Communications: New Business Opportunities and Challenges*. *Economics of Development* 23(4), 60–71 (2024). DOI: 10.57111/econ/4.2024.60
34. Labadze, L., Grigolia, M., Machaidze, L.: Role of AI chatbots in education: Systematic literature review. *International Journal of Educational Technology in Higher Education* 20(1), 56 (2023)
35. *Yeoh, C.W., Oh, Z.J.:* Integrating AI-Powered Digital Marketing Strategies for Enhancing Educational Communication and Technopreneurship. *International Journal of Academic Research in Business and Social Sciences* **15**(2), 1159–1176 (2025). DOI: 10.6007/IJARBSS/v15-i2/24835.
36. Stöhr, C., Ou, A.W., Malmström, H.: Perceptions and usage of AI chatbots among students in higher education across genders, academic levels and fields of study. *Computers and Education: Artificial Intelligence* 7, 100259 (2024)
37. Bilovodska, O., Kravchuk, T., Ponomarenko, I., Bliumska-Danko, K., Kononenko, A.: *Artificial Intelligence for Marketing Product Strategy in the Online Education Market*. *Economics of Development* 23(3), 18–31 (2024). DOI: 10.57111/econ/3.2024.18
38. Tarapata, J.*:* Marketing Communication of Modern Enterprises in the Era of Digital Economy. *Zeszyty Naukowe. Organizacja i Zarządzanie / Politechnika Śląska* 174, 355–370 (2023). DOI: 10.29119/1641-3466.2023.174.25

THEMATIC SECTION

Short papers

Resilient and Intelligent Organizations
in Times of Permanent Disruption

SYMORG 2026 • CONFERENCE PROCEEDINGS

SAP REINSURANCE MANAGEMENT MODULE CUSTOMIZATION FOR AUTOMATED INTEGRATION OF STRUCTURAL ELEMENTS

Katarina Mladenović[1*] [0009-0005-6156-1682], Ivan Jovanović[1] [0009-0003-1097-7720],
Srđa Bjeladinović[1] [0000-0001-6404-3497]
[1] University of Belgrade, Faculty of Organizational Sciences, Belgrade, Serbia
*Corresponding author, e-mail: katarina.mladenovic10@gmail.com

Abstract: *Enterprise Resource Planning (ERP) systems integrate different business processes, but their standard functionalities do not always meet all user-specific requirements. The need for automating the connection of structural elements was identified to reduce errors during data entry, and to provide more refined data as inputs to the system. Although SAP offers the possibility of purchasing and integrating Reinsurance Management (RM) module with other modules, and thus reduces input errors, this approach is not financially acceptable to certain companies. Therefore, this poster explores and demonstrates the possibilities of customizing the RM module using ABAP programming language, BAdI (Business Add-In) technology, and screen extensions to achieve the same goal.*

Keywords: *ERP systems; SAP systems; reinsurance module; system customization; ABAP programming language*

1. INTRODUCTION

Information systems are used by companies to support daily operations, improve performance, and achieve business objectives more effectively (Zimmermann, 2015). In many organizations, distinct business processes were supported by separate information systems, which often led to duplicate data and inconsistent processes. ERP (Enterprise Resource Planning) systems helped address these issues by integrating business activities within a single system environment (Bradford, 2014). However, their standard functionalities may still not fully match the requirements of each organization in practice (Jegorova, 2025). In today's business environment, organizations need to continuously adapt to the specific requirements of the market and their users (Sultana, 2025; Rahman et al., 2024).

SAP (Systems, Applications, and Products) is one of the leading ERP systems worldwide, used by companies across different industries to support a wide range of business needs. It includes a large number of modules that address different business processes (Ghulaxe, 2024).

In an industrial case involving the SAP Reinsurance module, users identified a need to automatically link structural elements to improve system management and ensure cleaner, more consistent data and process operations. SAP can support this through the introduction of additional modules and further system integration, but such approach entails additional implementation and financial costs (Sivaprasad et al., 2024).

In this paper, we present an approach to customizing the existing SAP Reinsurance Management (RM) module to achieve similar functionality while avoiding the cost of acquiring an additional module or functionality. Also, the proposed approach improves input data quality by automating structural elements through customization, thereby reducing the risk of user errors. In this paper, we discuss the customization of the SAP RM module through two identified use cases using ABAP programming language, Business Add-In (BAdI) technology, and screen extensions, in order to extend standard ERP functionality, improve data consistency across related business objects, and reduce organizational costs. The first identified use case is that users needed a clearer way to enter and store additional risk-related information for reinsurance treaties (i.e., contracts). Second, the Class of Business and Line of Business, which are standard structural elements, were entered manually prior to applying customization presented in this paper. This could lead to inconsistent data across treaties, accounts, and loss of documents. These issues may cause incomplete risk data, incorrect transaction classification, and additional manual work for users. The mentioned challenges differentiate the present work from

existing SAP enhancement practices by demonstrating how multiple enhancement mechanisms can be combined in a specific SAP RM scenario to improve data consistency and reduce manual input errors without implementing an additional module.

2. BACKGROUND

Organizations still need to balance the use of standard functionalities with customization when adapting the system to their needs (Patange, 2025). Fit-Gap analysis is used to identify differences between the functionalities offered by ERP systems and users' specific needs (Parthasarthy, 2007; Ganesh et al., 2014). Sultana (2025) higlights that ABAP based development enables effective customization of the system to meet these requierements. Prior work frames SAP customization as a response to the limitations of standard ERP functionality. Sivaprasad et al. (2024) show that ABAP, enhancements, and user exits can extend standard SAP Sales and Distribution pricing and billing processes, while Patange (2025) contrasts traditional extension mechanisms, such as user exits and enhancement spots, with predefined extension points aimed at preserving system integrity and upgrade compatibility. Schwaninger (2011) explains how User Exits and BAdIs can extend SAP functionality without changing the standard code. The same idea is applied in this paper to the SAP RM module. The reviewed literature mainly considers these mechanisms in general SAP customization contexts or in other SAP modules. However, in the available literature, we were not able to identify studies addressing the specific combination of customization approaches presented in this paper. The proposed approach extends SAP RM through additional risk-data screens and the automated linking of Class of Business (CoB) and Line of Business (LoB) across treaty, account, and loss levels.

3. SAP ENHANCEMENTS MECHANISMS

The SAP system can be extended to meet user requirements using different approaches. In this paper, we focus on enhancements, which are extensions of standard SAP functionality that do not affect the standard system (Fewer, 2006). Program exit, one of the enhancement approaches, is a point in the code that connects the standard SAP program with user-defined code. This approach allows users to implement additional functionality in the program without modifying the standard SAP code. It is typically implemented using function modules, subroutines, and methods (BAdIs). BAdI is an object-oriented approach for customizing SAP systems using classes, interfaces, and methods (Fewer, 2006). The system allows the use of BAdI technology and other enhancement approaches to extend standard functionality (Agrawal, 2012). Another enhancement approach is the screen exit, which allows adding new fields to the screen using function modules and BAdIs (SAP AG, 2007).

A screen is a user interface that enables users to perform various functions in the system. Screens serve as containers for different elements such as checkboxes, buttons, and input fields (SAP AG, 2023). The screen's business logic is divided into two phases: PBO (Process Before Output) and PAI (Process After Input). The fields on the screen are linked to global variables and are populated during the PBO phase. When a user changes a field on the screen, the process transitions to the PAI phase (SAP AG, 2007).

4. IMPLEMENTATION OF CUSTOMISATION OF SAP REINSURANCE MANAGEMENT

Based on conducted research and analysis of users' daily needs, functionalities were developed to facilitate work in the SAP system. These functionalities enable the storage of additional data, eliminate the need for an additional module implementation, and automatically populate fields to minimize human error. As a result, features that were not available in the standard SAP system are now available to users. The ABAP programming language was used to develop these functionalities. The solutions were implemented in SAP GUI and SAP Fiori environments.

Table 1: Use Cases for System Enhancement

Use case	Use case description	Implementation method
Use case 1 (UC1)	Extending the treaty data model with additional risk- related data.	The risk information screen is developed in a custom ABAP program and integrated via the corresponding BAdI to extend treaty functionality.
Use case 2 (UC2)	Automatic linking and population of CoB and LoB fields across treaty, invoice, and claim levels.	To implement automatic linking, a custom table defining all valid combinations of CoB and LoB is created. Depending on the business object, the appropriate BAdI or feeder class is used for implementation.

4.1. UC1: Design and Implementation of a New Screen for Storing Additional Risk Information

A facultative treaty in reinsurance is a specific type of treaty that is often highly complex, as it covers risks with a higher potential for loss. The standard SAP Reinsurance module does not fully support facultative treaty

management, which makes business processing more difficult. Prior to this implementation, users were unable to store additional risk-related information, which led to the need to develop a solution. This solution avoided the need to implement an additional module, which would have increased overall implementation costs. For this purpose, a new screen has been developed and linked to the main treaty screen. This screen is available only for facultative treaties.

New fields have been added to the screen to enter basic information about the insured person, as well as a table to store insured sum, currency, and change date data. Almost all fields on the screen are marked as mandatory. Creating a new screen involves identifying the BAdI used to call the screen in the treaty and linking it to the custom screen created to store risk data. The BAdI definition used is R_V_VTG_BADI, and the method implemented in the class that implements the BAdI interface is PUT_DATA_TO_SCREEN_9026, which enables access to the treaty screen for customization. In addition, a custom class, module pool program, structures, and tables were created to store data on insured risks. Validations were implemented to prevent the storage of incorrectly entered data. Figure 1 shows the result displayed in the User Exit Subscreen tab, which extends the basic functionality and contains a screen with elements for entering and displaying data.

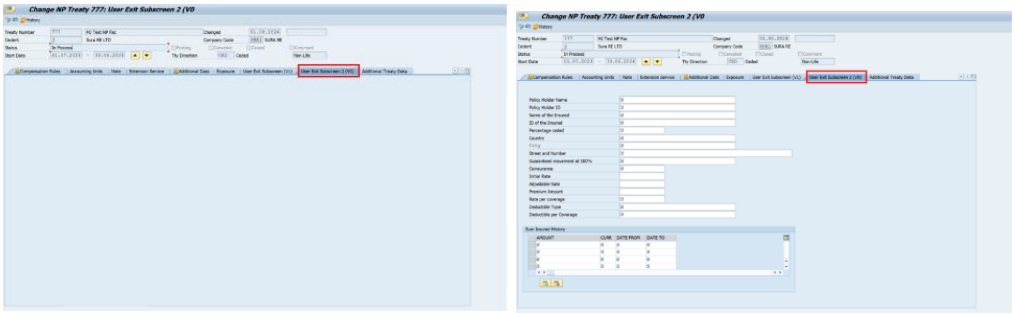


Figure 1: Results displayed in the User Exit Subscreen tab: a) Treaty subscreen before customization b) Subscreen for entering additional risk-related data

4.2. UC2: Automatic Linking of Class of Business and Line of Business Fields at Treaty, Account, and Insured Risk Levels

During manual data entering while system testing, discrepancies were identified between Business Class and Business Line. To address this issue, automatic population of these fields was implemented at contract, account, and insured risk levels. Although the issue was the same across all three objects, the solution was implemented differently depending on the system environment (SAP GUI or SAP Fiori) and underlying business processes. The corresponding mappings are defined in a custom table created for this purpose. This approach reduces the risk of incorrect data entry and shortens the time required for creating treaties, accounts, and insured risk records.

Depending on the business object on which the automatic connection is implemented, the corresponding BAdI is identified and a method is implemented to ensure the required changes. For the treaty object, the BAdI definition R_V_VTG_BADI is used, as in the previous use case, with the PREPARE_SAVE_VTG method implemented. The main difference from the previous use case lies in the method used, while the BAdI remains the same since the extension is applied to the same treaty object. In the previous case, the method was used to connect the main screen to the custom screen, whereas in this case the method is triggered during treaty saving. With this implementation, users are able to correctly maintain the mandatory Class of Business and Line of Business fields when creating a treaty. An additional validation step generates an error message if an incorrect value is entered and automatically corrects the value. In addition, values in the Line of Business field are automatically cleared when the corresponding Class of Business is removed, ensuring data consistency.

The issue with the account is in the postings section, where the system automatically overrides the main Line of Business defined in the treaty, disregarding the Class of Business selected by the user. For this use case, the class ZCL_IM_RI_A_ABR is used, which implements the interface /MSG/IF_EX_R_A_ABR_BADI. To create an automatic link on the account, the BEFORE_SAVING method from the interface is implemented. As a result of this implementation, the user will, in the future, select only the Class of Business, while the Line of Business will be generated based on the previously selected Class of Business. The Loss module is a component of reinsurance that handles the insured's reporting of loss events. The issue concerns the Line of Business field, which is not required in the FGU Balance section. To implement automatic connection at the loss level, a Feeder class connects the SAP Fiori interface to the backend business logic. For this purpose, the DO_CHECK method is implemented.

After implementing a combination of the customization techniques proposed in this paper, the specified use cases were tested. The results in the tested environment were compared to the previous results, before customization. Based on the conducted testing and internal evaluation, it was estimated that the implementation of the proposed solution reduced contract, account, and loss data entry time by approximately 20% and decreased the number of manual errors by approximately 35%. The results were obtained in the testing

environment but can be treated as non-strict indications of possible gains in the production environment, which will be tested as part of future research.

5. CONCLUSION

In this poster, we present how the SAP Reinsurance Management module can be extended when standard ERP functionality does not fully meet user requirements. The manuscript has explored customization options to automatically connect structural elements as an alternative to introducing a new module. The selected use cases show how screen exits and BAdIs can add new functionality without modifying the standard SAP code.

During the system customization process, the screen exit approach was used to implement additional screen functionality. This example demonstrates the possibility of combining a standard SAP screen with a custom-developed screen without modifying the core system code. A BAdI was used to connect the standard and custom screens, illustrating the integration of different enhancement mechanisms. In addition, the second use case demonstrates the use of a BAdIs and a feeder class to automatically link and populate fields between different objects. The key difference between these objects lies in the technical implementation, as different objects and environments require different enhancement approaches. This approach enables the system to meet user requirements while ensuring that the core code remains unchanged and future upgrades are not affected. The main contribution of this poster is to show that different SAP enhancement mechanisms can be combined in a specific SAP RM scenario to extend functionality, improve data consistency, and reduce manual input errors as a fraction of the cost of buying the new SAP module.

Although customization requires additional effort, it makes the system easier to use and may help reduce costs. Future work should focus on optimizing these solutions and examining how other SAP customization mechanisms affect system performance. Furthermore, other techniques for achieving similar goals will be explored, potentially in other modules. In addition, future work may include conducting quantitative evaluations to measure the impact of the proposed extensions. This could involve comparing the cost savings from customization, in relation to production systems or previously assessed solutions, rather than acquiring an additional module, as well as analyzing reductions in user errors and improvements in process performance.

Disclaimer. This paper includes results obtained during Katarina's master thesis research at the Faculty of Organizational Sciences, University of Belgrade, Belgrade, Serbia.

REFERENCES

[1] Agrawal, A. (2012). Customizing materials management processes in SAP ERP (2nd ed.). Galileo Press.
[2] Bradford, M. (2014). Modern ERP: select, implement, and use today's advanced business systems. Lulu.com.
[3] Fewer, B. (2006). SAP ABAP interview questions, answers, and explanations: ABAP certification review. Equity Press.
[4] Ganesh, K., Mohapatra, S., Anbuudayasankar, S. P., & Sivakumar, P. (2014). Enterprise resource planning: fundamentals of design and implementation. Springer.
[5] Ghulaxe, V. (2024). SAP: Comprehensive industrial process solutions for businesses. Research Journal of Engineering and Technology, 15(1), 13-24.
[6] Jegorova, A. (2025). Resolving system-organisational misfits: development and assessment of a misfit resolution framework for off the-shelf ERP systems. In Joint Proceedings of the BIR 2025 Workshops and Doctoral Consortium, co-located with 24th International Conference on Perspectives in Business Informatics Research (BIR 2025) (Vol. 4034, pp. 61-69).
[7] Parthasarthy, S. (2007). Enterprise resource planning (ERP): A managerial and technical perspective. New Age International.
[8] Patange, P. P. (2025). Empowering Extensibility in SAP S/4HANA: A Study on the'Custom Fields and Logic'Fiori App and Its Role in Technical Consulting. Journal of Computer Science and Technology Studies, 7(11), 30-36.
[9] Rahman, M. A., Bhowmik, J., Ahamed, M. S., & Rahman, R. (2024). Opportunities and challenges in data analysis using SAP: a review of ERP software performance. Global Mainstream Journal, 1(4), 50-67.
[10] SAP AG. (2007). Enhancement Framework. SAP Documentation.
[11] SAP AG. (2023). ABAP Workbench Fundamentals - Part 2. SAP Documentation.
[12] Schwaninger, J. (2011). ABAP Development for Materials Management in SAP: User Exits and BAdIs. Galileo Press.
[13] Sivaprasad, N., Archit, J., Shalu, J., Krishna, K. T., & Akshun, C. (2024). Advanced techniques in SAP SD customization for pricing and billing. Innovative Research Thoughts , 9(1), 421-449.
[14] Sultana, T. (2025). Customizing SAP: an internship journey in SAP development and customization of SAP ABAP and HCM module at Eitekh ERP Limited (Doctoral dissertation, Brac University).
[15] Zimmermann, A. S. (2015). Evolving enterprise architectures for digital transformations (pp. 183-194). Gesellschaft für Informatik.

MOTIVATION FOR DIGITALIZATION IN AGRICULTURE: SIGNIFICANCE AND OPPORTUNITIES FOR ENHANCEMENT IN THE SOUTHERN GREAT PLAIN REGION OF HUNGARY

Patrik Urbán* [0009-0007-6587-8227]
PhD student, University of Szeged, Faculty of Economics and Business Administration, Hungary
*Corresponding author, e-mail: urbangyorgypatrik@gmail.com

Abstract: *This study examines how motivation, trust, knowledge transfer and education shape agricultural digitalization in Hungary's Southern Great Plain region. The analysis is based on eight semi-structured expert interviews conducted in 2026 with agricultural advisory, institutional, producer-organization and educational actors. The findings show that digitalization is progressing unevenly: larger, capital-intensive, horticultural and greenhouse farms are more advanced, while smaller family farms face capital constraints, weak digital skills, long payback periods and delayed generational renewal. Digitalization is currently driven mainly by external pressures, including subsidies, rising input prices, drought risk, labour shortages and administrative obligations. However, sustained use requires visible results, technological and interpersonal trust, practical advisory support and targeted competence development. The paper argues that support systems should combine funding with field demonstrations, reference farms, mentoring and education-oriented knowledge transfer.*

Keywords*: Agricultural digitalization, Motivation, Trust, Knowledge transfer, Precision farming*

1. INTRODUCTION

Digitalization is reshaping agricultural production, decision-making and farm management (Schwab, 2016; Kraus et al., 2021). Precision technologies such as GPS-based guidance, sensors, drones, yield mapping, automated machinery and data-driven decision-support systems can improve productivity, reduce input use and support more sustainable farming (Gebbers & Adamchuk, 2010; Zhang & Kovacs, 2012; Parviainen et al., 2017). These tools are particularly relevant in regions where drought, volatile input prices, labour shortages and climate risks increase pressure on farms.

The Southern Great Plain region of Hungary is a relevant empirical setting because of its agricultural importance, drought sensitivity and heterogeneous farm structure. In this context, digitalization is not only a modernization opportunity but also a competitiveness and adaptation issue. Previous literature suggests that adoption is shaped not only by capital and infrastructure, but also by digital skills, motivation, trust and knowledge transfer (Ryan & Deci, 2000; Dessart et al., 2019; Mokhtar et al., 2022; Rupeika-Apoga & Petrovska, 2022). Accordingly, this paper asks how motivation, trust and knowledge transfer influence agricultural digitalization in the Southern Great Plain.

2. LITERATURE REVIEW

Digitalization is a central element of the Fourth Industrial Revolution, transforming value creation, business models and organizational processes (Kagermann et al., 2013; Schwab, 2016). At the firm level, digital tools may contribute to cost reduction, process control, resource efficiency and faster decision-making (Bharadwaj et al., 2013; Parviainen et al., 2017). In agriculture, precision farming supports site-specific decisions and more accurate responses to soil, crop and weather variability (Gebbers & Adamchuk, 2010; Zhang & Kovacs, 2012).

However, agricultural digitalization often spreads more slowly than technological potential would suggest. Earlier studies identify barriers such as high investment costs, weak infrastructure, limited digital literacy, insufficient advisory support and cautious attitudes toward innovation (Barnes et al., 2019; Mokhtar et al., 2022;

Rupeika-Apoga & Petrovska, 2022). Motivation is therefore treated here as an activating condition: capital, technology and training do not automatically lead to adoption if farmers do not perceive digital tools as useful, necessary or relevant to their own production problems (Ryan & Deci, 2000; Fishbach & Woolley, 2022).

Trust is another central factor. Farmers need technological trust in the reliability of digital systems, interpersonal trust in advisors and providers, and institutional trust in funding, data management and support structures (Mayer et al., 1995; Zucker, 1986; Riegelsberger et al., 2005). Behavioural economics also explains why adoption may remain limited even when long-term benefits appear rational: bounded rationality, status quo bias, uncertainty avoidance and the overvaluation of short-term costs can slow technological change (Simon, 1955; Thaler & Sunstein, 2008; Thaler, 2016). Knowledge transfer, field demonstrations, mentoring and well-timed incentives may therefore reduce perceived risk and support adoption (Duflo et al., 2011; Dessart et al., 2019; European Commission, 2019).

3. RESEARCH METHODOLOGY

The study applies a qualitative research design based on eight semi-structured expert interviews conducted in 2026. The interviewees included agricultural chamber and village advisor network representatives, institutional experts, a producer-organization leader, and professionals from agricultural vocational and higher education. Their advisory, educational and institutional positions provided regional insight into digitalization patterns, producer-level barriers and knowledge-transfer mechanisms.

The sample was selected through purposive expert sampling, focusing on actors with direct professional insight into agricultural digitalization in the Southern Great Plain. The interview guide was developed from the literature on precision agriculture, motivation, trust, technology acceptance and knowledge transfer (Mayer et al., 1995; Ryan & Deci, 2000; Riegelsberger et al., 2005; Dessart et al., 2019). The interviews explored the regional status of digitalization, technologies used, motivational drivers, demotivating factors, trust-related issues, advisory systems, education and generational renewal.

The material was analysed through thematic qualitative content analysis. First, open coding was used to identify themes related to motivation, demotivation, trust, digital skills, knowledge transfer, education and generational renewal. Second, these codes were grouped into broader analytical categories to identify recurring patterns across interviews. The study does not aim at statistical representativeness; its purpose is to provide analytical insight into the regional mechanisms shaping agricultural digitalization.

4. RESULTS AND DISCUSSION

The interviews show that agricultural digitalization in the Southern Great Plain is progressing, but unevenly. Larger, capital-intensive and better organized farms, especially horticultural and greenhouse producers, are more advanced in adopting digital solutions. For them, digitalization often appears as a competitiveness requirement. In contrast, many smaller family farms remain only partially digitalized or not digitalized at all. Their main barriers include limited capital, weak digital skills, long payback periods and delayed generational renewal.

A key finding is that digitalization is currently driven more by external than internal motivation. Subsidy opportunities, rising input prices, labour shortages, administrative obligations, climate pressure and the consequences of the 2022 drought all increase the perceived relevance of digital tools. However, these external pressures do not necessarily create long-term intrinsic commitment. Subsidies may trigger investment, but they do not guarantee sustained use if farmers lack skills, advisory support or clear evidence of usefulness.

Visible results and local references are particularly important. The interviews suggest that farmers are more likely to trust a technology when they can observe its benefits in a comparable local environment. The 'neighbour effect' may motivate more strongly than abstract professional arguments: if a nearby farmer works more efficiently or achieves better results with digital tools, adoption becomes more credible. Agricultural digitalization is therefore not only an information problem, but also a social and behavioural process.
Trust emerged as a differentiated phenomenon. Farmers tend to accept mechanical or directly observable technologies more easily, such as guidance systems, automatic steering or sensors, because their effects can be checked quickly. More complex software-based or data-management systems raise greater caution, especially when data security, return on investment or system transparency is unclear. Thus, trust is closely linked to visibility, usability and the credibility of the person or institution recommending the technology.

Knowledge transfer is one of the strongest practical findings. Conferences, study visits, product demonstrations, advisory systems and demonstration farms can support adoption if they are practical and connected to real farm-level problems. Farmers need technologies to be shown, tested, explained and followed up in their own production context. Education also has a strategic role: agricultural vocational and higher education can shape the mindset of future farmers, advisors and professionals by integrating precision farming, robotics, drone use and data-based farm management into training.

5. CONCLUSION

The study concludes that agricultural digitalization in the Southern Great Plain is not merely a technological or investment issue. It is a complex adaptation process shaped by motivation, trust, skills, farm structure, generational renewal, education and credible professional mediation. Digitalization may become a development opportunity for better prepared farms, but it may also widen structural gaps if smaller or less skilled farms do not receive targeted support.

The practical implication is that advisory and support systems should not rely on financial incentives alone. Investment funding should be combined with field demonstrations, reference farms, mentoring, independent advisory support, practical training and follow-up. Digital benefits should be linked to concrete farm-level problems such as input saving, water management, labour organization, risk reduction and administrative simplification. For smaller farms, shared service models and regional mentor networks may lower the entry threshold. In this sense, digital motivation is built through visible results, trusted intermediaries and practical learning experiences.

REFERENCES

[1] Barnes, A. P., Soto, I., Eory, V., Beck, B., Balafoutis, A. T., Sánchez, B., Vangeyte, J., Fountas, S., Van der Wal, T. and Gómez-Barbero, M. (2019). Influencing incentives for precision agricultural technologies within European arable farming systems. Environmental Science & Policy, 93, 66-74.

[2] Bharadwaj, A., El Sawy, O. A., Pavlou, P. A. and Venkatraman, N. (2013). Digital business strategy: Toward a next generation of insights. MIS Quarterly, 37(2), 471-482.

[3] Dessart, F. J., Barreiro-Hurlé, J. and Van Bavel, R. (2019). Behavioural factors affecting the adoption of sustainable farming practices: A policy-oriented review. European Review of Agricultural Economics, 46(3), 417-471.

[4] Duflo, E., Kremer, M. and Robinson, J. (2011). Nudging farmers to use fertilizer: Theory and experimental evidence from Kenya. American Economic Review, 101(6), 2350-2390.

[5] European Commission. (2019). A smart and sustainable digital future for European agriculture and rural areas. Brussels: European Commission.

[6] Fishbach, A. and Woolley, K. (2022). The structure of intrinsic motivation. Annual Review of Organizational Psychology and Organizational Behavior, 9, 339-365.

[7] Gebbers, R. and Adamchuk, V. I. (2010). Precision agriculture and food security. Science, 327(5967), 828-831.

[8] Kagermann, H., Wahlster, W. and Helbig, J. (2013). Recommendations for implementing the strategic initiative INDUSTRIE 4.0: Final report of the Industrie 4.0 Working Group. Frankfurt: Acatech.

[9] Kraus, S., Schiavone, F., Pluzhnikova, A. and Invernizzi, A. C. (2021). Digital transformation in healthcare: Analysing the current state-of-research. Journal of Business Research, 123, 557-567.

[10] Mayer, R. C., Davis, J. H. and Schoorman, F. D. (1995). An integrative model of organizational trust. Academy of Management Review, 20(3), 709-734.

[11] Mokhtar, W. N. H. W., Tengku Izhar, T. A., Zaini, M. K. and Hussin, N. (2022). The importance of digital literacy skills among farmers for sustainable food security. International Journal of Academic Research in Progressive Education and Development, 11(1). https://doi.org/10.6007/IJARPED/v11-i1/12104

[12] Parviainen, P., Tihinen, M., Kääriäinen, J. and Teppola, S. (2017). Tackling the digitalization challenge: How to benefit from digitalization in practice. International Journal of Information Systems and Project Management, 5(1), 63-77.

[13] Riegelsberger, J., Sasse, M. A. and McCarthy, J. D. (2005). The mechanics of trust: A framework for research and design. International Journal of Human-Computer Studies, 62(3), 381-422.

[14] Rupeika-Apoga, R. and Petrovska, K. (2022). Barriers to sustainable digital transformation in micro-, small-, and medium-sized enterprises. Sustainability, 14(21), 13577.

[15] Ryan, R. M. and Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. Contemporary Educational Psychology, 25(1), 54-67.

[16] Schwab, K. (2016). The Fourth Industrial Revolution. Geneva: World Economic Forum.

[17] Simon, H. A. (1955). A behavioral model of rational choice. Quarterly Journal of Economics, 69(1), 99-118.

[18] Thaler, R. H. (2016). Behavioral economics: Past, present, and future. American Economic Review, 106(7), 1577-1600.

[19] Thaler, R. H. and Sunstein, C. R. (2008). Nudge: Improving decisions about health, wealth, and happiness. New Haven: Yale University Press.

[20] Zhang, C. and Kovacs, J. M. (2012). The application of small unmanned aerial systems for precision agriculture: A review. Precision Agriculture, 13(6), 693-712.

[21] Zucker, L. G. (1986). Production of trust: Institutional sources of economic structure, 1840-1920. Research in Organizational Behavior, 8, 53-111.

ACKNOWLEDGMENT / AI USE DISCLOSURE

The author acknowledges the use of ChatGPT by OpenAI for language editing, proofreading, improving readability, and refining the academic style of selected parts of the manuscript. The tool was not used to generate empirical data, fabricate or modify results, conduct autonomous analysis, create citations, or replace the author's scholarly judgment. All AI-assisted edits were reviewed, verified, and approved by the author. The author takes full responsibility for the final content, arguments, citations, interpretation, and conclusions of the manuscript.

Original Research Article *https://doi.org/10.5281/zenodo.23397161*

QUO VADIS, HUMANITAS? MAPPING THE HUMAN COST OF ENVIRONMENTAL DEGRADATION THROUGH THE SDG FRAMEWORK

Nataša Petrović[*1] [0000-0002-9576-7102], Jelena Andreja Radaković[2] [0000-0001-8771-8215], Aleksandar Jovović[3] [0000-0003-2294-5729], Boris Najman[4] [0000-0003-2632-556X], Polona Šprajc[5] [0000-0003-1356-3438]

[1] University of Belgrade – Faculty of Organizational Sciences, Belgrade, Serbia
[2] University of Belgrade – Faculty of Organizational Sciences, Belgrade, Serbia
[3] University of Belgrade – Faculty of Mechanical Engineering, Belgrade, Serbia
[4] University Paris-East Créteil, Créteil, France
[5] University of Maribor – Faculty of Organizational Sciences, Kranj, Slovenia

*Corresponding author, e-mail: natasa.petrovic@fon.bg.ac.rs

Abstract: *The paper presents the environmental dimension of Sustainable Development (SD) through an analysis of the Sustainable Development Goals (SDGs), their corresponding environmental pressures, environmental impacts, and human cost. From the analysis it is clearly evident that there is a discrepancy between international agreements and actual results. For these reasons, the paper emphasizes that the environmental dimension of SD is not only an ecological problem, but a serious developmental challenge that affects all peoples, societies as a whole, and the very survival of our civilization.*

Keywords: *Environmental pillar, environmental limits, Sustainable Development, Sustainable Development Goals*

1. INTRODUCTION

It is a naive misconception that humans began to exert a negative influence on the environment only at the start of the Industrial Revolution and the development of modern technology. The first significant environmental impacts were the discovery of agriculture (8000-7500 B.C.) and the domestication of animals (7000 B.C.). These two stages in the history of human civilization brought profound changes to the animal and plant worlds across humanity's habitats (e.g., the clearing of large areas of forest to create agricultural land and to feed the ever-growing human population) (Petrović, 2024).

The negative impact of the human species on nature has continued throughout history, and the poor relationship between humans and the environment persists despite all agreements, declarations, and agendas. Unfortunately, according to a UN global assessment report, the devastation of nature has increased over the past few decades (Petrović et al., 2024; Watts, 2021). Thus, “global warming, climate change, and natural disasters pose substantial and tragically expanding challenges for the development of societies across the globe and to the very sustainability of the planet” (Radaković et al., 2025).

The growing number of global environmental problems is often interpreted as a progressive loss of ecological stability. This interpretation assumes that the cumulative effects of environmental crises can lead to direct or indirect disruption of the ultimate boundaries that is the ecosystem of planet Earth. Unfortunately, humanity is increasingly approaching those limits. In the last decades of the 20th century, Sustainable Development (SD) became the new paradigm for human development, simply because the previous paradigm, which focused on economic development without regard for the environment and social structures, was unsustainable.

2. SD AND SDGS

The concept of SD began to gain stronger international recognition during the 1980s, particularly “after the establishment of the World Commission on Environment and Development in 1983. The Commission published the report *Our Common Future* in 1987.” (Radaković et al., 2020). The report emphasised the effects of unrestrained population and economic expansion, as well as the need for a new development model. In the report, “SD is defined as development that meets the needs of the present without compromising the ability of future generations to meet their own needs” (World Commission on Environment and Development [WCED], 1987). This definition underlines all efforts of international policies and laws that pertain not only to environmental protection but also to the global sustainability of societies (Borojević et al., 2017). The interdisciplinarity of the concept of SD is reflected in its interconnected pillars: “economic growth, environmental protection, and social equality” (WCED, 1987).

“The new millennium started with the United Nations Millennium Declaration (United Nations General Assembly, 2000), which called on world leaders to address poverty, hunger, sickness, illiteracy, environmental degradation, and gender inequality. Millennium Development Goals (MDGs), which lasted from 2000. until 2015.” (United Nations General Assembly, 2000) achieved significant success in reducing extreme poverty, but it had several key shortcomings that the new Agenda 2030 was meant to address. Globally, the UN's Agenda for Sustainable Development, announced in 2015, is the principal political instrument for addressing these problems. The Agenda sets 17 Sustainable Development Goals (SDGs) that highlight the interdependence of the social, economic, and environmental elements of progress (United Nations, 2015).

3. RESULTS AND DISCUSSION

In this paper, SDG 6, SDG 13, SDG 14, and SDG 15 are treated as core environmental goals, in line with the biosphere logic of the SDGs “wedding cake” model (Stockholm Resilience Centre, 2016). SDG 3, SDG 7, SDG 11, SDG 12, and SDG 16 are included as environmentally related goals because their targets address health impacts of pollution, clean energy, sustainable cities, responsible consumption, and the environmental consequences of conflict and weak governance (Stockholm Resilience Centre, 2016; United Nations, 2015; United Nations, n.d.).

Table 1 provides a systematic mapping of chosen environmental pressures and their negative environmental impacts, human costs, and their alignment with relevant SDGs.

Table 1: The Human Cost of Environmental Pressures through the SDG Framework

Environmental pressure	Negative environmental impact	Human cost	Relevant SDG
Air pollution	Emissions of pollutants (e.g., PM2.5, NO_x, SO_2... from transport, industry, energy production, and households)	Diseases, premature mortality, decline in quality of life, impacts on vulnerable population (Indicator: e.g., ~6.7 million premature deaths/year coused by air pollution)	SDG 3, SDG 11, SDG 13
Water pollution	Discharges of wastewater (e.g., with pesticides, heavy metals, industrial pollutants etc.)	Waterborne diseases, exposure to toxins, mortality, limited access to safe drinking water and sanitation (Indicator: e.g., 3.4–3.5 million deaths per year from water-related diseases)	SDG 3, SDG 6, SDG 14
Land pollution	Solid and hazardous waste (e.g., plastics, residuals from industry)	Exposure to toxins and contaminated food, health risks (Indicator: e.g., projected generation of 3.8 billion tonnes of municipal solid waste by 2050)	SDG 3, SDG 11, SDG 12, SDG 15
Climate changes	Greenhouse gases, global warming, extreme weather, and natural disasters (floods, droughts and heatwaves)	Diseases and mortality, food and water insecurity, climate migration, disaster risks (Indicator: e.g., the death rate from heat waves with an average of 546,000 deaths per year)	SDG 3, SDG 7, SDG 11, SDG 13
Land degradation	Soil erosion, deforestation, desertification, and loss of soil fertility	Decreasing of food security, rural poverty, migration, loss of livelihoods, vulnerability to disaster risks (Indicator: e.g., 3 billion people affected)	SDG 2, SDG 12, SDG 13, SDG 15
Loss of biodiversity	Habitat destruction, species extinction, and loss of genetic diversity	Diseases risks, loss of ecosystem services, reduced food and medicine resources, reduced resilience (Indicator: e.g., 1 million animal and plant species threatened with extinction)	SDG 3, SDG 14, SDG 15
Chemical pollution	Pesticides, organic pollutants, PFAS, heavy metals, and hazardous chemicals	Diseases (Indicator: e.g., about 9 million deaths annually)	SDG 3, SDG 6, SDG 12, SDG 15

Environmental pressure	Negative environmental impact	Human cost	Relevant SDG
Water scarcity	Overuse of freshwater resources, groundwater depletion, droughts and unequal access to water.	Diseases, mortality risk, food insecurity, conflicts, migration (Indicator: e.g., 2.2 billion people without safe access to drinking water)	SDG 3, SDG 6, SDG 13, SDG 16

**Source: Author's own elaboration based on (Bohn et al., 2025), Food and Agriculture Organization of the United Nations (2022), Intergovernmental Panel on Climate Change (2023), Persson et al., (2022), United Nations Environment Programme (2021), World Health Organization (2023, 2026a, 2026b, 2026c, 2026d).*

4. CONCLUSION

From the report *Sustainable Development Goals Report 2025* (United Nations, 2025), it is apparent that there is an alarming inadequacy in the progress made towards achieving the SDGs, only five years before the 2030 goal. Just 35% of the SDG targets are either on track or show signs of making moderate progress, whereas almost 50% are falling behind and 18% are worse than their 2015 baseline. The most pressing issue here is the achievement of the environmentally related SDGs mentioned above, which are regressing due to environmental degradation, climate change, pollution, etc. The authors of the paper note that the environmental issue of the SD is not only an ecological problem, but also has a direct impact on human well-being, societies, and the sustainability of our civilization. It is concluded that although Agenda 2030 and its SDGs provide an international framework for the very survival and long-term development of the planet and humanity, there is a significant gap between the agreements, commitments, and actual achievements.

REFERENCES

[1] Bohn, F. J., Bastos, A., Martin, R., Rammig, A., Koh, N. S., Sioen, G. B., Buscher, B., Carver, L., DeClerck, F., Drupp, M., Fletcher, R., Forrest, M., Gasparatos, A., Godoy-Faúndez, A., Hagedorn, G., Hänsel, M. C., Hetzer, J., Hickler, T., Krug, C. B., . . . Yoshida, Y. (2025). Reviews and syntheses: Current perspectives on biosphere research 2024–2025 – eight findings from ecology, sociology, and economics. *Biogeosciences, 22*(10), 2425–2460. https://doi.org/10.5194/bg-22-2425-2025

[2] Borojević, T., Maletič, M., Petrović, N., Radaković, J.A., Senegačnik, M., & Maletič, D. (2017). Youth Attitudes Towards Goals of a New Sustainable Development Agenda. *PROBLEMY EKOROZWOJU*, *12*(2), 161–172.

[3] Food and Agriculture Organization of the United Nations. (2022). *The state of the world's forests 2022: Forest pathways for green recovery and building inclusive, resilient and sustainable economies*. FAO. https://doi.org/10.4060/cb9360en

[4] Intergovernmental Panel on Climate Change. (2023). *Climate change 2023: Synthesis report*. IPCC. https://doi.org/10.59327/IPCC/AR6-9789291691647

[5] Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. (2019). Global assessment report on biodiversity and ecosystem services: Summary for policymakers. IPBES.

[6] Persson, L., Carney Almroth, B. M., Collins, C. D., Cornell, S., de Wit, C. A., Diamond, M. L., Fantke, P., Hassellöv, M., MacLeod, M., Ryberg, M. W., Søgaard Jørgensen, P., Villarrubia-Gómez, P., Wang, Z., & Hauschild, M. Z. (2022). Outside the safe operating space of the planetary boundary for novel entities. *Environmental Science & Technology, 56*(3), 1510–1521. https://doi.org/10.1021/acs.est.1c04158

[7] Petrović N., Radaković, J. A., & Šprajc, P. (2024, June 12–15). Educational Pathways Towards Sustainability: Student Literacy in Environmental Studies. In M. Kostić-Stanković, I. Mijatović & J. Krivokapić (Eds.) Proceedings of XIX International Symposium SymOrg 2024 *"Unlocking the Hidden Potentials of Organization Through Merging of Humans and Digitals"*, (pp. 181–187), Zlatibor, Serbia. University of Belgrade – Faculty of Organizational Sciences: Belgrade. ISBN 978-86-7680-464-1

[8] Radaković, J. A., Makajić-Nikolić, D., & Nikolić, N. (2025). Bridging the Gap: A Novel Approach to Flood Risk Assessment for Resilience. *Water*, *17*(13), 1848. https://doi.org/10.3390/w17131848

[9] Radaković, J. A., Petrović, N., Kručičan, A., Ćirović, M. & Tomić, A. (2020, September 23–24). Environmental sustainability and libraries: What do students really think? In P. Šprajc, I. Podbregar, D. Maletič, D. Tomić, Y. Ziegler, V. Uli & O. Arsenijević (Eds.), Proceedings of the 39th International Conference on Organizational Science *"Organizations at innovation and digital transformation roundabout"*, (pp. 605–616), Kranj, Slovenia, online, MS Teams. University of Maribor Press: Maribor. https://doi.org/10.18690/978-961-286-388-3.48

[10] Stockholm Resilience Centre. (2016, June 14). *The SDGs wedding cake*. Retrieved from https://www.stockholmresilience.org/research/research-news/2016-06-14-the-sdgs-wedding-cake.htm.

[11] United Nations Convention to Combat Desertification. (2025). Drought and desertification factsheet. UNCCD.

[12] United Nations Environment Programme. (2021). *Making peace with nature: A scientific blueprint to tackle the climate, biodiversity and pollution emergencies*. UNEP.

[13] United Nations Environment Programme. (2024). Global Waste Management Outlook 2024: Beyond an age of waste. UNEP.

[14] United Nations General Assembly. (2000). *United Nations Millennium Declaration* (A/RES/55/2). https://docs.un.org/en/a/res/55/2

[15] United Nations. (2015). *Transforming our world: The 2030 Agenda for Sustainable Development*. https://documents.un.org/doc/undoc/gen/n15/291/89/pdf/n1529189.pdf.

[16] United Nations. (2023). The Sustainable Development Goals Report 2023: Goal 6. United Nations.

[17] United Nations. (2025). *The Sustainable Development Goals report 2025*. United Nations. Retrieved from https://unstats.un.org/sdgs/report/2025/

[18] United Nations. (n.d.). *The 17 goals*. Department of Economic and Social Affairs, Sustainable Development. Retrieved from https://sdgs.un.org/goals.

[19] Watts, J. (2021, October 29). *Human society under urgent threat from loss of Earth's natural life*. The Guardian. Retrieved from https://www.theguardian.com/environment/2019/may/06/human-society-under-urgent-threat-loss-earth-natural-life-un-report?CMP=share_btn_url

[20] World Commission on Environment and Development (WCED). (1987). *Our common future*. Oxford University Press.

[21] World Health Organization. (2026a). *Climate change and health*. World Health Organization. Retrieved from https://www.who.int/news-room/fact-sheets/detail/climate-change-and-health

[22] World Health Organization. (2026b). *Air pollution*. World Health Organization. Retrieved from https://www.who.int/health-topics/air-pollution

[23] World Health Organization. (2026c). *Drinking-water*. World Health Organization. Retrieved from https://www.who.int/news-room/fact-sheets/detail/drinking-water

[24] World Health Organization. (2026d). *Chemical safety*. World Health Organization Regional Office for Europe. Retrieved from https://www.who.int/europe/health-topics/chemical-safety#tab=tab_1

[25] World Meteorological Organization (2023). *Atlas of mortality and economic losses from weather, climate and water-related hazards (1970–2021)* (WMO-No. 1313). World Meteorological Organization. Retrieved from https://wmo.int

Original Research Article *https://doi.org/10.5281/zenodo.23397263*

ARE CHINESE UNIVERSITIES ENTREPRENEURIALLY ORIENTED?

Zelimir Todorovic*[1][0000-0001-6662-2396]
[1]Doermer School of Business, Purdue University Fort Wayne, 2101 East Coliseum Blvd, Fort Wayne, IN, United States of America
*Corresponding author, e-mail: todorovz@purdue.edu

Abstract: *This study examines the Entrepreneurial Orientation (EO) of Chinese universities using the Entre-U scale developed by Todorovic et al. (2011). Survey data were collected from department chairs in science and engineering-oriented disciplines across Chinese universities, yielding 335 valid responses. Confirmatory Factor Analysis and Structural Equation Modeling were conducted using AMOS to validate the measurement model and examine the relationship between EO and university performance. The findings support the applicability of the Entre-U scale in the Chinese higher education context. However, most EO dimensions did not show significant relationships with performance measures, including patents and start-ups. Only industry collaboration demonstrated a statistically significant positive relationship with university entrepreneurial performance.*

Keywords*: Entrepreneurial Orientation, Entre-U Scale, Chinese entrepreneurial universities*

1. INTRODUCTION

In the CWTS Leiden Ranking (2025), Chinese universities have claimed the top two positions, surpassing Harvard University, which has slipped to third place. The CWTS's newest ranking has surprised many observers and raises important questions about the underlying drivers of university research performance. While existing explanations emphasize funding levels, institutional expansion, and national innovation policy, less attention has been given to the internal strategic orientation of universities themselves. In particular, it remains unclear whether Chinese universities exhibit a strong Entrepreneurial Orientation (EO)—reflected in innovativeness, proactiveness, and risk-taking—which may help explain their rapid rise in global research rankings.

Entrepreneurial Orientation is theorized to enhance organizational performance by fostering opportunity recognition, encouraging experimentation, and strengthening the capacity to transform knowledge into impactful outputs. In the context of higher education institutions, EO may translate into more active research mobilization, stronger industry collaboration, and more flexible institutional policies that support innovation and commercialization. As such, universities with higher EO are expected to achieve stronger performance outcomes, not only in traditional academic metrics such as publications and citations, but also in applied outputs such as patents, start-ups, and spin-offs.

Against this background, this study investigates whether Chinese universities are entrepreneurially oriented and whether such orientation is associated with measurable performance outcomes. Specifically, it examines whether EO—as operationalized through the Entre-U scale—helps explain variation in university performance across Chinese institutions. In the rest of this article, we provide a literature review of EO and the Entre-U scale development, followed by the research methodology and empirical results. The final discussion outlines the study's limitations and future research directions to further understand EO in higher education institutions.

2. LITERATURE REVIEW FOR ENTREPRENEURIAL ORIENTATION AND ENTRE-U SCALE

2.1 Entrepreneurial Orientation

Entrepreneurial Orientation (EO) is a central construct in entrepreneurship and strategic management research that explains how firms adopt entrepreneurial behaviors to achieve competitive advantage and improve performance (Miller, 1983). Miller (1983) defined an entrepreneurial firm as one that engages in product-market innovation, undertakes risky ventures, and is first to introduce proactive innovations in the

marketplace. Building on this foundational work, Covin and Slevin (1989) operationalized EO as a strategic posture reflecting a firm's tendency to be innovative, risk-taking, and proactive, and developed one of the most widely used measurement scales for empirical research. These three dimensions became the dominant conceptualization of EO in early research. Subsequent studies expanded the construct to a five- dimensional model of EO by adding competitive aggressiveness and autonomy to the original dimensions (Lumpkin & Dess, 1996). This multidimensional perspective has been widely adopted because it allows researchers to examine EO either as a unified construct or as separate dimensions depending on the research context. A large body of empirical research has examined the relationship between EO and firm performance. Studies consistently show that firms with stronger entrepreneurial orientation tend to achieve higher growth, innovation output, and competitive advantage (Rauch, Wiklund, Lumpkin, & Frese, 2009; Wiklund & Shepherd, 2005). More recent research has extended EO to new contexts, including international, digital, and corporate entrepreneurship. Scholars argue that EO enhances firms' ability to recognize and exploit emerging opportunities in rapidly changing markets, particularly in technology-driven industries (Covin & Wales, 2019). Additionally, EO has been linked to innovation capability, knowledge management, and strategic flexibility, which enable organizations to adapt to competitive pressures and technological disruptions. As a result, EO remains a key theoretical lens for understanding how firms develop entrepreneurial strategies and sustain long-term competitive performance.

2.2 Entre-U Scale

The transformation of traditional universities into entrepreneurial institutions has been widely discussed in the literature on innovation systems and higher education. Early theoretical foundations emphasize the interaction among universities, industry, and government in fostering innovation and economic development (Etzkowitz & Leydesdorff, 2000). The entrepreneurial university was conceptualized as an institution that strategically adapts its organizational structure, funding mechanisms, and culture to support innovation and entrepreneurial activities (Clark, 1989). While these early contributions provided important conceptual foundations, they did not offer standardized quantitative measures to empirically evaluate how entrepreneurial universities are.

To address this gap, Todorovic et al. (2011) developed the Entre-U scale, which became the most widely used instrument to measure the EO of a university. The objective of the scale was to capture institutional factors that influence entrepreneurial intentions and behaviors. The scale captures four key dimensions: research mobilization, unconventionality, industry collaboration, and university policies. Subsequent studies have used the Entre-U scale in different cultural contexts to measure universities' EO and its relationship with performance. This study adopts the Entre-U scale in China to examine Chinese universities' Entrepreneurial Orientation and its relationship with performance.

3. METHODOLOGY

3.1 Measures

We adopted the Entre-U scale from Todorovic et al. (2011). The Entre-U scale has a total of 23 items measuring four dimensions of EO in universities: Research Mobilization (6 items), Unconventionality (8 items), Industry Collaboration (5 items), and University Policies (4 items). All measures used a 7-point Likert scale ranging from Strongly Disagree to Strongly Agree (1-7). Since the survey was conducted in China, the original scales were translated into Chinese (Mandarin) by a Chinese student who is also fluent in English. Then, the Chinese questionnaire was back-translated into English (Brislin, 1970). The back- translation was performed by a bilingual professor in the United States. The Chinese version of the questionnaire was pre-tested by sending it to individual professors and a group of potential respondents. Comments and feedback were considered in the final version of the questionnaire.
University performance was measured by asking survey respondents about university performance. All activities were assessed by numbers or proportions during the last three years. The survey measured the number of patents and the number of spin-offs/start-ups owned by faculty members from the department.

3.2 Sampling and Data Collection Method

Data were collected via an online survey distributed to department chairs via email. The sample comprised chairs of departments in engineering, natural sciences, biochemistry, computer science, health sciences, aeronautical and astronautical sciences, and agricultural sciences at Chinese universities. A total of 130 universities were selected based on the 2014 Chinese University Alumni Association rankings, including the top 100 universities and the bottom 30 within the top 600. The target population included 4,578 individuals across disciplines: engineering (1,919), natural sciences (715), biochemistry (724), computer science (404), health sciences (662), aeronautical and astronautical sciences (78), and agricultural sciences (76). Of these,

1,013 could not be reached via email or telephone. Two reminder emails were also sent, one week and three weeks after the initial invitation.

A total of 593 responses were received (16.63% response rate). After excluding 253 responses due to excessive missing data (>20%) and five outliers, the final sample consisted of 335 valid cases. The final sample was 83.8% male and 16.2% female; only one respondent was under 30 years old, and 48.1% were 46 or older. Most respondents (81.2%) had served as department chairs for at least three years. By discipline, respondents were drawn from engineering (43.6%), biochemistry (18.8%), natural sciences (15.2%), computer science (9%), and health sciences (10.4%). Non-response bias was assessed by comparing early (first 25%) and late (last 25%) respondents, with results indicating no significant concern.

4. RESULTS

4.1 Confirmatory Factor Analysis

AMOS was used to analyze the data. Confirmatory Factor Analysis was conducted to test the first-order model with four latent factors, and all latent factors were correlated. The preliminary results showed that the factor loadings of UC8 and IC5 were less than 0.50. The two items were dropped from further analysis. After dropping the two items, the Confirmatory Factor Analysis was conducted again. The reliability analysis showed that all measures in the measurement model achieved good reliability. The model fit indices also showed that the measurement model was acceptable (CFI = .841, RMSEA = .106).

Table 1: Reliability Analysis

Research Mobilization		**Unconventionality**		**Industry Collaboration**		**University Policy**	
Items	Standardized Regression Coefficient	Items	Standardized Regression Coefficient	Items	Standardized Regression Coefficient	Items	Standardized Regression Coefficient
RM1	0.565	UC1	0.696	IC1	0.788	UC1	0.725
RM2	0.648	UC2	0.547	IC2	0.669	UC2	0.75
RM3	0.756	UC3	0.717	IC3	0.794	UC3	0.825
RM4	0.784	UC4	0.736	IC4	0.696	UC4	0.577
RM5	0.709	UC5	0.756				
RM6	0.677	UC6	0.71				
		UC7	0.676				
CR=.846;AVE=.481; α=.85		CR=.866;AVE=.482; α=.861		CR=.827;AVE=.546; α=.819		CR=.813;AVE=.525; α=.805	

4.2 Relationship between EO and performance

Using the number of IPs and the number of start-ups as performance measures, SEM was submitted to AMOS for analysis. The model fit indices showed that the model had a good fit to the data (CFI = .844, RMSEA = .096). The correlations between the four dimensions of EO and performance showed that only industry collaboration had a significant impact on performance at the .1 level. The other three dimensions did not have a significant impact on performance.

5. DISCUSSION AND FUTURE RESEARCH DIRECTIONS

The results of this study provide empirical support for the applicability of the Entre-U scale in measuring the Entrepreneurial Orientation of universities in China, thereby extending its validity to a new national and institutional context. However, the findings also indicate that Entrepreneurial Orientation does not have a statistically significant direct impact on traditional performance indicators, such as the number of intellectual property outputs and university spin-offs. This suggests that commonly used performance metrics may not fully capture the entrepreneurial contributions of universities within the Chinese higher education system.

From a theoretical perspective, the study contributes to the EO literature by highlighting the contextual limitations of the EO–performance relationship in higher education institutions. From a practical standpoint, the results have important implications for higher education management and policy design, suggesting that university performance evaluation systems may need to be broadened beyond narrow commercialization outcomes. In particular, more comprehensive measures of entrepreneurial impact are required to better

reflect the diverse ways in which universities contribute to innovation and regional economic development in China.

Future research should therefore focus on developing and validating more holistic performance indicators for entrepreneurial universities. Doing so will improve our understanding of how universities contribute to innovation ecosystems and long-term economic growth.

REFERENCES

[1] Brislin, R. W. (1970). Back-Translation for Cross-Cultural Research. Journal of Cross-Cultural Psychology, 1(3), 185-216. doi:10.1177/135910457000100301

[2] Clark, B. (1989). Creating Entrepreneurial Universities: Organizational Pathways of Transformation. Oxford:IAU Press/Pergamon.

[3] Covin, J. G., & Slevin, D. P. (1989). Strategic Management of Small Firms in Hostile and Benign Environments. Strategic Management Journal, 10(1), 75-87. doi:10.1002/smj.4250100107

[4] Covin, J. G., & Wales, W. J. (2019). Crafting High-Impact Entrepreneurial Orientation Research: Some Suggested Guidelines. Entrepreneurship: Theory & Practice, 43(1), 3-18. doi:10.1177/1042258718773181

[5] Etzkowitz, H., & Leydesdorff, L. (2000). The Dynamics of Innovation: From National Systems and "Mode 2" to a Triple Helix of University-Industry-Government Relations. Research Policy, 29(2), 109-123.

[6] Lumpkin, G. T., & Dess, G. G. (1996). Clarifying the Entrepreneurial Orientation Construct and Linking It to Performance. The Academy of Management Review, 21(1), 135-172. doi:10.2307/258632

[7] Miller, D. (1983). The Correlates of Entrepreneurship in Three Types of Firms. Management Science, 29(7), 770-791. doi:10.1287/mnsc.29.7.770

[8] Rauch, A., Wiklund, J., Lumpkin, G. T., & Frese, M. (2009). Entrepreneurial Orientation and Business Performance: An Assessment of Past Research and Suggestions for the Future. Entrepreneurship: Theory & Practice, 33(3), 761-787. doi:10.1111/j.1540-6520.2009.00308.x

[9] Todorovic, Z., McNaughton, R. B., & Guild, P. (2011). ENTRE-U: An Entrepreneurial Orientation Scale for Universities. Technovation, 31(2/3), 128-137. doi:10.1016/j.technovation.2010.10.009

[10] Wiklund, J., & Shepherd, D. (2005). Entrepreneurial Orientation and Small Business Performance: A Configurational Approach. Journal of Business Venturing, 20(1), 71-91. doi:10.1016/j.jbusvent.2004.01.001

Original Research Article *https://doi.org/10.5281/zenodo.23397304*

BARGAINING SOLUTIONS FOR FAIR CLASSIFICATION: A COMPARISON OF NASH AND KALAI-SMORODINSKY SOLUTIONS

Sandro Radovanović*[1][0000-0002-9975-8844], Milan Vukićević[1][0000-0002-1631-6531],Boris Delibašić[1][0000-0002-6153-5119], Milija Suknović[1][0009-0003-0449-5258]
[1]Univesity of Belgrade, Faculty of Organizational Sciences
*Corresponding author, e-mail: sandro.radovanovic@fon.bg.ac.rs

Abstract: *Fairness-aware machine learning often involves a trade-off between predictive performance and the disparate impact of predictions across groups of people. In practice, efforts to improve fairness can come at the expense of accuracy. In addition, models optimized for accuracy may even reinforce existing disparities. In this work, we examine a bargaining-based approach to fair classification, where classification performance and fairness are treated as competing objectives jointly optimized through a dynamic weighting scheme. We focus on two cooperative bargaining solutions, the Nash Bargaining and the Kalai-Smorodinsky solutions. Our findings indicate that the choice of bargaining rules does more than determine the final trade-off point. It also shapes the trajectory of the optimization process. This suggests that different bargaining formulations may systematically favor different regions of the performance-fairness spectrum. These observations point to several directions for future research. In particular, understanding when Nash bargaining solution tends toward fairness-dominant outcomes and when Kalai-Smorodinsky solution produces more balanced results remains an open question.*

Keywords*: Fairness-aware machine learning, multi-objective optimization, Nash bargaining, Kalai-Smorodinsky bargaining, statistical parity.*

1. INTRODUCTION

Machine learning systems are increasingly deployed in contexts where predictions influence employment, credit, education, social services, and the criminal justice system [1]. In such settings, high predictive performance is often insufficient. In other words, a model may achieve strong overall accuracy while still producing systematically unequal outcomes across demographic groups. This tension has made fairness-aware learning one of concerns in modern machine learning [4].

A key difficulty is that fairness and accuracy rarely improve simultaneously. Optimizing solely for predictive performance can preserve or even amplify group disparities. On the other hand, aggressively enforcing fairness constraints may reduce overall performance. Many existing approaches address this by combining prediction loss and fairness loss into a single objective with a fixed trade-off parameter [4]. Although straightforward, this formulation assumes that the appropriate balance is known in advance. In practice, however, this balance is often unknown because it is context-dependent or sensitive to the data itself [8]. In this work, we explore bargaining-based solutions. Rather than fixing the trade-off between classification performance and fairness in advance, we allow their relative importance to change during training according to a cooperative bargaining solution. We focus on two classical solutions, the Nash Bargaining and the Kalai-Smorodinsky solutions. This perspective evaluates the fairness-accuracy trade-off as a structured negotiation between competing objectives. Our goal is not only to compare the final outcomes in terms of performance and fairness, but also to examine how different bargaining rules shape the optimization trajectory. In particular, we investigate whether the Nash Bargaining and the Kalai-Smorodinsky solutions lead to balanced compromises or tend to favor one objective over the other.

2. METHODOLOGY

Given samples $(\mathrm{x_i}, \mathrm{y_i}, \mathrm{s_i})$ where $x_i \in \mathrm{R}^d$, $y_i \in \{0, 1\}$, and sensitive attribute $s_i \in \{0, 1\}$, the model f_θ outputs a logit $z_i = f_\theta(x_i)$ and probability $p_i = \sigma(z_i)$. We optimize model parameters θ with a weighted multi-objective loss (Formula 1).

$$L_{total}(\theta; w) = w_{cls} L_{cls}(\theta) + w_{fair} L_{fair}(\theta)$$
$$w_{cls}, w_{fair} \geq 0 \quad w_{cls} + w_{fair} = 1 \tag{1}$$

Unlike fixed scalarization, the weights $w = (w_{cls}, w_{fair})$ are updated during training by a bargaining mechanism. In addition to jointly optimizing two inherently conflicting objectives [1], the optimization problem is non-stationary, as the objective function evolves over time.

For the experimental part, we used two well-known datasets in fair machine learning literature. Namely, COMPAS (recidivism prediction) [5] and Adult (income prediction) [7] datasets. For each dataset, we partition the data into training, validation, and test sets to ensure a clear separation between model fitting, model selection, and final evaluation. To preserve the joint distribution of outcomes and sensitive groups, we apply stratification over the label-sensitive attribute pair (y, s). To ensure reproducibility, all experiments are repeated across three independent random seeds. Reported results are aggregated across these runs.

We use a standard multilayer perceptron neural network with two hidden layers and ReLU activations [3]. The model outputs probabilities via a sigmoid activation. The learning objective combines binary cross entropy loss and fairness loss (Formula 2) into a single weighted function. The fairness training objective uses squared statistical-parity surrogate:

$$L_{fair}(\theta) = \left(\frac{1}{|\mathcal{J}_1|} \sum_{i \in \mathcal{J}_1} p_i - \frac{1}{|\mathcal{J}_0|} \sum_{i \in \mathcal{J}_0} p_i \right)^2 \tag{2}$$

where $\mathcal{J}_1 = \{i : s_i = 1\}$ and $\mathcal{J}_0 = \{i : s_i = 0\}$ in the minibatch. It is worth mentioning that bargaining weights are updated every $K = 2$ epochs.

The training follows a bilevel structure. In other words, model parameters and objective weights are updated separately. At the inner level, model parameters are optimized using gradient-based updates on the current loss. At the outer level, the weights for classification and fairness are adjusted based on validation performance (Formula 3). Consequently, this creates a feedback loop where the model learns under a given trade-off, and then the trade-off is updated based on observed value of loss. We note that this procedure does not guarantee a global optimum. It only provides a practical and effective approximation for multi-objective optimization [6].

$$U_{cls}(t) = -L_{cls}^{val}(\theta_t), \quad U_{fair}(t) = -L_{fair}^{val}(\theta_t) \tag{3}$$

For updating objective weights, we consider the Nash Bargaining solution, which selects weights that maximize the joint improvement of both objectives relative to the disagreement point. This corresponds to maximizing the product of gains, which tends to amplify the objective that currently exhibits larger improvements [2]. Let disagreement utilities be d_{cls}, d_{fair}, then:

$$d_k(t+1) = \rho d_k(t) + (1-\rho) U_k(t), \quad k \in \{cls, fair\} \tag{5}$$

Where ρ is the exponential moving average decay factor (set to 0.95 in the experiment). From there one can formulate numerically stable formulation of the Nash bargaining solution we used in training as:

$$J_{NBS} = \log(U_{cls} - d_{cls} + \epsilon) + \log(U_{fair} - d_{fair} + \epsilon) \tag{6}$$

Another objective weights solution, the Kalai-Smorodinsky solution, focuses on equalizing relative progress toward the best observed (ideal) utilities. Instead of maximizing joint gains, it balances the proportion of improvement achieved by each objective. In addition to disagreement points, ideal utilities z_{cls}, z_{fair} are maintained as running best values (Formula 7).

$$z_k(t+1) = \max(z_k(t), U_k(t)), \quad k \in \{cls, fair\} \tag{7}$$

Using both disagreement and ideal points, we define normalized gains that measure progress relative to the attainable range (Formula 8).

$$r_{cls} = \frac{U_{cls} - d_{cls}}{z_{cls} - d_{cls} + \epsilon}, \quad r_{fair} = \frac{U_{fair} - d_{fair}}{z_{fair} - d_{fair} + \epsilon} \tag{8}$$

The Kalai-Smorodinsky bargaining solution seeks to equalize relative improvements across objectives, i.e., $r_{cls} \approx r_{fair}$.

In addition, to approximate the empirical Pareto frontier, we run fixed weight experiments with $w_{fair} \in \{0.0, 0.1, \ldots, 1.0\}$ and $w_{cls} = 1 - w_{fair}$. These are conducted to approximate the empirical Pareto frontier.

For each dataset and random seed, we record three types of outputs. First, we calculate frontier points (test accuracy and statistical parity) so we can construct Pareto plots with uncertainty bars with an idea to visualize the fairness-accuracy trade-off. Second, we store optimization paths so we can track and analyze weight trajectories to track how objective importance evolves. Finally, we calculate aggregated summaries (mean and standard deviation across seeds), so we can assess stability and convergence.

3. RESULTS AND DISCUSSION

As already said, we report average scores over three random seeds on both datasets and compare dynamic Nash and Kalai-Smorodinsky bargaining fair classification solutions with fixed-weight baselines.

Adult. The results on Adult dataset show a rather smooth fairness-accuracy frontier. One can see (see Figure 1) that most of fixed-weight models maintain high accuracy (~0.84) and they can reduce the statistical parity gap from 0.072 to 0.035 as fairness weight increases. On the other hand, Kalai-Smorodinsky mostly stays in high-accuracy region (0.845, gap 0.070), while achieving modest fairness. Nash, in contrast, converges to near-perfect fairness (gap ≈ 0.000) with a notable drop in accuracy (0.759).

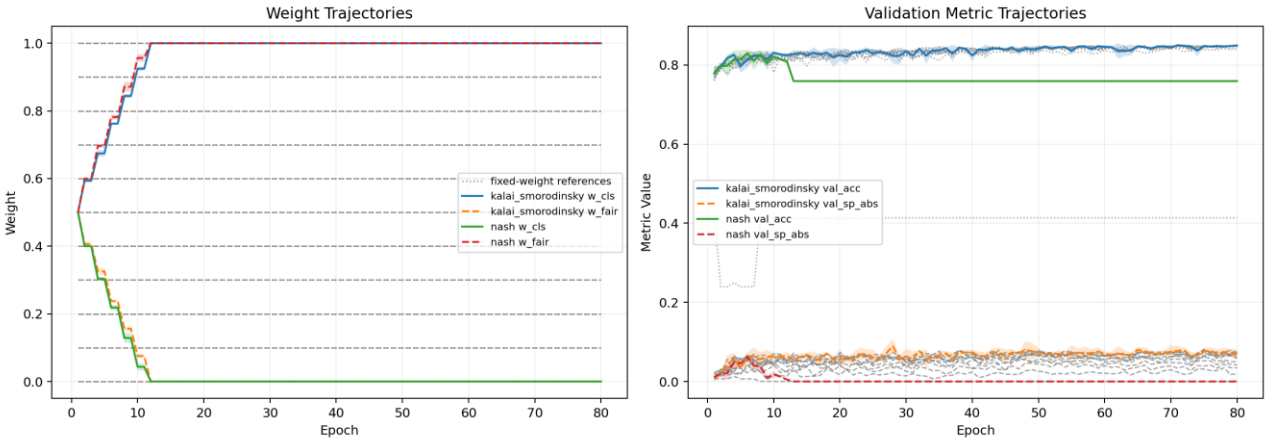


Figure 1: Adult Optimization Paths

COMPAS. On the COMPAS dataset, the frontier is narrower and steeper. This indicates that the trade-off is stronger. If the model focuses on accuracy, then the accuracy is around 0.688 with statistical parity gap near 0.18. Increasing fairness reduces the statistical parity gap but lowers accuracy. Kalai-Smorodinsky stays near the best accuracy (0.684, gap 0.182), while Nash converges to near-perfect fairness (gap ≈ 0.000) with a sharp drop in accuracy (0.493).

Beside accuracy-fairness trade-off, the proposed approach shapes the optimization path as well. Nash bargaining solution amplifies relative gains, thus leading to a feedback effect that pushes toward fairness. On the other hand, Kalai-Smorodinsky enforces proportional progress. In other words, it prevents dominance of a single solution and therefore leads to a more balanced accuracy-fairness trade-off.

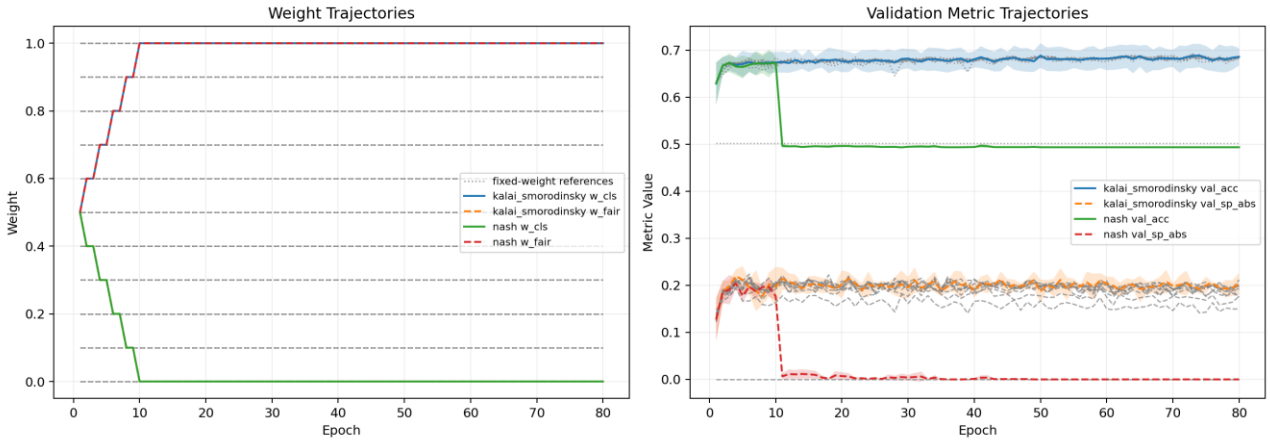


Figure 2. COMPAS Optimization Paths

4. CONCLUSION

This paper examines a bargaining-based formulation of fair classification. The idea is that prediction loss and fairness loss are treated as competing objectives whose weights evolve during training. Experiments are conducted on two well-known datasets, COMPAS and Adult. One can observe that Nash bargaining solution consistently converged to fairness objective, which came at a substantial cost in accuracy. On the other hand, Kalai-Smorodinsky solution remained closer to solutions that satisfy both objectives to some degree. In other words, Kalai-Smorodinsky solution preserved strong predictive performance while providing only moderate fairness improvement.

The results also show that bargaining solutions are not interchangeable. They embody different definitions of compromise and therefore induce different optimization behaviour. Future work will extend this research to other fairness notions, different normalization schemes, and different families of models. However, the main question future work will try to answer is understanding why Nash tends to satisfy fairness objective and why Kalai-Smorodinsky more often results in a jointly acceptable compromise.

ACKNOWLEDGEMENT

This research was supported by the University of Belgrade - Faculty of Organizational Sciences and, in part, by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia through institutional funding (grant number: 200151).

The authors used artificial intelligence tools to improve the language clarity and grammatical accuracy of this work.

REFERENCES

[1] Barocas, S., Hardt, M., & Narayanan, A. (2019). *Fairness and Machine Learning*. fairmlbook.org.
[2] Binmore, K., Rubinstein, A., & Wolinsky, A. (1986). The Nash bargaining solution in economic modelling. *The RAND Journal of Economics*, 176-188.
[3] Bishop, C. M., & Bishop, H. (2023). *Deep Learning: Foundations and Concepts*. Springer Nature.
[4] Caton, S., & Haas, C. (2024). Fairness in machine learning: A survey. *ACM Computing Surveys, 56*(7), Article 166.
[5] Dressel, J., & Farid, H. (2018). The accuracy, fairness, and limits of predicting recidivism. *Science Advances, 4*(1), eaao5580.
[6] Franceschi, L., Frasconi, P., Salzo, S., Grazzi, R., & Pontil, M. (2018, July). Bilevel programming for hyperparameter optimization and meta-learning. In *International Conference on Machine Learning* (pp. 1568-1577). PMLR.
[7] Kohavi, R. (1996, August). Scaling up the accuracy of naive-bayes classifiers: A decision-tree hybrid. In *Knowledge Discovery in Data* (Vol. 96, pp. 202-207).
[8] Radovanović, S., & Ivić, M. (2023). Enabling equal opportunity in logistic regression algorithm. *Management: Journal of Sustainable Business and Management Solutions in Emerging Economies, 28*(2), 55-66.

Original Research Article *https://doi.org/10.5281/zenodo.23397354*

FROM PREDICTION TO DECISION: A POSITION ON INTEGRATING MACHINE LEARNING AND DECISION SUPPORT SYSTEMS

Boris Delibašić*[1][0000-0002-6153-5119], Sandro Radovanović[1][0000-0002-9975-8844],Milija Suknović[1][0009-0003-0449-5258]
[1]University of Belgrade, Faculty of Organizational Sciences
*Corresponding author, e-mail: boris.delibasic@fon.bg.ac.rs

Abstract: *This position paper presents a step toward transitioning from machine learning to decision support systems through budget-aware intervention design. While machine learning models are commonly used to identify individuals at risk, they do not directly address the question of how to allocate limited resources effectively. To bridge this gap, we test an ML-to-DSS pipeline that combines prediction, intervention simulation, and policy evaluation under budget constraints. The approach is evaluated in two domains with multiple intervention policies, including risk-based, segment-based, tiered, cost-efficient, and counterfactual strategies. The results show that predictive performance alone is insufficient for decision quality, and that the effectiveness of intervention strategies depends on domain characteristics and budget formulation. Outcome-oriented metrics such as retention uplift provide a more meaningful evaluation than standard predictive measures. The findings highlight the importance of integrating prediction with decision logic, constraints, and actions.*

Keywords*: Decision Support Systems, Machine Learning, Intervention Policies, Budget Allocation*

1. INTRODUCTION

Machine Learning has become a dominant paradigm in many application domains. A large body of work focuses on estimating the probability of an outcomes such as customer churn [4] or student dropout [5]. While predictions are valuable, they do not address the central question faced by decision makers. *Given limited resources, which interventions should be applied, to whom, and at what cost*?

This gap between machine learning (ML) and decision support systems (DSS) is, arguably, one of the most important in automated decision making. ML models optimize predictive performance, while DSS supports decision-making under constraints [2]. As a result, high predictive accuracy ML models alone don't guarantee decision quality. In other words, even if we can confidently predict the outcome, we might not make the right decision.

In this paper, we study the transition from ML to DSS by testing out different budget-aware intervention strategies. The framework is evaluated in two domains, namely, telecom customer churn and early student dropout. The key contribution of this poster paper is a discussion of ML-to-DSS pipeline that integrates prediction, intervention simulation, and outcome-oriented evaluation.

2. RELATED WORK

We intentionally select customer churn and student dropout as two representative domains for evaluating the intervention approaches. These domains differ substantially in their operational context, intervention mechanisms, and decision constraints. Yet, they share a common structural problem, identifying at-risk individuals and allocating limited resources to prevent adverse outcomes.

2.1. From Prediction to Decision-Making

Some recent work in ML moved beyond viewing prediction as a final goal and instead treat it as an input to a downstream decision problem [3]. In this paradigm, models are trained not only to minimize predictive error

but to improve the quality of the decisions. For example, instead of optimizing for accuracy in predicting churn, a model may be evaluated based on how well it makes final decisions. This approach acknowledges that different decisions may have different consequences and resource requirements, and that optimal choices depend on both predicted outcomes and operational constraints. As a result, the focus moves from *who is at risk* to *what should be done*.

In parallel, DSS research highlights the importance of actionability, interpretability, transparency, and alignment with constraints [2], [9]. These properties are needed as stakeholders must understand, trust, and potentially challenge the recommendations. Unlike ML systems, DSS are explicitly designed to support decision-makers in selecting feasible and justifiable actions. These actions are often linked to interventions. This implies that recommendations must be operationally implementable, or, in other words, comply with organizational rules and resource limitations.

2.2. Budgeted and Policy-Based Interventions

Budget-constrained decision-making has been studied mostly in decision theory and operations management [1]. However, the integration of such optimization frameworks with ML predictions remains limited. ML models are often used to score or rank instances, while the subsequent allocation of interventions is handled through heuristics or fixed business rules. This results in a gap between ML and DSS. Existing approaches often ignore the fact that interventions differ in cost, effectiveness, and feasibility. Consequently, the same ML model can lead to very different outcomes depending on how interventions are selected and prioritized.

This position paper aims to challenge this gap by comparing multiple intervention policies under budget constraints. Rather than assuming a fixed decision rule, we model different strategy logics, from simple risk-based targeting to cost-aware and counterfactual strategies and evaluate their performance in terms of outcome reduction and resource utilization.

3. METHODOLOGY

Two datasets were used to evaluate the ML-to-DSS approaches. The first dataset is a telecom churn dataset containing 3,333 customers with a churn rate of 14.5% [8]. It includes variables related to service usage, billing, calling plans, and customer service interactions, which are standard features in churn prediction tasks. The second dataset is a student dropout dataset comprising 30,586 students, constructed from Open University Learning Analytics [6] using the first 21 days of activity. Each instance represents a unique student-course combination, with features capturing early assessment behavior and engagement in the virtual learning environment.

For the ML modelling, we used four widely used tabular classification models. More specifically, logistic regression, decision trees, random forests, and gradient boosting. Model performance was assessed using ROC-AUC and PR-AUC, as well as Brier score. Across repeated train/test runs, gradient boosting consistently achieved the best overall performance in both domains.

Building on the ML layer, we define five intervention policies that represent different decision-making logics.

- The *risk-ranked outreach policy* serves as a baseline. We contact the people who are most at risk and give them all the same basic help.
- The *segment-based policy* groups (cluster) customers by similarity and assign the best intervention to a group.
- The *tiered retention policy* divides the population into three risk bands and allocates a progressively stronger incentive to higher-risk groups.
- The *ROI-greedy policy* chooses actions that give the biggest impact per unit cost, so the budget is used as efficiently as possible. In other words, we design an optimization problem to maximize the retention of the group given budget constraints.
- Finally, the *counterfactual policy* takes an individualized approach by identifying the minimal set of feasible actions required to shift a predicted adverse outcome toward a favorable one [7].

To evaluate intervention we, for each instance in the dataset, use the ML model to produce a baseline probability of the outcome. Then, we simulate candidate interventions by modifying relevant feature values and get their assumed effects. After that, we re-evaluate instance using the same ML model to obtain a post-intervention probability. The value of an intervention is then defined as the reduction in predicted risk.

Intervention policies generate recommended actions, which are then selected under explicit budget constraints. Evaluation metrics include total expected outcome reduction, total cost, number of treated cases, and retention uplift. Each action has an assumed cost and feature-level effect. Also, the trained predictive

model is used to estimate expected outcome reduction after the intervention is applied. Consequently, the DSS layer is simulation-based rather than causal. Every experiment was repeated five times with different budget constraints Results are aggregated across runs with 95% confidence intervals.

4. RESULTS AND DISCUSSION

In churn, gradient boosting achieved mean ROC-AUC 0.912, mean PR-AUC 0.858, and mean Brier score 0.042. In dropout, it achieved mean ROC-AUC 0.811, mean PR-AUC 0.829, and mean Brier score 0.176. This means that both problems are somewhat predictable.

The results, summarized in Figure 1 (churn) and Figure 2 (dropout), show how intervention policies translate ML outputs into decision outcomes under varying budget constraints.

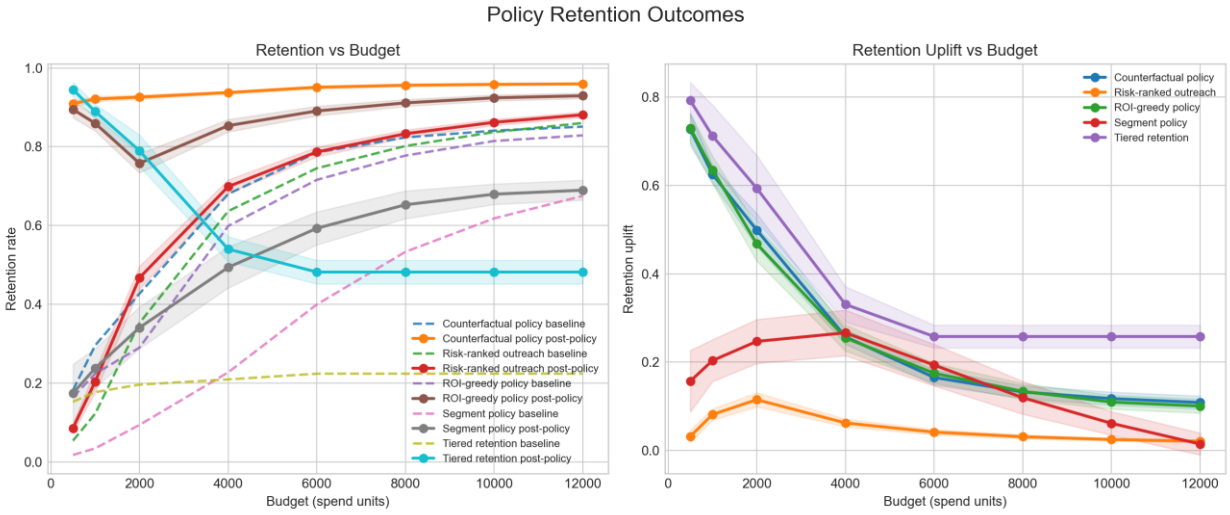


Figure 1: Churn dataset policy retention outcomes

In the churn setting (Figure 1), the counterfactual policy performs strongly across budgets. However, it does not uniformly dominate. It provides high retention at low and moderate budgets, though its gains diminish as budget increases, suggesting diminishing returns. The ROI-greedy and segment-based policies achieve comparable performance at higher budgets and in some cases slightly outperforming the counterfactual approach. This can be interpreted that cost-efficient allocation and structured targeting can become more competitive as more resources are available. The tiered policy performs well at low budget. Meanwhile, the risk-ranked outreach baseline remains the weakest.

In the dropout setting (Figure 2), the tiered retention policy achieves the highest uplift across most budget levels. The counterfactual and ROI-greedy policies perform competitively but remain slightly below the tiered approach. The segment-based policy provides moderate and relatively stable improvements. As expected, the risk-ranked outreach baseline remains the weakest.

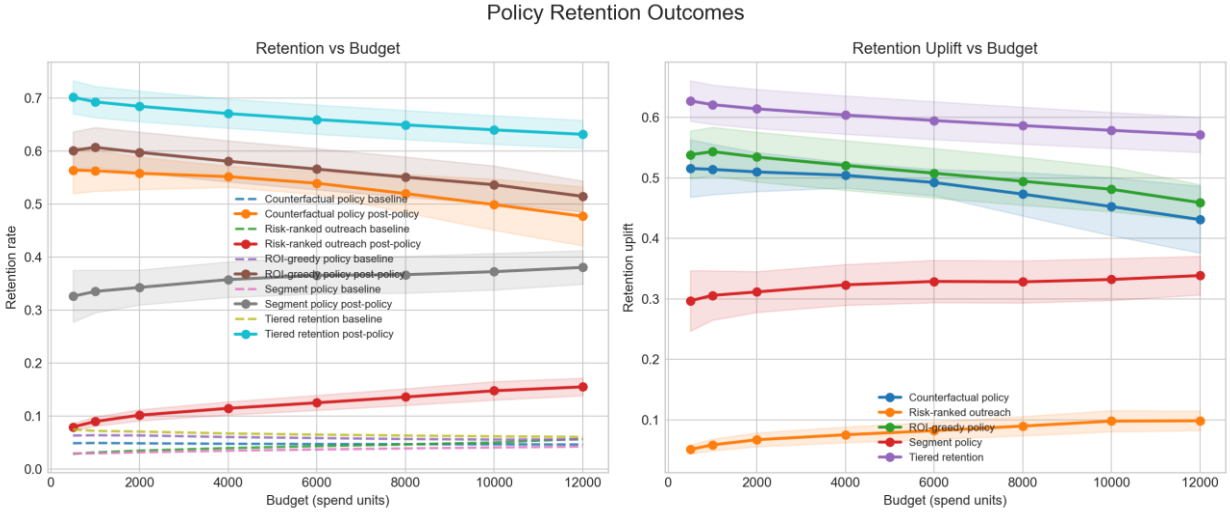


Figure 2: Student dropout dataset policy retention outcomes

These results highlight what can be perceived as a shift from ML toward DSS system design. While accurate predictions are an important prerequisite, it represents only the first stage of the decision-making process. Across both domains, we observe that ML models with strong predictive performance can lead to substantially different outcomes depending on how their outputs are translated into actions. In other words, good ML models alone are not a sufficient proxy for decision quality.

A central finding of this poster paper is that policy choice matters. Even when the same ML model is used, different intervention strategies results in different outcomes. In the churn domain, individualized counterfactual policies leverage strong feature–action relationships to achieve high impact, while in the dropout domain, cost-efficient or structured policies such as ROI-greedy and tiered retention perform better. Consequently, we can state that there is no universally best strategy for any given application. The best-performing policy depends on the domain and the budget regime. In other words, DSS should not rely on a single fixed policy, but rather test and evaluate different policy designs.

Another key insight of the paper concerns the role of budget constraints. The results indicate that the way constraints are modelled significantly influences the policy performance. It is not the same if budget is modelled as total amount spend or as treatment capacity (number of people). Finally, expressing outcomes in terms of retention and uplift might provide a more meaningful and interpretable basis for comparison than traditional ML metrics.

4. CONCLUSION

This paper shows a transition from ML to DSS. We present three conclusions. First, ML predictive performance is necessary, but insufficient for achieving good decision quality. Second, incorporating budget-aware policy simulation alters which strategies are performing better. Third, outcome-oriented metrics such as retention uplift might provide a more actionable link between ML and DSS. A key limitation is that the intervention layer is simulation-based rather than causally identified. However, the approach remains valuable as it demonstrates how ML models can be integrated with policy selection, budget allocation, and intervention evaluation.

ACKNOWLEDGEMENTS

This research was supported by the University of Belgrade - Faculty of Organizational Sciences and, in part, by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia through institutional funding (grant number: 200151).

The authors used artificial intelligence tools to improve the language clarity and grammatical accuracy of this work.

REFERENCES

[1] Bertsimas, D., & Kallus, N. (2016). The power and limits of predictive approaches to observational-data-driven optimization. *arXiv preprint arXiv:1605.02347*.

[2] Bohanec, M. (2021). From data and models to decision support systems: lessons and advice for the future. In *EURO Working Group on DSS: A tour of the DSS developments over the last 30 years* (pp. 191-211). Cham: Springer International Publishing.

[3] Elmachtoub, A. N., & Grigas, P. (2022). Smart "predict, then optimize". *Management Science, 68*(1), 9-26.

[4] Khattak, A., Mehak, Z., Ahmad, H., Asghar, M. U., Asghar, M. Z., & Khan, A. (2023). Customer churn prediction using composite deep learning technique. *Scientific Reports, 13*(1), 17294.

[5] Krueger, J. G. C., de Souza Britto Jr, A., & Barddal, J. P. (2023). An explainable machine learning approach for student dropout prediction. *Expert Systems with Applications, 233*, 120933.

[6] Kuzilek, J., Hlosta, M., & Zdrahal, Z. (2017). Open university learning analytics dataset. *Scientific Data, 4*(1), 170171.

[7] Straitouri, E., & Rodriguez, M. G. (2023). Designing decision support systems using counterfactual prediction sets. *arXiv preprint arXiv:2306.03928*.

[8] Umayaparvathi, V., & Iyakutti, K. (2016). A survey on customer churn prediction in telecom industry: Datasets, methods and metrics. *International Research Journal of Engineering and Technology (IRJET), 3*(04), 1065-1070.

[9] Ustun, B., Spangher, A., & Liu, Y. (2019, January). Actionable recourse in linear classification. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (pp. 10-19).

Original Research Article *https://doi.org/10.5281/zenodo.23397375*

REVENUE OPTIMIZATION OF NEW YORK CITY TAXI SERVICES USING LINEAR PROGRAMMING AND METAHEURISTICS

Isidora Gaćeša*[1][0009-0002-0907-7170], Gordana Savić[1][0000-0001-6233-1815], Violeta Cvetkoska[2][0000-0002-0472-5522]
[1]University of Belgrade, Faculty of Organizational Sciences
[2]Ss. Cyril and Methodius University in Skopje, Faculty of Economics-Skopje
*isidora.gacesa@fon.bg.ac.rs

Abstract: *Modern society is characterized by rapid urbanization and technological development. Urban mobility has become a significant part of everyday functioning in the urban environment, especially due to the growing demand for fast and easily accessible transportation. Taxi services represent a crucial segment for meeting the transportation needs of people living in urban core areas, but their versatility generates challenges in efficient vehicle allocation and utilization. In this research, the problem of adaptive taxi vehicle allocation by months using real-life data for New York Yellow Taxi services from January through November 2025 is considered. A revenue-oriented optimization model is proposed and formulated as a mixed integer linear programming (MILP) problem, and additionally solved using metaheuristic methods - Genetic Algorithm (GA) and Simulated Annealing (SA). The results indicate two possible directions in choosing methodology, depending on whether the priority is maximization of the objective or stability of the solution.*

Keywords*: Taxi fleet allocation, Mixed Integer Linear Programming (MILP), Metaheuristics, Revenue maximization, Genetic Algorithm*

1. INTRODUCTION

Modern society is characterized by significantly rapid urbanization and technological development. Urban mobility has become a significant part of everyday functioning in the urban environment, especially when it comes to growing demand for fast and easily accessible mobility, followed by a vast number of challenges such as traffic jams and pollution (Miskolczi et al., 2021). Therefore, in the domain of urban mobility structures, taxi services represent a crucial segment for meeting the transportation needs of people living in urban core areas (Zhang et al., 2024). Even though taxi services are capable of adapting to demand fluctuations caused by different locations, time of day, and user activities, this versatility generates consequential challenges in the domain of efficient vehicle allocation and utilization. Inefficient allocation of taxi vehicles can result in higher operational costs, longer service waiting times, and a decrease in overall system efficiency (Alam et al., 2025; Adegoke et al., 2025). Within this framework, the issue of taxi fleet distribution optimization becomes rather significant, particularly due to the dynamic nature of demand fluctuations during various time periods. The core research problem addressed in this paper is how to optimally distribute a fixed fleet of taxi vehicles across urban zones over time, in order to maximize total revenue while accounting for monthly variations in spatial demand. The primary research motivation stems from the gap in existing literature: most studies address individual aspects of taxi optimization in static settings, without systematically comparing multiple optimization approaches on real, temporally varying data. This research aims to bridge that gap by proposing and comparing exact and metaheuristic methods applied to eleven consecutive months of real New York City taxi data.

In this research, the question of adaptive taxi vehicle allocation by months using real-life data for New York Yellow Taxi services from January through November 2025 is considered. In order to solve the problem stated above, both exact and approximate methods are used, namely, a mixed integer linear programming (MILP) model, a Genetic Algorithm (GA), and Simulated Annealing (SA). The proposed model is a revenue-oriented formulation aimed solely at maximization of the total fleet revenue.

2. RELATED WORK

There is evidence on urban taxi systems and on-demand mobility, which can be divided into several groups, including demand prediction, fleet size and allocation, route optimization, and the application of metaheuristic approaches. Demand prediction is considered by Alam et al. (2025), whose study explores several machine learning algorithms in the context of New York City, demonstrating their importance for vehicle allocation. In addition, Adegoke et al. (2025) find a TCN-based approach to achieve better precision and stability compared to classical approaches using higher-resolution data. Although these studies propose effective forecasting techniques, they do not incorporate demand estimation into actual fleet allocation decisions, which represents a gap addressed by this research.

When studying fleet size and allocation, researchers have developed several methods to calculate the number of taxis in the system. For example, in their most recent research, Malidis et al. (2026) propose a newly developed algorithm specifically designed for vehicle allocation, which takes into consideration both waiting time and operational costs. In the same domain of fleet allocation, but focusing on different aspects of transportation, Mondal et al. (2023) strive to achieve the minimum possible idle time and CO_2 emissions by employing different metaheuristics. Consistent with previous authors, Kronmueller et al. (2024) solve a related problem by developing and implementing a MILP model. The results of their research suggest the possibility of reducing the number of vehicles if a compromise between additional delay and execution time can be achieved. Although the above-mentioned papers employ methods that are efficient when it comes to fleet size estimation and vehicle allocation, they are strictly time-fixed or single-period oriented while also neglecting allocation changes between different time periods due to changes in demand, which opens up space for the research conducted in this paper.

When it comes to the usage of metaheuristic methodologies, a few existing studies have analyzed different approaches to their application in taxi systems. Among the first researchers, Nunes et al. (2011) utilize a GA-based algorithm designed to solve the Traveling Salesman Problem in the domain of taxi service operations. On the other hand, Yao et al. (2018) take a step further by proposing a hyperheuristic approach intended to enable the efficient solution of large-scale transportation problems. Alongside these two papers, Lee (2026) compares a few different metaheuristic methods by applying them to New York City commute data. His research concluded that the SA algorithm is significant when it comes to generating the initial solution. Lastly, Li et al. (2017) employ MDP to develop an algorithm for sequential empty vehicle routing. Although all of the above works showcase the performance of different metaheuristics in taxi service optimization, they mostly focus on employing a single technique or comparing their results with one another. None of them compare these methods with exact methods, such as MILP, in terms of goal achievement and solution stability. Furthermore, none of the above papers analyze changes in vehicle allocation due to demand variations across different time periods. This represents another difference between the existing literature and this study, which examines the problem by taking all of these factors into account.

3. METHODOLOGY

3.1. Mixed Integer Linear Programming (MILP)

Both in theory and in practice, a large number of different optimization problems can be transformed into their representative linear program designed to enable finding of an optimal solution (Cohen et al., 2021). Linear programming (LP) represents one of the most widely known techniques used for solving real-world problems. Standard LP is based on the assumption that all decision variables can take continuous values, which significantly narrows the number of practical applications that can be justified. In some cases, such as vehicle allocation or assignment tasks, decision variables are subject to integer constraints, which calls for a specific branch of linear programming known as integer linear programming (ILP). Mixed integer linear programming (MILP) further allows for the simultaneous existence of both integer and continuous decision variables within the same optimization problem, contributing to greater modeling flexibility while preserving the linear structure of the model (Taha, 2020). Given that vehicle allocation inherently involves integer decision variables, while the later used auxiliary variables are continuous, MILP represents a natural and well-justified modeling choice for the problem at hand.

3.2. Genetic Algorithm (GA) and Simulated Anneling (SA)

Genetic algorithm (GA) represents a metaheuristic optimization method which is based on principles of natural selection and heredity. GA logic is based on working with a population of potential solutions, where each solution represents a separate individual encoded in the form of chromosomes. When evaluating quality and generating the next generation, primary biologically inspired operators of GA are applied: selection, crossover and mutation. Due to its flexibility and ability to provide high-quality approximate solutions for large-scale optimization problems within a reasonable time frame, the GA approach is particularly suitable when exact methodologies become time-consuming and computationally expensive (Sadeeq & Abdulazeez, 2023).

Similarly to GA, Simulated annealing (SA) represents a metaheuristic optimization method whose foundations are inspired by metal annealing in physics. A key advantage of the SA approach lies in its ability to avoid local optima and explore a larger solution space in order to find a better solution. Beginning from an initial solution, in each iteration the algorithm generates a new one in the vicinity of the previous one. If the generated solution is of better quality, it is accepted unconditionally; otherwise, the worse solution may be accepted with a certain probability, which depends on the difference in the objective function value and a control parameter known as temperature (Boussaid et al., 2013). GA and SA were selected as two structurally different metaheuristic approaches – population-based and trajectory-based, respectively, enabling a meaningful comparison of approximate methods against the exact MILP benchmark.

4. MODEL FORMULATION

The central problem considered in this paper is the optimal allocation of a limited number of taxi vehicles per set of urban zones in order to maximize the total revenue of the fleet, taking into account spatial and temporal variations of transport demand.

Let $Z = \{1,2,3,\dots,n\}$ denote the set of urban zones in which taxi services are provided. For each zone $i \in Z$, the following parameters are defined: D_i – demand in zone i, expressed as the cumulative number of trips in the observed time period; R_i – average revenue per trip in zone i, in US dollars; N – total number of taxi vehicles available for allocation (fleet size). The decision variables (x_i) represent the number of taxi vehicles allocated in zone i. Since the generated revenue is limited by real demand, if x_i is greater than D_i the excess vehicles in zone i do not generate additional revenue. The contribution of each zone to the total revenue can therefore be expressed as $min(x_i, D_i) \times R_i$. To linearize this nonlinear contribution, a set of auxiliary variables (y_i), one for each zone $i \in Z$, was introduced. The model is defined as follows:

$$(max) \sum_{i=1}^{n} R_i y_i \tag{1}$$

$$u.c.$$

$$\sum_{i=1}^{n} x_i = N \tag{2}$$

$$y_i \leq x_i, \forall i = 1,2,\dots,n \tag{3}$$

$$y_i \leq D_i, \forall i = 1,2,\dots,n \tag{4}$$

$$x_i \in Z^+, y_i \geq 0, \forall i = 1,2,\dots,n \tag{5}$$

Each y_i represents the effective utilization of zone i, that is, $y_i = min(x_i, D_i) \times R_i$. Since y_i appears in the objective function with a positive coefficient R_i, the search for an optimal solution naturally pushes yi toward its maximum.

For the GA and SA heuristics, the fleet capacity constraint (2) cannot be imposed directly and is implemented through the fitness function via a quadratic penalty $P(x) = (x_i - N)^2 \times 10^6$, which strongly discourages inadmissible solutions. The heuristics were implemented in R: package GA with population size 100, maximum 500 generations and stop after 50 generations without improvement; package GenSA for SA with a maximum of 5000 iterations and starting point $x_i = \frac{N}{n}$.

5. DATA PREPARATION

The data used for the study were obtained from the New York City TLC trip record database (NYC TLC, 2026). Yellow Taxi trip records from January through November 2025 were used. The data preparation process was carried out using the R programming language and included filtering, removing extreme values, aggregation by zones, selection of the most significant zones and demand scaling. For each month of the analysis, the 10 zones with the largest number of realized rides were considered. The fleet size N was calculated as the median demand across all 10 selected zones, ensuring that the fleet is large enough to cover typical demand, but not large enough to provide maximum capacity to each zone.

6. RESULTS AND DISCUSSION

The solution provided by the MILP approach, which focuses only on revenue maximization, expectedly dominates the metaheuristic approaches. The MILP revenue values range from 7.85 million dollars in August to 10.48 million dollars in May, while the average monthly revenue amounts to 9.55 million dollars. This result is expected, given that the model exclusively focuses on revenue maximization while completely ignoring the disproportion in vehicle allocation. Implemented metaheuristic methods (GA and SA) group their values in a narrow range from 2.99 to 4.80 million dollars, with the average revenue of 3.80 million dollars for GA and 3.73 million dollars for SA. Relatively small differences in the obtained revenue values among metaheuristics

suggest that implementing the squared penalty for fleet size consequently leads to mostly similar local optima. Table 1 summarizes the revenue performance of all methods.

Table 1: Summary of revenue performance by method (averaged over 11 months)

Method	Average Revenue	Min	Max	Std Range
MILP	$9.55M	$7.85M	$10.48M	$2.63M
GA	$3.80M	$2.99M	$4.80M	$1.80M
SA	$3.73M	$3.02M	$4.52M	$1.49M

According to the obtained results, the primary structure of the proposed model leads to a pronounced concentration of taxi vehicles in a small number of highly profitable zones. The subject distribution, which was anticipated before and during the model testing, leads to a complete disregard of the zones characterized with a lower individual revenue value, resulting in a significant disproportion in the coverage of the demand within the considered zones. Consequently, MILP also showcases the most noticeable oscillations in monthly allocation, which is a direct consequence of concentrating vehicles in a small number of highly profitable zones. In contrast, the SA method achieves the most stable allocations among all used methods, while the GA remains relatively stable with moderate variability across all months (Figure 1).

Variation of revenue value over months contributes to a better understanding of the stability of the obtained solutions. The metaheuristic methods showcase subtle changes in revenue values which are consistent over the observed time period. This highlights that, although the revenue values obtained from GA and SA are lower than those obtained from MILP, they are more consistent and stable over time, which validates the applicability of GA and SA when stability of the solution is prioritized over absolute objective value.

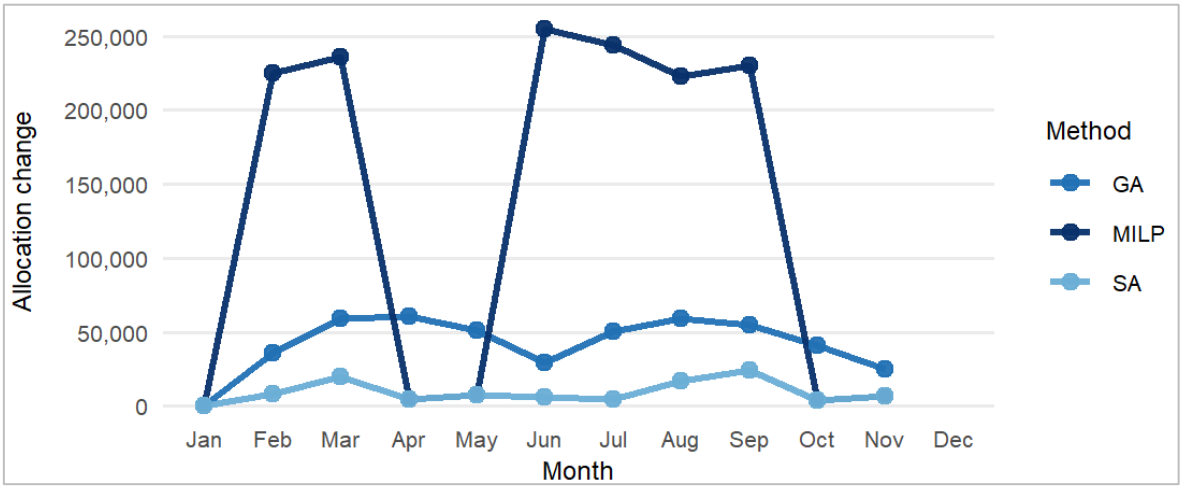


Figure 1: Monthly allocation changes

7. CONCLUSION

Based on the obtained results, it can be concluded that the exact, MILP-based, approach to problem solving generates significantly better results when it comes to revenue maximization. Compared to the employed GA-based and SA-based metaheuristic approaches, the MILP solution results in a revenue value that is 2.5 times higher than the values achieved by the metaheuristic methods. Additionally, taking into consideration all 11 months analyzed, the mentioned difference becomes even more pronounced when monthly revenue is analyzed. For MILP, average monthly revenue stands at $9.55M and exceeds those of GA and SA, which amount to $3.80M and $3.73M, respectively. Therefore, if the focus is solely on revenue maximization, the MILP approach wins a convincing victory over the GA and SA approaches. On the other hand, the situation changes considerably when it comes to allocation stability, which favours the metaheuristic algorithms. Upon reviewing the number of vehicles allocated to different city zones across different months, it was noticed that the metaheuristic approaches, especially the SA-based one, require significantly less vehicle reallocation each month compared to MILP. Additionally, when analyzing the deviation of the monthly change, the portrayed difference in allocation stability is even more substantial. The deviation for MILP revenue amounts to $2.63M, while the deviation for GA and SA is almost half as large, with values of $1.80M for GA and $1.49M for the SA approach. Therefore, whether the priority is to maximize revenue or obtain operational stability, either MILP or metaheuristics can be employed. Future studies could focus on the introduction of balancing mechanisms that

penalize uneven vehicle allocation, as well as the incorporation of additional real-world constraints into the existing model.

REFERENCES

[1] Adegoke, O., Adepitan, A. A., & Kwakye, S. (2025). High-resolution mobility data science for urban transportation policy, demand forecasting, and infrastructure investment prioritization. International Journal of Science and Engineering Applications, 14(7), 82-97.

[2] Alam, S., Ayub, M. S., Cui, H., & Khan, M. A. (2025). A comparative study of machine learning models for taxi-demand prediction using a big data framework. Public Transport, 17(3), 803-833.

[3] Boussaid, I., Lepagnot, J., & Siarry, P. (2013). A survey on optimization metaheuristics. Information Sciences, 237, 82-117.

[4] Cohen, M. B., Lee, Y. T., & Song, Z. (2021). Solving linear programs in the current matrix multiplication time. Journal of the ACM, 68(1), 1-39.

[5] Kronmueller, M., Fielbaum, A., & Alonso-Mora, J. (2024). Reducing the minimal fleet size by delaying individual tasks. IEEE Transactions on Intelligent Transportation Systems, 25(9), 11011-11022.

[6] Lee, S. (2026). Heuristic bus route optimization incorporating taxi-derived route preferences under NYC congestion pricing. American Journal of Student Research.

[7] Li, J. P. K., Bhulai, S., & van Essen, T. (2017). Optimization of the revenue of the New York City taxi service using Markov decision processes. In 6th International Conference on Data Analytics (pp. 1-13). Barcelona, Spain.

[8] Mallidis, I., Aguilera-Garcia, A., Salanova Grau, J. M., Siomos, A., & Vassallo, J. M. (2026). Optimizing fleet size for on-demand taxi and ridehailing services: Application to the case of Madrid. Transport Policy, 103868.

[9] Miskolczi, M., Foldes, D., Munkacsy, A., & Jaszberenyi, M. (2021). Urban mobility scenarios until the 2030s. Sustainable Cities and Society, 72, 103029.

[10] Mondal, M., Sano, K., Kato, T., & Puppateravanit, C. (2023). Optimization of taxi allocation for minimizing CO2 emissions based on heuristic algorithms. Smart Cities, 6, 1589-1611.

[11] Nunes, J., Matos, L., & Trigo, A. (2011). Taxi pick-ups route optimization using genetic algorithms. In International Conference on Adaptive and Natural Computing Algorithms (pp. 1-13). Springer.

[12] NYC TLC. (2026). NYC TLC Trip Record Data. Retrieved from: https://www.nyc.gov/site/tlc/about/tlc-trip-record-data.page.

[13] Sadeeq, H. T., & Abdulazeez, A. M. (2023). Metaheuristics: A review of algorithms. International Journal of Online and Biomedical Engineering, 19(9), 142-164.

[14] Taha, H. A. (2020). Integer programming. In Mathematical Programming for Operations Researchers and Computer Scientists (pp. 41-69). CRC Press.

[15] Yao, Y., Peng, Z., & Xiao, B. (2018). Parallel hyper-heuristic algorithm for multi-objective route planning in a smart city. IEEE Transactions on Vehicular Technology, 67(11), 10307-10318.

[16] Zhang, M., Tian, D., Liu, J., & Li, X. (2024). Analysis of taxi demand and traffic influencing factors in urban core area based on data field theory and GWR model: A case study of Beijing. Sustainability, 16(17), 7386.

Original Research Article *https://doi.org/10.5281/zenodo.23397420*

INTEGRATING NEURAL NETWORKS IN CARDIAC BIOPHYSICAL MODELING

Bogdan Milićević[1,2*[0000-0002-0315-8263]], Miloš Ivanović[2,3[0000-0002-8974-2267]], Boban Stojanović[2,3[0000-0003-2901-0061]], Miljan Milošević[1,2,4[0000-0003-3789-2404]], Miloš Kojić[2,5[0000-0003-2199-5847]], Nenad Filipović[2,6[0000-0001-9964-5615]]

[1]Institute for Information Technologies, University of Kragujevac, Kragujevac 34000, Serbia
[2]Bioengineering Research and Development Center (BioIRC), Kragujevac 34000, Serbia
[3]Faculty of Science, University of Kragujevac, Kragujevac 34000, Serbia
[4]Belgrade Metropolitan University, Belgrade 11000, Serbia
[5]Serbian Academy of Sciences and Arts, Belgrade 11000, Serbia
[6]Faculty of Engineering, University of Kragujevac, Kragujevac 34000, Serbia
*Corresponding author, e-mail: bogdan.milicevic@uni.kg.ac.rs

Abstract: *This paper presents an overview of recent advancements in the application of Artificial Neural Networks for modeling cardiac muscle mechanics and electrophysiology, specifically focusing on the transition from traditional biophysical solvers to surrogate models. While biophysical models offer high physiological fidelity, their integration into multi-scale clinical simulations is often restricted by computational costs. We examine two strategies to address this: data-driven surrogates and Physics-Informed Neural Networks (PINNs). In the data-driven domain, we explore recurrent and convolutional architectures, and we evaluate their ability to capture history-dependent muscle behavior, and to achieve speed-ups in finite element frameworks. PINNs embed governing partial differential equations directly into the network's loss function. By synthesizing these approaches, this work demonstrates how AI-driven frameworks can augment traditional numerical solvers, providing a scalable pathway to real-time biomechanical simulations.*

Keywords: *convolutional neural networks, recurrent neural networks, physics-informed neural networks, cardiac modeling, finite element modeling*

1. INTRODUCTION

High-fidelity biophysical models, such as the Huxley-type models for cross-bridge dynamics and systems for electrophysiology, provide physiological detail. These models allow for the simulation of patient-specific conditions, such as cardiomyopathies and arrhythmias, by linking molecular-level changes to organ-level performance [1, 2]. However, the integration of these models into multi-scale finite element (FE) frameworks presents a computational challenge. A single simulation of a cardiac cycle can require thousands of time steps, with micro-level biophysical equations being solved at each integration point, leading to execution times that are prohibitive for real-time clinical applications [3, 4]. To bridge this gap, Artificial Neural Networks (ANNs) have emerged as powerful tools for creating high-performance surrogate models. These neural frameworks are generally divided into two paradigms: data-driven surrogates and Physics-Informed Neural Networks (PINNs). For history-dependent materials Gated Recurrent Units (GRU) and Temporal Convolutional Networks (TCN), are used to learn the mapping between mechanical inputs (e.g., stretch and activation) and outputs (e.g., stress and stiffness). These models are particularly adept at capturing the history-dependent nature of muscle tissue, where current tension is a function of the muscle's previous states [4, 6]. Parallel to data-driven methods, Physics-Informed Neural Networks (PINNs) have redefined how biophysical equations are solved. Unlike traditional black-box models, PINNs embed the governing partial differential equations (PDEs) directly into the neural network's loss function. This allows the network to act as a meshless solver that respects the underlying physics of the system [7]. While initially popularized for fluid dynamics, PINNs are now being increasingly applied to both cardiac electrophysiology—to capture the rapid depolarization "spikes"—and muscle mechanics, where they solve the Huxley rate equations without the need for traditional numerical integration [7]. A significant hurdle in deploying these neural models is the selection of optimal architectures. Recent research has focused on automating this process through Bayesian Optimization (BO), which systematically searches for the most efficient network depth and width to balance accuracy and speed [8]. By combining these AI-driven frameworks with conventional FE solvers, it is

possible to achieve significant speed-ups while maintaining high correlation with gold-standard biophysical simulations [4]. Our research aligns with broader advancements in scientific machine learning aimed at blending neural function approximators with physical principles. Beyond classical PINNs, the field has rapidly expanded into neural operator learning, such as Deep Operator Networks, which learn operators mapping between infinite-dimensional function spaces rather than finite vectors [9]. From an analytics perspective, these surrogates function as predictive tools that trade a marginal amount of accuracy for orders-of-magnitude gains in computational throughput. This work constitutes a comparative synthesis of published results from the authors' research programme on ANN surrogates for cardiac simulation, identifying methodological trade-offs and outlining directions for future integration within clinical-scale finite element workflows.

2. METHODS

The development of high-performance computational models for complex biological systems relies on two distinct methodological frameworks: the construction of data-driven surrogates and the implementation of physics-informed neural networks. To capture the behavior of systems where current states are dependent on previous time-series data, deep learning architectures with temporal memory are employed. Gated Recurrent Units (GRU) are utilized due to their specialized gating mechanisms, comprising reset and update gates [4]. These gates allow the model to selectively retain or discard information from previous time steps, which is critical for modeling the "memory" inherent in physiological processes where output is a function of the entire deformation history [6]. In parallel, Temporal Convolutional Networks (TCN) offer an alternative approach by capturing long-range patterns through causal, dilated convolutions. Unlike standard convolutional layers, TCNs ensure that predictions for a specific time point depend only on past data, preventing information leakage from the future. By using dilated layers, the network maintains a high receptive field with fewer layers than traditional recurrent models, enabling the efficient processing of long sequences while maintaining lower computational overhead during the inference phase of multi-scale simulations [4].Physics-Informed Neural Networks (PINNs) incorporate the residuals of the governing partial differential equations directly into the network's loss function, and the model is forced to find solutions that are physically consistent [7]. To resolve sharp, non-linear transitions within the time domain, a sequential causal training strategy can be implemented. In case of muscle modeling presented in this paper, inputs to GRU and TCN are mechanical stretch and activation history sequences, while outputs are stress and stiffness. Training data are generated by running the Huxley model over a library of loading trajectories, with loss being mean squared error on output sequences. In TCN causal dilated convolutions replace recurrent layers to improve inference speed. For PINNs, inputs are spatial coordinates of myosin heads relative to binding sites, muscle activation, stretch and time. The loss function combines the PDE residual at collocation points with initial condition residuals.

3. RESULTS AND DISCUSSION

The performance of the developed neural frameworks was evaluated by comparing their predictions against gold-standard numerical solutions across several complex mechanical and electrophysiological scenarios. Mechanical response of the muscles was tested with the GRU and TCN architectures, in place of traditional Huxley muscle model within the finite element analysis framework. As illustrated in Figure 1, the networks were subjected to diverse loading conditions including isotonic contraction, quick release, prescribed force, and prescribed displacement. In all cases, the surrogates showed exceptional agreement with the original biophysical model [4]. The quick release (Figure 1b) test was particularly significant, as it required the networks to process rapid changes. Quantitatively, the surrogate models achieved Pearson correlation coefficients exceeding 0.99, while providing a computational speed-up of approximately 25 to 50 times compared to traditional sequential solvers [4, 6]. This means that a cardiac cycle simulation requiring hours with the reference solver can be completed in minutes using the surrogate.Beyond isolated mechanical tests, the integration of the GRU surrogate into a three-dimensional finite element model of the left ventricle demonstrated the scalability of the approach. During a full cardiac cycle simulation, the surrogate successfully predicted the complex distribution of active stress across the ventricular wall. The multi-scale framework handled the transition from the microscopic cross-bridge dynamics to the macroscopic organ deformation with high numerical stability. This integration is vital for simulating the left ventricle's ejection phase, where the surrogate model maintained the physiological pressure-volume relationship characteristic of healthy cardiac function. The implementation of Physics-Informed Neural Networks (PINNs) provides a robust alternative to data-driven surrogates by embedding the governing biophysical laws directly into the learning process. Unlike traditional black-box models, the PINN acts as a continuous solver by minimizing the residual of the Huxley rate equations alongside initial conditions. For isometric contraction, the framework utilized a deep neural network with 8 hidden layers (20 neurons each). To ensure numerical stability and a unique solution, the Neural Tangent Kernel (NTK) method was employed to adaptively weight the loss function components, balancing the residuals from the collocation points against those from the initial conditions. As shown in the comparison with the method of characteristics, the PINN successfully

emulated the stress generation with high fidelity over the spatial domain (ranging from -20.8 nm to 62.4 nm). The isotonic contraction scenario presented a greater challenge due to varying muscle activation and non-zero contraction velocities. For this case, a larger architecture of 8 hidden layers with 200 neurons was required. The network was trained on a dataset enriched with randomly generated activation functions and integrated as a material model at the macro level within the finite element software. Despite the increased complexity, the PINN demonstrated a significant correlation with traditional solvers, accurately capturing the stress transitions during the contraction and recovery phases (Figure 2).

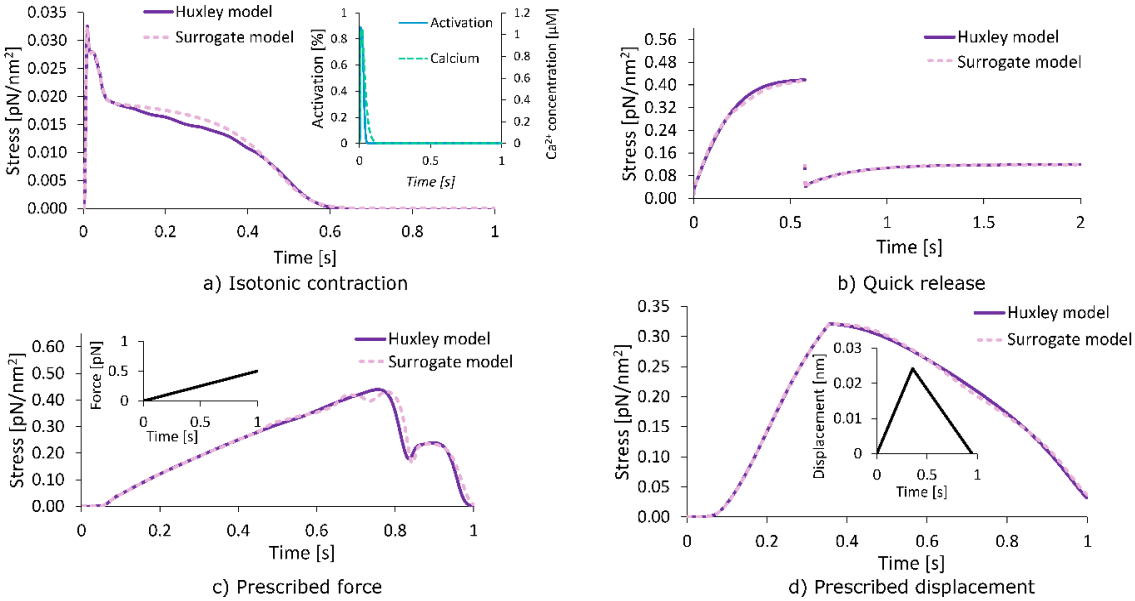


Figure 1: Stresses obtained in numerical experiments used to test the neural network. a) Isotonic contraction), b) quick release, c) prescribed force, d) prescribed displacement (according to [4], with permission)

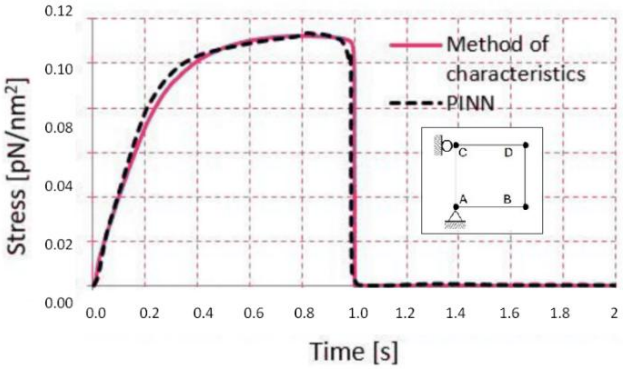


Figure 2: Comparison of PINN and method of characteristics in the case of isotonic muscle contraction

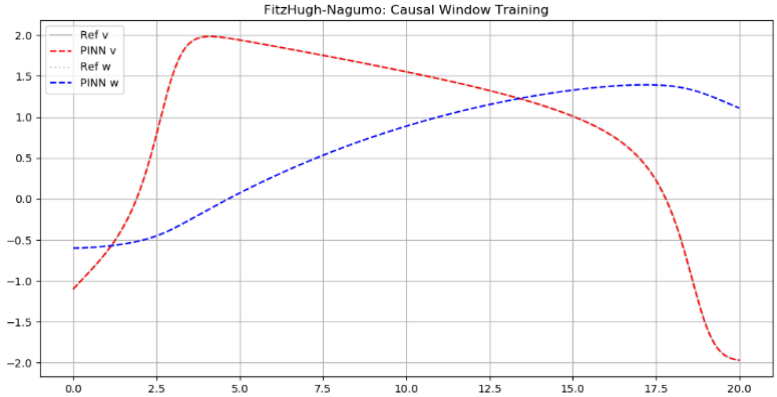


Figure 3: Comparison of solutions to FitzHugh-Nagumo system obtained by PINN and RK45.

The implementation of Physics-Informed Neural Networks (PINNs) for the FitzHugh-Nagumo (FHN) systems successfully resolved the steep gradients associated with cardiac action potentials. A critical finding in the results was the necessity of the sequential causal training approach. Without this strategy, the networks often struggled to converge on the rapid "spikes" of depolarization. By partitioning the time domain, the PINN was able to respect the temporal causality of the system, leading to a drastic reduction in residual error for the voltage and gating variables (Figure 3). In the FHN test cases, the PINN architecture effectively captured the recovery phase and the refractory period of the excitable media, which are essential for identifying potential triggers for cardiac arrhythmias. In summary, GRU surrogates are compact, fast at inference, and integrate directly into existing FE solvers as material subroutines, but require a pre-computed training library and may generalize poorly outside the training distribution. TCNs offer parallelizable training and competitive inference speed, but their receptive field must be carefully tuned to the sequence length. PINNs require no training data beyond the governing equations and are inherently physically consistent, but training convergence on stiff systems (e.g., rapid action potential spikes) requires careful use of sequential causal training, and inference is currently slower than data-driven surrogates due to input tensor size.

4. CONCLUSION

The findings presented in this study underscore the potential of integrating Artificial Neural Networks with traditional biophysical modeling to overcome the computational barriers in cardiac simulations. The results demonstrate that neural networks are effective at emulating the history-dependent behavior of Huxley-type muscle models. By capturing the temporal evolution of stress and stiffness, these data-driven surrogates enable multi-scale finite element simulations that are significantly faster than those relying on sequential numerical integration, without compromising physiological accuracy. Furthermore, the application of Physics-Informed Neural Networks (PINNs) provides a robust alternative for modeling electrophysiological systems and muscle mechanics without the need for extensive pre-computed datasets. By embedding governing partial differential equations directly into the network's loss function, PINNs ensure that the predicted action potentials and contraction dynamics remain physically consistent. While biophysical solvers remain necessary for baseline validation, the accelerated inference provided by neural surrogates bridges the gap between raw scientific simulation and actionable prescriptive analytics. In conclusion, the synergy between data-driven surrogates, and physics-informed learning, provides a scalable and efficient pathway for real-time cardiac biomechanical simulations.

ACKNOWLEDGEMENTS: This study has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101080905 (STRATIFYHF). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or granting authority – European Health and digital executive agency (HADEA). Neither the European Union nor the granting authority can be held responsible for them. This research was also supported by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia, contract number [451-03-33/2026-03/200378 (Institute of Information Technologies, University of Kragujevac) and 451-03-34/2026-03/200107 (Faculty of Engineering, University of Kragujevac)]. AI Use Disclosure: During the preparation of this manuscript, the authors used AI-assisted tools for language editing and text refinement purposes. All scientific content, results, analyses, and conclusions are the sole responsibility of the authors, who have reviewed and verified the final manuscript.

REFERENCES

[1] Huxley, A. F. (1957). Muscle structure and theories of contraction. *Progress in Biophysics and Biophysical Chemistry*, *7*, 255–318.

[2] Luo, C. H., & Rudy, Y. (1991). A model of the ventricular cardiac action potential: Depolarization, repolarization, and their interaction. *Circulation Research*, *68*(6), 1501–1526.

[3] Stojanovic, B., Milosevic, M., Kojic, M., & Filipovic, N. (2021). Multiscale finite element simulation using the Huxley muscle model. *Journal of the Serbian Society for Computational Mechanics*, *15*(1), 1–23.

[4] Milićević, B., Ivanović, M., Stojanović, B., Milošević, M., Kojić, M., & Filipović, N. (2022). Huxley muscle model surrogates for high-speed multi-scale simulations of cardiac contraction. *Computers in Biology and Medicine*, *149*, Article 105963. https://doi.org/10.1016/j.compbiomed.2022.105963

[5] Whittaker, D. G., Ni, H., Benson, A. P., Hancox, J. C., & Zhang, H. (2020). Calibration of ionic and cellular cardiac electrophysiology models. *WIREs Systems Biology and Medicine*, *12*(3), Article e1482. https://doi.org/10.1002/wsbm.1482

[6] Kojic, M., Milosevic, M., Ziemys, A., & Filipovic, N. (2022). *Computational models in biomedical engineering: Finite element models based on smeared physical fields: Theory, solutions, and software*. Amsterdam: Academic Press (Elsevier). ISBN: 978-0-12-821909-2

[7] Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, *378*, 686–707. https://doi.org/10.1016/j.jcp.2018.10.045

[8] Snoek, J., Larochelle, H., & Adams, R. P. (2012). Practical Bayesian optimization of machine learning algorithms. In F. Pereira, C. J. Burges, L. Bottou, & K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems 25 (NIPS 2012)* (pp. 2951–2959). Curran Associates, Inc.

[9] Lu, L., Jin, P., Pang, G., Zhang, Z., & Karniadakis, G. E. (2021). Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. Nature Machine Intelligence, 3(3), 218-229.

Original Research Article *https://doi.org/10.5281/zenodo.23397542*

INTEGRATING BIOMARKERS, VOICE ANALYSIS, AND SENSORS: STRATIFYHF DECISION SUPPORT PLATFORM

Tijana Geroski*[1,2][0000-0003-1417-0521], Anđela Blagojević[1,2][0000-0002-8652-3827], Ognjen Pavić[2,3][0000-0003-2533-1079], Lazar Dašić[2,3][0000-0002-8055-100X], Nemanja Marković[1,2][0009-0003-8991-780X], Nduka Okwose[4][0000-0003-0145-2538], Đorđe Jakovljević[4][0000-0003-2686-6542], Nenad Filipović[1,2][0000-0001-9964-5615]

[1]Faculty of Engineering, University of Kragujevac, Serbia
[2]Bioengineering Research and Development Center (BioIRC), Kragujevac, Serbia
[3]Institute for Information Technologies, University of Kragujevac, Serbia
[4]Institute for Cardio Metabolic Medicine, University Hospitals Coventry and Warwickshire, UK
*Corresponding author, e-mail: tijanas@kg.ac.rs

Abstract: *Heart failure (HF) remains a leading cause of morbidity and mortality worldwide, yet early detection in primary care settings is challenging due to non-specific symptoms and limited diagnostic resources. This study proposes an artificial intelligence-based decision support system integrating clinical biomarkers, voice analysis, and wearable sensor data for early HF detection and risk stratification. The system uses data from retrospective and prospective studies from clinical partners under STRATIFYHF project. Machine learning models, primarily based on random forest and XGBoost algorithms, were trained for early HF detection using primary care parameters, LVEF-based classification (HFpEF, HFmrEF, HFrEF), and future risk prediction. Voice biomarker extraction employed a multi-test recording protocol, capturing time-domain, pitch, loudness, spectral, and perturbation features linked to cardiorespiratory compromise. A custom wearable device integrating ECG and seismocardiogram (SCG) sensors was developed with dual-core architecture, achieving high-fidelity signal acquisition. The platform roadmap demonstrates significant potential for transforming HF management, enabling accessible screening across primary and secondary care settings.*

Keywords*: heart failure, blood biomarkers, voice, wearable sensors, artificial intelligence*

1. INTRODUCTION

Heart failure (HF) is characterized by the heart's inability to pump enough blood to meet the body's metabolic demands [1]. The development and use of new technologies and user-friendly tools for HF patients are revolutionizing the detection of HF, offering significant potential to reduce healthcare costs and facilitate self-management of patients. Despite advances in biomarker discovery and wearable technology, HF risk assessment remains largely dependent on clinician expertise and clinical encounters. Artificial intelligence offers potential to integrate diverse data streams, from blood biomarkers to novel non-invasive biomarkers, into actionable clinical insights that can transform early detection strategies.

2. MATERIALS AND METHODS

We propose a decision support system that combines machine learning algorithms applied to clinical biomarkers, acoustic voice features, and physiological sensor data to enable early identification and risk stratification of HF patients, motivated by the multi-modal systems for cardiovascular diseases [2].

2.1. Early detection and Risk stratification model

2.1.1 Dataset

The dataset was compiled under the STRATIFYHF project* across five European clinical institutions, including longitudinal data from 6,039 patients (3,911 with HF, 2,128 without) across 1 to 4 clinical visits, yielding 11,314 diagnostic records (9,000 HF-positive, 2,314 HF-negative). Data included physical examination findings, blood

* https://stratifyhf.eu/

biomarkers, ECG and echocardiography results, and medical history. For left ventricle ejection fraction (LVEF) classification, the dataset tracked 742 HF with preserved ejection fraction (HFpEF), 386 HF with mid-range ejection fraction (HFmrEF), and 608 HF with reduced ejection fraction (HFrEF) patients, with longitudinal sampling capturing disease progression for a training set of 2,745 samples (1,159 HFpEF, 554 HFmrEF, 1,032 HFrEF).

2.1.2 Proposed methods

Early detection utilized primary care data (physical examination, blood biomarkers, medical history), while risk stratification classified patients by LVEF (HFrEF, HFmrEF, HFpEF) using random forest. Tabular data from multiple visits were restructured such that each visit/timepoint represented a separate row. Features with >50% missing values were excluded, rows without HF/LVEF ground-truth labels were removed, categorical variables were integer-encoded, and missing values were imputed using Multiple Imputation by Chained Equations (MICE) with separate imputers for categorical and continuous variables. The final dataset contained 202 features, including 198 clinical variables derived from echocardiography, physical examination, and demographic data, alongside anonymized patient identifiers and visit information. Data were split at the patient level into 80% training and 20% testing subsets to prevent leakage across visits and ensure that all samples from a given patient were present in only one subset. The classification pipeline was implemented in two stages: first, HF versus non-HF detection, followed by subclassification into HFpEF, HFmrEF, and HFrEF for HF-positive patients. Internal validation was performed on the held-out test cohort. Additionally, for patients initially suspected but not diagnosed with HF, an XGBoost model predicted future HF risk at time point 3 based on data from time points 1 and 2.

2.2. Extracting Voice as non-invasive biomarker for HF

Heart failure affects respiratory and cardiovascular systems, influencing voice through fluid retention and pulmonary congestion [3]. This congestion alters airflow through vocal cords, causing breathiness, hoarseness, or gurgling sounds [4]. A multi-test voice recording protocol (Table 1) was designed to capture vocal capabilities, with feature extraction preferred over raw audio analysis.

Table 1. Recording protocol overview

Task description	Test type	Primary objective
Reading of the text out in loud in the normal speaking tempo	Verbal	To assess controlled speech patterns, articulation clarity, prosody, and fluency under a standardized linguistic load.
Free speech (30 seconds monologue describing previous day)	Verbal	To evaluate naturalistic, spontaneous speech production, including vocal effort, fluency, and potential cognitive-linguistic impacts.
Number counting up (1 to 30, as fast and as accurate as possible)	Verbal	To probe articulatory precision, speech motor control, and respiratory endurance during rapid, repetitive vocalizations.
Number counting down (30 to 1, as fast and as accurate as possible)	Verbal	To assess articulatory precision, speech motor control, and respiratory endurance under repetitive vocalizations, with an added mild cognitive load.
Maximum Phonation Test (holding "aaa" for as long as possible, 3 times)	Vocal	To directly measure glottal efficiency, vocal fold integrity, and the capacity of respiratory support for phonation.

By identifying voice biomarkers indicative of HF [5], healthcare providers can detect the condition in its early stages, allowing for timely diagnosis and treatment. Feature extraction reduces computational complexity, filters noise and improves model generalization compared to raw audio signals.

2.3. Wearable sensors for HF diagnosis

Third aspect of investigation is using designing a wearable that can be used for real-time continuous monitoring and screening. A single, compact, multipurpose sensory device is designed integrating electrocardiogram (ECG) and seismocardiogram (SCG) to measure electrical activity of the heart and physical recoil caused by blood pumping. The proposed architecture integrates specialized biomedical front-end with a dual-core processing unit to achieve high-fidelity, synchronous acquisition of electrocardiographic (ECG) and seismocardiographic (SCG) data. Central to the system is a Raspberry Pi Pico, which utilizes its dual-core architecture to effectively decouple time-critical hardware interfacing from high-level data management, ensuring robust and efficient processing. The electrical components of the cardiac cycle are digitized via an ADS1293 analog front-end, while the corresponding mechanical vibrations are captured by MPU-6000 MEMS accelerometer mounted on a sternum. To ensure morphological integrity and temporal alignment, both signals are sampled at 200 Hz using an event-driven firmware strategy. Event-driven firmware with hardware interrupts ensures temporal alignment and eliminates polling latency. To mitigate the influence of noise on physiological signals such as ECG and SCG, digital filtering is essential for preserving signal integrity. These recordings are frequently corrupted by motion artifacts, power-line interference, and respiratory-induced baseline wander. To address this, a bandpass filter is employed, while a high-pass filter with a 0.5 Hz cutoff eliminates low-

frequency drifts and electrode polarization, and a 40 Hz low-pass filter suppresses high-frequency disturbances, such as electromyographic noise and environmental interference.

3. RESULTS AND DISCUSSION

We present the results in three phases: results from early detection and risk stratification models, results related to extraction of vocal biomarkers and results related to sensor data processing.

3.1. Early detection and Risk stratification results

Early detection and risk stratification models within the decision support system were graded using standard evaluation metrics, which include F1 score, precision, recall and classification accuracy. Results for early detection of heart failure are shown in
Table **2** and Figure 1.

Table 2. Early detection model performance (binary classification)

Class	F1 Score	Precision	Recall	Accuracy
HF not present	92%	92%	92%	97%
HF present	98%	98%	98%	

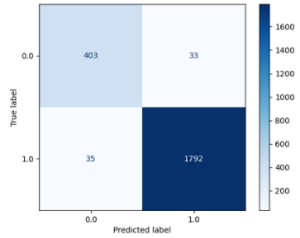


Figure 1: HF classification confusion matrix

Early detection based solely on primary care parameters shows great promise for uncovering signs of heart failure early without the need for secondary care characteristics. Results for HF type classification are shown in Table 3 and Figure 2.

In the three-class classification, the HFmrEF class cannot be distinguished well enough due to the fact that it has less training sample, as well as the nature of the class in practice. Because HFmrEF class is borderline, it is very similar to both HFpEF and HFrEF classes.

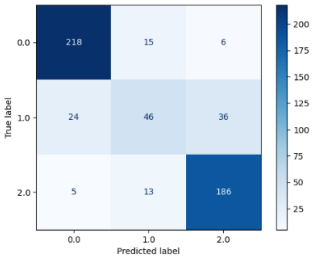


Figure 2: Three-class classification model for HFpEF/HFmrEF/HFrEF

Table 3. Results on early detection of HF types (multiclass classification)

Three class risk stratification model				
Class	**F1 Score**	**Precision**	**Recall**	**Accuracy**
HFpEF	90%	88%	90%	82%
HFmrEF	51%	62%	51%	
HFrEF	86%	82%	86%	
Classification for HFpEF vs (HFrEF and HFmrEF)				
Class	**F1 Score**	**Precision**	**Recall**	**Accuracy**
HFpEF	96%	94%	98%	96%
HFmrEF + HFrEF	95%	98%	93%	
Classification for HFrEF vs HFmrEF				
Class	**F1 Score**	**Precision**	**Recall**	**Accuracy**
HFmrEF	81%	89%	75%	87%
HFrEF	90%	86%	94%	

Risk stratification results are shown in **Error! Not a valid bookmark self-reference.**.

Table 4. Risk stratification model evaluation

Class	Precision	Recall	F1 Score
Low risk of HF	78%	82%	80%
High risk of HF	99%	98%	99%
Macro-average	88%	90%	89%

These results show that the risk stratification module within the decision support system can uncover high risk patients and raise alarms. This is very important since it can show the clinicians which patients require immediate treatment to divert the progression of the disease/stop development of HF. Proposed model is validated via five-fold cross validation through retrospective and prospective clinical data.

3.2. Voice as non-invasive biomarker for HF

Five types of features have been extracted and will be further used as training data for ML models (Figure 3):

- **Time-domain/prosodic:** Total duration, phonation time, pause metrics (number >100ms, >500ms, longest), phonation segment lengths and statistics, capturing respiratory compromise and speech motor control
- **Pitch/fundamental frequency**: Mean and SD of pitch and f0, reflecting vocal fold mass, stiffness, and neuromuscular control
- **Loudness/amplitude**: Mean and SD of loudness, shimmer, indicating respiratory muscle weakness and glottal efficiency
- **Spectral/timbral**: MFCCs and CPP (mean, SD), describing vocal tract resonance and vibration efficiency
- **Perturbation**: Jitter of longest phonation, quantifying vocal fold irregularities.

This method is in its early integration stage and while extracted features exhibit potential in predicting HF, current models are still updated and require further validation. Voice features will be tested as standalone input features, as well as an addition to the collected information from patient's health records. These features will be integrated into the multi-modal decision support system to enhance diagnostic accuracy and disease monitoring.

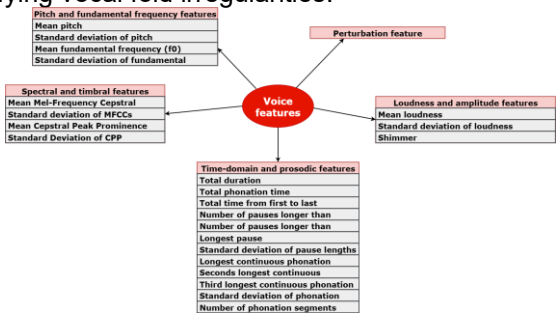


Figure 3: Extracted vocal features

3.3. Wearable sensors for HF diagnosis

The developed system demonstrated high fidelity in the acquisition and processing of both ECG and SCG signals. The filtered ECG waveforms exhibit clearly defined P, QRS, and T complexes, showing a strong morphological correlation with established reference signals and confirming the reliability of the electrical measurement pipeline across all channels. Simultaneously, the filtered SCG recordings reveal distinct mechanical oscillations corresponding to the systolic and diastolic phases of the cardiac cycle. The close structural and spectral alignment between the captured SCG data and standard reference recordings validates the efficacy of the accelerometer coupling, the 200 Hz sampling rate, and the specific digital filtering strategy employed to isolate chest wall vibrations. While currently validated under controlled laboratory conditions, the robustness of the acquisition pipeline and the fidelity of the recorded data suggest a scalable foundation for future development. Subsequent optimization could facilitate the transition of this technology into a cost-effective, portable solution for continuous remote monitoring, significantly enhancing early intervention strategies and the longitudinal management of chronic cardiovascular conditions.

4. CONCLUSION

The STRATIFYHF decision support system successfully integrates multiple data modalities, clinical biomarkers, voice features, and wearable sensor signals, into a comprehensive AI-driven decision support system capable of early heart failure detection and accurate risk stratification across care settings. This approach represents a significant step toward personalized, preventive cardiovascular care, with potential to reduce healthcare costs and improve patient outcomes through timely intervention before clinical decompensation occurs.

Acknowledgements: This paper has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101080905 (STRATIFYHF project). The research is supported by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia, contract numbers [451-03-33/2026-03/200107 (Faculty of Engineering, University of Kragujevac) and 451-03-33/2026-03/200378 (Institute for Information Technologies, University of Kragujevac)].
Disclosure of interests: The author(s) declare no competing interests.
AI-use disclosure statement: AI tools were only used for language editing and grammar checking.

REFERENCES

[1] G. Agnetti, et al. "New insights in the diagnosis and treatment of heart failure," *BioMed research international,* vol. 1, p. 265260, 2015.

[2] S. Miao, et al. "The construction and application of a clinical decision support system for cardiovascular diseases: multimodal Data-Driven development and validation study," *JMIR medical informatics,* vol. 13, p. e63186, 2025.

[3] O. M. Murton, et al. "Acoustic speech analysis of patients with decompensated heart failure: A pilot study," *The Journal of the Acoustical Society of America,* vol. 142, no. 4, pp. EL401-EL407, 2017.

[4] E. E. Mohamed and A. El Maghraby, "Voice changes in patients with chronic obstructive pulmonary disease," *Egyptian Journal of Chest Diseases and Tuberculosis,* vol. 63, no. 3, pp. 561-567, 2014.

[5] A. Fuller, et al. "Voice recognition as an innovative method to identify and monitor heart failure–a review," *Heart Failure Reviews,* vol. 31, no. 1, p. 58, 2026.

Original Research Article *https://doi.org/10.5281/zenodo.23397627*

APPLICATION OF FUZZY LOGIC AND RISK MATRIX IN THE PRELIMINARY ASSESSMENT OF INDIVIDUALS' CREDITWORTHINESS

Ilija Ivezić[1,2][0009-0004-1981-1249], Dragan Pamučar[2][0009-0004-2488-9403], Bisera Andrić Gušavac[2][0000-0001-9438-5906]

[1]Transfera Transport and Logistics, Belgrade, Republic of Serbia

[2]University of Belgrade Faculty of Organizational Science

*Corresponding author, e-mail: ilija777ivezic@gmail.com

Abstract: *This paper suggests the application of fuzzy logic and the risk matrix in the preliminary assessment of individuals' creditworthiness. The main goal of the research is development of a model that enables a more flexible and transparent evaluation of credit risk in conditions of inaccuracy and partial uncertainty of input data. In the paper, the Mamdani fuzzy model is preliminary developed in the MATLAB environment, where the input variables are length of service, monthly income, loan amount, repayment period and age, while the output of the model represents the applicants' credit score. Based on the defined membership functions and fuzzy rules, a preliminary assessment of creditworthiness can be performed, which enables the classification of applicants according to the level of risk and the definition of appropriate corrective actions. The combination of fuzzy logic and risk matrix represents a useful methodological framework for the preliminary assessment of creditworthiness, with significant potential for further development and application in the field of credit decision support.*

Keywords*: Fuzzy Logic, Risk Matrix, Creditworthiness, MATLAB, Credit Risk.*

1. INTRODUCTION

In modern business conditions, risk management is one of the most important elements of decision-making, especially in the financial sector. The assessment of individuals' creditworthiness plays a particularly important role, because its accuracy depends on the long-term financial stability of the credit user. The approvement of a loan depends on several factors that jointly affect the overall risk assessment. Among the most important factors are the amount of income, length of service, the age of the client, the amount of the requested loan and the planned repayment period (Basel Committee on Banking Supervision, 2000; Thomas et al., 2017; Van Gestel & Baesens, 2009).

In practice, the assessment of creditworthiness is not always completely precise and strictly defined, as it includes elements of uncertainty, experiential judgment and subjective assessment. Recent studies have explored the application of intelligent analytical and decision-support approaches in complex systems characterized by uncertainty and multiple influencing factors (Barati Bagherabad et al., 2026). The boundary between acceptable and unacceptable risk is often not clearly defined but depends on the interrelationship of several criteria. Fuzzy logic is suitable for imprecise and linguistically expressed data and represents a good approach for modeling such problems (Zadeh, 1965; Ross, 2010). In addition to fuzzy logic, an important tool in the risk assessment process is the risk matrix. Its main advantage is reflected in the fact that it enables a simple and transparent classification of observed cases according to the level of risk, considering the probability of occurrence of an undesired event and the impact of its consequences. In the context of credit assessment, the risk matrix can serve as an effective mechanism for classifying individuals into low, medium and high-risk zones, thus facilitating the understanding of their credit position and the need for possible corrective measures (International Organization for Standardization, 2018; International Electrotechnical Commission, 2019).

2. METHODOLOGY

2.1. Mamdani Fuzzy Logic Model

The Mamdani model stands is widely used in theory and practice as a method for mapping human experience into algorithmic frameworks, thus bridging the gap be-tween rigid binary systems and the fluidity of real processes (Zadeh, 1965; Mamdani & Assilian, 1975; Ross, 2010). Unlike rigid mathematical functions, this approach operates with linguistic variables that allow complex dependencies to be described through intuitive 'if-then' rules. The key advantage of the selected model in this re-search is reflected in its ability to successfully treat uncertainty and noise in the data, transforming subjective assessments into precise control signals. Recent research has investigated the use of intelligent and analytical models in decision-making environments requiring flexibility and interpretability (Sarkar & Goswami, 2026).

Within the framework of the preliminary developed fuzzy model, five input variables are defined that describe the basic characteristics of an individual relevant for the preliminary assessment of creditworthiness, as well as one output variable that represents the credit rating. All membership functions are modeled by trapezoidal functions, to provide a gradual transition between adjacent linguistic values and a more realistic representation of uncertainty. Input variables that are defined are:

- Length of service,
- Monthly income, expressed in hundreds of euros,
- Loan amount, expressed in thousands of euros,
- Repayment period,
- Age.

The output variable of the model is the credit score. This variable can be divided into five linguistic categories: very bad, bad, medium, good and very good. The output defined in this way enables sufficiently precise differentiation of creditworthiness levels, while maintaining the interpretability of the model.

2.2. RISK MATRIX

Risk assessment methods are procedures and analytical approaches that determine the probability of adverse events and the severity of their consequences. Their application enables systematic risk recognition, analysis, and evaluation, providing a basis for appropriate management decisions. Qualitative, quantitative, and combined methods are usually distinguished, whereby the choice of a specific method depends on the available data, the complexity of the system, and the goals of the assessment itself.

The risk matrix is a standardized instrument for qualitative risk assessment, the basic function of which is to enable a systematic assessment of the relationship between the probability of the occurrence of an unwanted event and the severity of its consequences. Its application contributes to the structured classification of risks according to the degree of importance, which facilitates the determination of priorities and the selection of adequate management measures (International Organization for Standardization, 2018; International Electrotechnical Commission, 2019). In this sense, the risk matrix not only serves as an analytical tool but also as a basis for more rational and transparent decision-making. The methodological value of the risk matrix is reflected in the fact that it enables a clear unification of the key elements of the assessment into a single analytical framework. By connecting the probability assessment and the consequence assessment, the risks are classified into appropriate categories, which indicate the level of their acceptability and the need for further action. Such an approach is particularly important in conditions of limited resources, when it is necessary to clearly determine which risks require immediate attention and which can be monitored and controlled within regular activities.

In addition to the analytical function, the risk matrix also has a pronounced practical value, as it enables a simpler presentation of the assessment results to different participants in the management process. Its transparency contributes to a better understanding of the nature and intensity of risks, as well as more consistent planning of preventive and corrective measures. Although its application implies a certain degree of simplification of complex phenomena, the risk matrix remains one of the most common and useful tools in modern risk management.

For the preliminary assessment of individuals' creditworthy, credit risk assessment can be used and can consist of two components:

- Probability of problems in repayment;
- The impact of the financial burden.

The probability of repayment problems is an assessment of the possibility that the debtor will not be able to pay off his financial obligations in an orderly, timely manner and in full. Within the risk matrix, this criterion serves to assess the frequency or likelihood of repayment difficulties, with a higher degree of probability indicating greater credit risk and the need for more careful management of such exposures (Basel Committee on Banking Supervision, 2000; Thomas et al., 2017; Van Gestel & Baesens, 2009). From the bank's perspective, this is a key indicator for the decision to accept a loan application.

3. DISCUSSION AND CONCLUSION

The presented idea indicate that the combination of fuzzy logic and the risk matrix can represent a useful framework for the preliminary assessment of the individual's creditworthiness. The special value of the proposed approach can be reflected in the fact that it allows variables that are not always strictly determined in practice and mutually independent, such as length of service, monthly income, loan amount, repayment period, and age, to be included in the assessment process. In this sense, the fuzzy model enables a more flexible assessment of credit eligibility than approaches based solely on sharp boundaries and binary decisions, while the risk matrix contributes to a clearer classification of users according to the level of assessed risk and easier definition of recommended corrective measures.

Some preliminary results of the fuzzy model did not remain at the level of numerical assessment but were connected to the risk matrix and a set of corrective actions. These results are not presented in the paper, due to the ongoing work on this paper. Some preliminary experiments showed that it is possible to transition from the assessment itself to recommendations for further action, which makes the model practically more relevant. In other words, the user not only receives information about which risk zone they belong to, but also guidelines on which elements of their financial profile should be improved, such as the amount of participation, the amount of the requested loan, the stability of employment, or the ratio of income and workload. This interpretability is one of the most important advantages of the proposed approach, as it enables the model to be comprehensible to both the user and the decision maker. Recent methodological studies have highlighted the importance of interpretable analytical frameworks for supporting complex decision-making processes (Fujita et al., 2025)

Further development of the work could be directed in several directions. The first and most important direction refers to the validation of the model on real data on credit users, which would make it possible to more accurately assess its accuracy, robustness, and discriminatory power. The second direction implies a finer calibration of membership functions and fuzzy rules, especially in those parts of the input area where greater deviations are observed, especially with lower and medium monthly incomes, as well as with certain combinations of loan amounts and repayment periods. The third direction of development is the expansion of the model with additional criteria, such as existing indebtedness, credit history, type of employment, marital status, or the number of dependent household members, which would achieve a more complete and realistic assessment of creditworthiness. In addition, it would be useful to compare the performance of the developed fuzzy system with classical statistical and machine learning methods, to more clearly determine the advantages and limitations of each approach.

In conclusion, it can be assessed that the application of fuzzy logic and the risk matrix represent a justified and methodologically useful approach to the preliminary assessment of the individuals' creditworthiness. The developed model shows that it is possible to connect imprecise and partly subjective input information with a structured system of risk classification and recommendations for further action.

REFERENCES

[1] Barati Bagherabad, M., Rivandi, E., & Javid Mehr, M. (2026). Machine Learning for Analyzing Effects of Various Factors on Business Economic. Applied Decision Analytics, 2(1), 41-54. https://doi.org/10.66972/ada2120269

[2] Basel Committee on Banking Supervision. Principles for the management of credit risk. Bank for International Settlements (2000)

[3] Fujita, T., Mehmood, A., & Ghaib, A. A. (2025). Hyperfuzzy OffGraphs: A Unified Graph-Based Theoretical Decision Framework for Hierarchical Under Off-Uncertainty. *Applied Decision Analytics*, *1*(1), 1-14. https://doi.org/10.66972/ada1120251

[4] International Electrotechnical Commission. (2019). IEC 31010:2019 Risk management—Risk assessment techniques (2008)

[5] International Organization for Standardization. ISO 31000:2018 Risk management—Guidelines (2018)

[6] Mamdani, E. H., & Assilian, S. An experiment in linguistic synthesis with a fuzzy logic controller. International Journal of Man-Machine Studies, 7(1), 1–13 (1975)

[7] Ross, T. J. Fuzzy logic with engineering applications (3rd ed.). Wiley (2010)
[8] Sarkar, A., & Goswami, S. S. (2026). A Review of the Application of MCDM Methods in Business Analytics. *Applied Decision Analytics*, *2*(1), 150-180. https://doi.org/10.66972/ada21202614
[9] Thomas, L. C., Crook, J. N., & Edelman, D. B. Credit scoring and its applications (2nd ed.). Society for Industrial and Applied Mathematics (2017)
[10] Van Gestel, T., & Baesens, B. Credit risk management: Basic concepts: Financial risk components, rating analysis, models, economic and regulatory capital. Oxford University Press (2009)
[11] Zadeh, L. A. Fuzzy sets. Information and Control, 8(3), 338–353 (1965)

Original Research Article *https://doi.org/10.5281/zenodo.23397680*

INNER GOVERNANCE: UTILIZING TRAIT EMOTIONAL INTELLIGENCE TO ANCHOR RESILIENCE IN ORGANIZATIONS

Teklehaimanot EMBAYE*[1][0000-0002-1306-392X]
[1]Affiliation, Dept. E-Business, Faculty of Organizational Sciences, University of Belgrade, Belgrade, Serbia
*Corresponding author, e-mail: teklehaimanot.yemane@gmail.com

Abstract: *Permanent disruption has become the structural operating condition for most organizations - not an interval between stable states, but a continuous environment in which the affective and cognitive resources sustaining adaptive performance are under persistent pressure. This paper introduces Inner Governance Framework as the multilevel self-regulatory mechanism through which Trait Emotional Intelligence theory is mobilized across individual, collective, and organizational levels to produce what we term Intelligent Organizational Resilience. Intelligent Organizational Resilience is the capacity to reconfigure under disruption and emerge with greater adaptive capability than before. Drawing on Conservation of Resources theory, Broaden and Build theory, Dynamic Capabilities, and the psychological safety literature, we present a theoretically grounded three-level framework with two empirically distinct emergence mechanisms and a bidirectional organizational learning loop. Individual level Trait Emotional Intelligence Questionnaire data across diverse professional samples provides preliminary evidence for the model's foundational layer. Cross-level and longitudinal testing remain the immediate future research priority. The framework offers both a theoretical contribution to organizational resilience scholarship and a diagnostic pathway for corporate practices.*

Keywords*: Trait Emotional Intelligence, Inner Governance, Organizational Resilience, Dynamic Capabilities, TEIQue, Multilevel Self-Regulation, VUCA, Permanent Disruption, Psychological Safety*

1. INTRODUCTION

With the rapid adoption of new digital technologies, organizational competitiveness and related economic dynamism are being redefined continuously. What were considered as primary engines of organizational competitive advantage like technological superiority, capital buildup, and supply chain flexibility are now being tested with the rapid deployment of machine learning algorithms and Gen-AI (Lanfranchi, 2026). The focus has now shifted towards intentional cultivation and development of human-centric skills like empathy, active listening, resilience, and emotional intelligence as examined by (Agarwal, 2025). With the commoditization of cognitive tasks and routine operations by Gen-AI, anchoring human-centric skills and most importantly resilience in digitally transforming organizations has become crucial more than any other time. In the contrary, as explained by (Akram, 2025), tasks that require interaction, judgment, and lived experiences are far from being fully or partially transformed by Gen-AI making technological investment and operational alignment alone inadequate to deliver expected productivity and organizational sustainability.

While human-centric skills are frequently categorized as durable competences that endure across technological cycles and geoeconomic changes, recent longitudinal research exposes the fragility of such skills and the constraint in resilience as analyzed by (Freedman, 2025) and (World Economic Forum, 2025). Technological disruption in general and Gen-AI in particular is one aspect of the disruption. Organizations today operate under conditions of continuous disruption rather than episodic crisis, placing sustained pressure on the affective and cognitive resources that underpin adaptive performance.

This paper presents Inner Governance Framework, a multilevel self-regulatory model grounded in Trait Emotional Intelligence (Trait EI) theory (K.V, 2010), through which individual, collective, and organizational capacities are mobilized to produce Intelligent Organizational Resilience (IOR). IOR is defined as the ability to reconfigure under pressure and emerge with greater adaptive capability than before (Teece, 1997).

Currently, this serious fragility is being intensified by a progressive atrophy of higher-order thinking, cognitive skills, and emotional regulation capabilities which is the result of overreliance on AI tools and digital assistants in the workforce (Gulyamov, 2025). In response to this eminent threat, organizations must deliberately assess, architect, and empower methodology to develop human-centric skills in general and emotional intelligence in particular to stay resilient, relevant and competitive (Azarov VN, 2025) and (Ruttenberg, 2026).

2. THEORETICAL CONTEXT

Emotional intelligence plays a dual role, first as a core skill in the human-centric skills domain and second as a methodology to create awareness, measure efficacy, and suggest grounded development strategies. There are a number of popular models of emotional intelligence both in the scientific literature and the corporate application of which, Trait EI, pioneered by K.V. Petrides stands out as it is based on comprehensively validated scientific theory and rigorous application (K.V, 2010). Trait EI locates emotional intelligence at the lower level of personality hierarchies. In contrast to the other model which claim emotional intelligence as an objective cognitive processing speed, Trait EI is defined as a constellation of emotional perceptions assessed through a rigorous self-report methodology, measured by Trait Emotional Intelligence Questionnaire (TEIQue). TEIQue measures individuals genuine believe about their own emotional capabilities and how they perform in emotion-laden, relentless environment. As examined by (Andrei, 2016), (Nikolaou, 2002), and (O'Boyle Jr, 2011), Trait EI demonstrates highly greater predictive validity with regards to workspace behaviors, performance, endurance to burnout, and organizational obligation. For instance, when employees face unexpected organizational transformation initiative or disastrous systems failure, their responses are not dictated by their theoretical knowledge of emotions, but by their emotional self-efficacy and distinctive coping mechanisms deeply rooted in their personality. This research paper introduces Inner Governance Framework, an applied Trait EI methodology by conceptualization of Trait EI as a foundation for resilience at individual, collective and organizational levels. As presented in Table 1, TEIQue consists of fifteen facets, four factors, and a Global Score.

Table 1: Factors and Facets of Trait Emotional Intelligence as Resilience Mechanics

TEIQue Factors and Facets	**High Scorer Self-perceptions**	**Resilience Mechanics for Organizational Resilience**
Factor Well-Being – Facets: Self-Esteem, Trait Happiness, and Trait Optimism	Confident, successful, inherently worthy regardless of externa validation; experiences full satisfaction with life; confident in positive future outcomes.	Provides the foundational emotional capacity to self-worth essential to endure professional failures and algorithmic intervention without identity collapse; sustains long-term intrinsic motivation; optimistic employees are significantly more likely to view artificial intelligence disruptions as opportunities for growth rather than existential threats.
Factor Self-Control – Facets: Emotion Regulation, Impulse Control, Stress Management	Capable of consciously controlling and regulating intense emotional states; expresses rational, deliberate thinking; and endures intense pressure effectively.	Acts as primary governor of cognitive bandwidth; essential in mitigating knee-jerk, reactive decision making, in high velocity digital environment; and directly counters burnout and chronic exhaustion, ensuring continuous sustainability during corporate crises and prolonged organizational transformation.
Factor Emotionality – Facets: Empathy, Emotion Perception, Emotion Expression, and Relationships	Understands own emotions and clearly reads the subtle emotional states of others; easily articulates and communicates internal emotional states; are capable of building, maintaining, and repairing personal and professional bonds; readily considers other people's emotional perspectives.	Serves as diagnosis to human operating system status, allowing leaders to detect team friction; transparency, reducing organizational ambiguity and paranoia; builds social capital and trust networks; prevents social erosion and toxic individualism; ensures cold algorithmic decisions and operational efficiencies are ethically contextualized by their human impacts; endure shared hardships and execute complex, cross-functional technological transformations.
Factor Sociability – Facets Social Awareness, Emotion Management, Assertiveness	Socially knowledgeable, skilled networking, able to read group dynamics; capable of influencing, comforting, inspiring, and motivating others; outspoken, forthright, and willing to defend their personal rights and professional boundaries.	Enables rapid, seamless integration into new teams and cross-functional task forces; forms the core asset of transformational leadership; allows managers to actively down regulate teams' collective anxiety during intense uncertainty; crucial in operational scope creep, managing unreasonable digital expectations and maintaining psychological equilibrium.

Adaptability - an independent facet	Highly flexible; easily and willingly adjusts to novel, unpredictable, or rapidly changing environments	Definitive survival trait in uncertain, volatile, complex, and ambiguous environments; conveniently accelerates the adoption and integration of new digital technology and related strategy implementation.
Self-Motivation - an independent facet	Driven by internal, intrinsic standards and motives rather than external rewards, demonstrating high perseverance and persistence	Enables organizational staff to move from passive default operational mode to an active, deliberate mode of operation, ensuring continuous, self-directed upskilling and transformation initiative, without the need for constant management oversight.

Organizational resilience scholarship has predominantly focused on structural, strategic, and process-level variables like redundancy, agility, modularity, and distributed command, while systematically underspecifying the human affective-regulatory substrate that enables those structures to function under real disruption pressure. The result is a well-described outer architecture for resilience with an underspecified inner mechanism. Inner Governance Framework provides an integrative mechanism for the identification, adjustment, contextual application, and configuration of emotional and cognitive responses of employees under high-pressure organizational ecosystems across diverse professional domains. Through a thorough analysis of psychometric data and organizational case studies, this paper illustrates how organizations can leverage Trait EI to architect resilient teams and organizations, capable of sustaining and thriving in highly volatile ecosystems and permanent disruption.

3. THE INNER GOVERNANCE CONCEPTUAL AND THEORETICAL DESIGN

Inner Governance is formally defined as the set of affective and self-regulatory processes, operating at individual, collective, and organizational levels, that sustain adaptive performance under conditions of continuous environmental pressure, without requiring centralized command or formal governance intervention. The construct is distinct from formal organizational governance in that it is emergent, affective, and self-organized. The Inner Governance Framework developed as a multilevel methodology for executive leadership coaching, talent development, academic counseling, and organizational cultural alignment represents a practical application of Trait EI theory to counter the diverse challenges in modern organizational behavior and disruptive processes in an intelligent way.

The Inner Governance Framework can be applied at three levels: at individual level, collective level, and organizational level as detailed in Table 2. Figure 1 illustrates the details of Level 1 – Individual Inner Governance mechanism.

Table 2: The three-level Inner Governance Architecture

Level	Core mechanism	Governance pole	Integration pole	Instrument
L1 — Individual	Self-regulatory loop	Emotion regulation: TEIQue Self-control facets prevent cognitive bandwidth collapse	Adaptive integration: TEIQue Well-being facets broaden attention and build durable resources (Fredrickson, 2001)	TEIQue full/short form
L2 — Collective	Compilational emergence (Kozlowski & Klein, 2000)	Shared co-regulation: explicit reframing norms, distributed stress absorption, panic contagion suppression	Psychological safety (Edmondson, 1999): voice behavior, collective efficacy, relational trust under load	TPSS, aggregate TEIQue, and Hülsheger et al. CER scale
L3 — Organizational	Transformational emergence via relational coordination (Gittell, 2002)	Sensing and seizing: weak-signal detection, pre-mortems, anticipatory reconfiguration	Reconfiguring and learning: adaptive routines, blame-free debrief, institutionalized ambidexterity (O'Reilly & Tushman, 2004)	ORI (Lee et al., 2013) and Adaptive Process Audit

3.1. Level 1 – Individual Inner Governance:

The structural logic of the loop within Level 1 consists of six phases. The phases are not a sequence that runs from start to finish and then repeats. They describe two simultaneously active subsystems: a

Governance pole and an Adaptive Integration pole that operate in dynamic tension, with each influencing the activity level of the other in real time. The directional flow through Phases 1 to 3 and phases 4 to 6 is a pedagogical simplification of something more accurately described as a continuous regulatory negotiation between containment and expression.

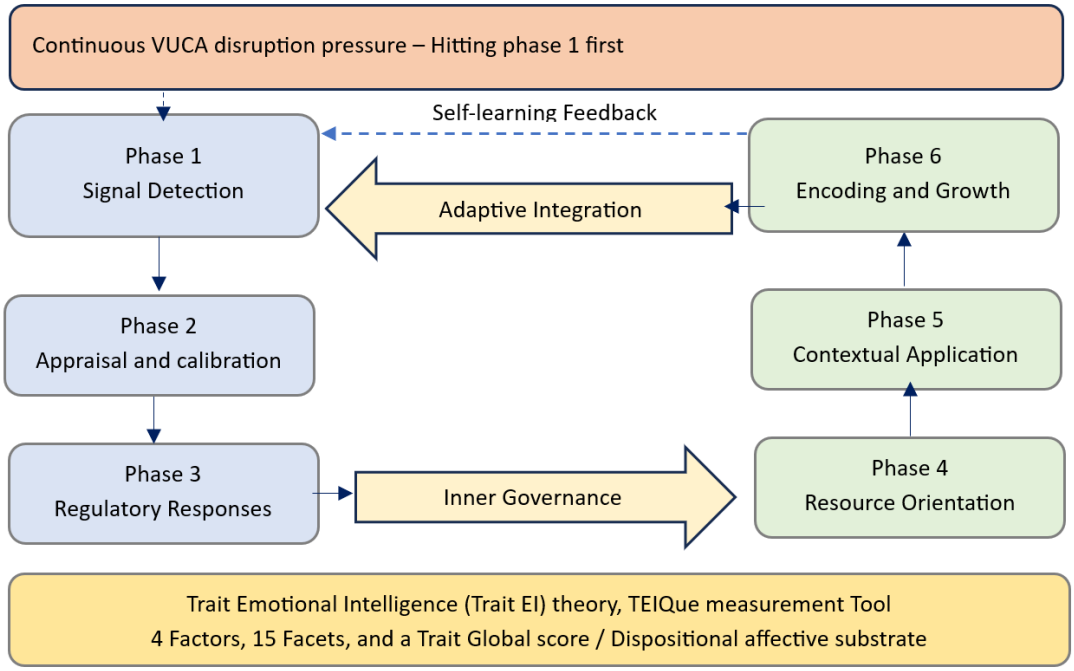


Figure 1: Level 1 – Individual Inner Governance Framework - Theoretical and Conceptual Design

At the foundation of the framework is the Trait EI layer, which enables to move the theoretical concepts to structural engineering by utilizing the four TEIQue factors and the fifteen specific facets. The Governance pole (Phases 1 to 3) is grounded in the Self-control factor of the TEIQue. It comprises emotion perception, cognitive reappraisal, stress management, and impulse inhibition. Its function is not suppression but calibration, ensuring that emotional signals reach the level of conscious processing where they can be used, without overwhelming the attentional bandwidth needed to use them. The Adaptive Integration pole (Phases 4 to 6) draws primarily on the Well-being and Sociability factors: optimism, happiness, empathy, assertiveness, and self-motivation. Its function is expansion rather than containment - building, expressing, applying, and encoding which lead to inner harmony.

The operating context chosen for the model is VUCA: Volatility, Uncertainty, Complexity, and Ambiguity, in digitally accelerating environment. VUCA conditions are continuous. Resilience is therefore not a recovery capacity but an ongoing adaptive intelligence, a structural property of how the organization senses, interprets, and responds to perpetual environmental pressure without degrading its core functioning. This operating context is decisive for the framework's logic. A model designed around discrete crisis events would target recovery performance, whereas a model designed for permanent disruption must target adaptive intelligence - the continuous capacity to sense, regulate, reconfigure, and learn without depleting the human resources that make such intelligence possible.

3.2. Level 2 — Collective Inner Governance

Collective Inner Governance is not aggregated individual Trait EI. It is a genuinely emergent team-level property that arises when individual Inner Governance behaviors interact through emotional contagion, explicit co-regulation, and the gradual sedimentation of those behaviors into shared implicit norms. The emergence mechanism is compilational that the configuration and interaction of individual Trait EI facets within a team create properties absent from any individual member.

3.3. Level 3 – Organizational Inner Governance

Organizational Inner Governance is the activation of Dynamic Capabilities (Teece et al. 1997) under disruption, made possible by the affective resource base established at levels 1 and 2. The sensing, seizing and reconfiguring triad maps cleanly onto the governance-integration polarity. Governance side (Sensing and Seizing): Weak signal detection in processes and workflow, structured process vulnerability audits, pre-

mortems before high-stakes operations, flexible protocol design, and anticipatory reconfiguration. These are vigilance functions which require psychological safety at Level 2 to function, because weak signals are only shared when people feel safe voicing them. Adaptive Integration (Reconfiguring): Adaptive routine activates when standard processes fail. It works by processing memory encoding through blame-free debrief cycles, institutionalization of adaptation as updated standard operating procedures, and organizational ambidexterity - the capacity to exploit current capabilities while simultaneously exploring adaptive alternatives (O'Reilly & Tushman 2004). Diagrams and illustrations for Level 2 and Level 3 of the framework are omitted for space reasons.

These three levels ensure organizational resilience, scaling from level 1 to level 3, while receiving organizational learning feedback from level 3 back to level 1 and 2. Traditionally, human resource development agencies frequently treat resilience as a monolithic vague outcome, rendering it nearly impossible to provide systematic training. By dissecting the facets of trait emotional self-efficacy into specific psychological levers and treating them as sophisticated configuration panel for human adaptability, the Inner Governance Framework enables organizations to identify exact points of intervention and provides relevant alignment prescriptions. As such, the framework's most original contribution is the proposition that Trait EI, through Inner Governance, is the enabler of Dynamic Capabilities under pressure. This explains why technically similar organizations, comparable in strategy, resources, and market position differ radically in adaptive performance during disruption. The difference is affective-regulatory, which is addressed by the multilevel Inner Governance framework.

4. CONCLUSION

The Inner Governance Framework advances organizational resilience scholarship by specifying the affective-regulatory mechanism that existing models leave underspecified: the multilevel process through which Trait Emotional Intelligence, measured via TEIQue, is mobilized across individual, collective, and organizational levels to produce Intelligent Organizational Resilience. Rather than treating resilience as a structural property or a generic human capacity, the framework grounds it in empirically testable cross-level dynamics - compilational emergence, relational coordination, and a bidirectional organizational learning loop that convert disruption exposure into expanded adaptive capability. The diagnostic and developmental pathway it offers moves organizations from reactive coping toward deliberate, self-reinforcing growth.

Acknowledgement and Disclosure: The Authors acknowledge and disclose that they have used an experimental AI tool, Google Scholar Labs to find relevant reference papers, publications, and related summaries.

5. REFERENCES

[1] Agarwal, P. D. (2025). Human-Centric Pedagogies in the Age of Intelligent Systems: Envisioning the Future with Industry 5.0. *Singapore: Springer Nature*, 75-97.
[2] Akram, W. A. (2025). The Role of Generative AI in Driving Sustainable Digital Transformation and Organizational Performance. *The Critical Review of Social Sciences Studies*, 2983-2996.
[3] Andrei, F. S. (2016). The incremental validity of the Trait Emotional Intelligence Questionnaire (TEIQue): A systematic review and meta-analysis. *Journal of personality assessment*, 261-276.
[4] Azarov VN, G. V. (2025). Cognitive Offloading or Atrophy? Risks of Skill Degradation and New Forms of Dependence in the Era of Pervasive Artificial Intelligence. *International Conference on Quality Management, Transport and Information Security, Information Technologies*, (pp. 166-168).
[5] Freedman, J. M. (2025). The Emotional Recession: global declines in emotional intelligence and its impact on organizational retention, burnout, and workforce resilience. *Frontiers in Psychology*.
[6] Gulyamov, S. S. (2025). The Problem of Cognitive Atrophy in the Environment of Permanent Digital Assistance: Developing a Methodology for Stimulating Independent Thinking. *International Conference on Technology Enhanced Learning in Higher Education*, (pp. 129-134).
[7] K.V, P. (2010). Trait emotional intelligence theory. *Industrial and organizational psychology*, 136-139.
[8] Khemraj, S. .. (2023). Enhancing Competitive Advantage through Learning Capabilities and Innovative Human Resource Management. *Intersecta Minds Journal, 2*(1), 26-41.
[9] Lanfranchi, G. C. (2026). Reconfiguring competitive advantage: a resource dynamic framework for generative AI adoption in digital content marketing. *European Journal of Innovation Management*, 770-798.
[10] Nikolaou, I. &. (2002). Emotional intelligence in the workplace: Exploring its effects on occupational stress and organizational commitment. *The international Journal of organizational analysis*, 327-342.

[11] O'Boyle Jr, E. H. (2011). The relation between emotional intelligence and job performance: A meta-analysis. *Journal of Organizational Behavior*, 788-818.
[12] Ruttenberg, D. (2026). *Cognitive Debt in AI-Augmented Research: Evidence from Neuroscience and Implications for Knowledge Production*. Retrieved from OSF: osf.io/preprints/psyarxiv/8hbp9_v1
[13] World Economic Forum. (2025). *New Economy Skills: Unlocking the Human Ad-vantage*. Retrieved from World Economic Forum: https://www.weforum.org/publications/new-economy-skills-unlocking-the-human-advantage/

Original Research Article *https://doi.org/10.5281/zenodo.24282722*

LEAN STARTUP METHODOLOGY FOR INNOVATION DEVELOPMENT AND PRODUCTION

Jovana Adžić [0009-0005-7017-7789]
Faculty of Organizational Sciences, Univesity of Belgrade

Abstract: *In modern business conditions characterized by rapid technological development, market globalization, and intense competition, organizations are faced with a constant need for innovation and improvement of production processes. Changes in consumer demands, as well as increasingly shorter product life cycles, force companies to develop new products and services faster, more efficiently, and with minimal risk. Faced with this reality, a large number of entrepreneurs decide to shut down their startups at early stages due to unfavorable market reactions to their solutions, lack of financial resources, or insufficient workforce. Drawing from this experience, Eric Ries chose to turn it into a driving force that would encourage entrepreneurs to continue forward and improve their solutions by applying knowledge and overcoming obstacles through listening to customers. As a response to this problem, the Lean approach emerged, enabling companies to quickly test their ideas without spending large amounts of time and money. This paper discusses the basic concepts of the Lean Startup methodology and the course of its development, as well as the way this approach functions within the innovation ecosystem.*

Keywords: *Lean startup, innovations, Startup ecosystem, Build-measure-learn.*

1. INTRODUCTION

The topic of this paper is the Lean Startup methodology, specifically its development and impact over time. The aim of this paper is to present how products and services can be aligned with the Lean Startup methodology, as well as which principles have been used from its inception to the present in the innovation of ideas and solutions. It will also address the ways in which the Lean Startup approach has been integrated and implemented within companies, and how startups have reached concrete solutions by applying this approach. One of the key topics will be the Build-Measure-Learn process, as well as its contribution to the further development of startups. This topic is important for several reasons, primarily because it enables early-stage startups to understand how they can improve their business, avoid failure in the initial stages, and prevent company shutdown due to lack of knowledge and experience, while also reducing the high costs associated with entering the market.

2. COMPOSITION

The Lean startup approach was developed as a response to numerous business failures of young startups. Lean methodology originally emerged as a model developed within manufacturing, with the aim of continuous improvement and startup development, minimizing the waste of resources and workforce while simultaneously creating value for customers. Over time, this principle began to be applied in the field of service provision as well, where customer feedback became of great importance for the further development of companies. The creator of this methodology is Eric Ries, an innovator who founded several startups himself and consolidated his knowledge and experience through decades of work, both from his own ventures and from observing other startups. The core idea of the Lean startup methodology is based on a cycle of continuous improvement (Olek, 2023). During the development of this business model, Ries attempted to integrate all tools used in practice, as well as those applicable in theory. In searching for answers, he attempted to identify the root causes behind startup failures. By collecting information, he realized that one of the fundamental problems was the waste of resources in the early stages of development, when the solution itself had not yet been tested. He also concluded that many companies wait too long before testing their solution, and therefore proposed a model in which an MVP (Minimum Viable Product) is launched on the market so that, based on user feedback, the solution can be improved without spending large amounts of money or time. For a long time, there was a search for a universally accepted definition of a startup that could precisely explain what constitutes a startup and how it functions. In addition

to raising numerous questions in the market, startups were also not interconnected and did not exchange experiences with one another. The startup ecosystem did not function as a community, meaning that startups were unable to learn from each other or share knowledge. Ries had a mission to break this cycle and create an environment that would enable learning and collective growth. For the first time, startups were confronted with a way of working that required rapid changes and adaptation to market demands. Ries demonstrated that the Lean startup methodology does not have to be expensive or unpredictable, but rather that it can address all obstacles encountered before the final launch of a product or service in a highly economical manner. Startups today represent a powerful mechanism for building knowledge-based businesses, although they continuously face numerous challenges (Mohammadi et al., 2026). The Lean startup approach originated in the field of production after World War II, as a response to increasing consumer demand combined with diverse market offerings (Faria et al., 2021). Considering the period and circumstances in which the Lean approach emerged, this itself serves as proof that it is a methodology with multiple applications and that, regardless of the field of activity, it can be applied across all startups. Another emerging approach involves redefining a startup's action plan in response to various challenges, such as insufficient financial resources, unfavorable user feedback, market conditions, or identified shortcomings in the initial development planning (Pillai et al., 2025). As previously mentioned, the concept of Lean methodology originated after World War II and is rooted in operations management, production, and quality management (Peralta et al., 2025). With the development of the market, various business models emerged within the startup ecosystem, but none proved to be as comprehensive as the Lean startup methodology, which demonstrated itself as a solution applicable to all segments of business operations. Startups continuously evolve in response to environmental demands. Innovation mechanisms change and improve over time, incorporating new technologies and methods while closely monitoring market trends. In this way, startups maintain competitiveness within an innovation ecosystem (Demir et al., 2025). Lean startup principles represent a set of guidelines that direct entrepreneurial activities toward rapid idea testing, continuous learning through iterations, and the development of innovations focused on user needs (Foggetti et al., 2025 according to Silva et al., 2020). Both in product development and in service activities, the system is continuously improved and represents the result of specific processes. The basic idea of this approach is to launch the MVP to the market as quickly as possible in order to improve the solution itself based on user feedback. In this way, the likelihood of investing large amounts of money in the development of a product or service that does not meet the expectations of the target audience is reduced. It is not uncommon for a startup to begin developing a product—especially a service—only to discover that the characteristics of the service do not meet user requirements. By applying the Lean Startup approach, the direction of development can be changed and redirected toward solving users' actual problems. Additionally, with the development of technology, innovations in the field of digitalization also emerge. "With the development of digital entrepreneurship, there is a growing need to integrate digital technologies into scaling frameworks in order to respond to the challenges of market engagement and competition" (Fon-Cot et al., 2025 according to Nambisan, 2017). With the advancement of innovation and the market, user demands have become greater, especially in the service sector. Startups operating in the field of service provision are fully oriented toward the application of Lean methodology, through which they obtain information of great importance for further development at a very early stage. Service provision has always been one of the most challenging ventures. The development of technologies has enabled fast digital solutions which, with the support of the Lean approach and based on feedback and co-creation, deliver rapid results without wasting time, money, or workforce. Over time, increasing attention has been directed toward the customer, and the value system of production can be understood through five key Lean principles: value creation for the customer, identification of the value stream, creation of flow, production based on customer pull, and the pursuit of perfection through continuous identification and elimination of waste (Ghezzi et al., 2020 according to Womack et al., 1997). Whether they are products or services, managing them involves a complex process that requires interaction with the environment, innovation, flexible solutions, and rapid response. Through interaction with the target audience, it was concluded that the obtained information needs to be validated, but the question arises—how? After extensive research, numerous authors have determined that validation alone is not sufficient, instead, it is necessary to follow the entire process: develop a solution, test it, and ultimately learn from the overall process. The process itself provides invaluable data, especially in the innovative sector where ideas are still in the testing phase and startups are uncertain about how users will respond. Since the path of a startup is highly uncertain from the very beginning, through this process of testing and learning from their own mistakes, companies survive by creating solutions that users truly need. The Build–Measure–Learn process represents the core mechanism of the Lean startup methodology. Its fundamental idea is to enable startups to gain experience, test their ideas, and reduce the risk of developing products or services that lack practical value, in a fast, efficient, and cost-effective manner. A startup must be positioned within an appropriate ecosystem to ensure that validation and testing are conducted effectively (Ayamga et al., 2026). When a flexible and iterative way of working is applied, these conditions are not only fulfilled but also enable the maximum utilization of the potential of the existing customer base (Jasemann et al., 2021). The nature of the activity itself, as well as the target audience, determines the operational environment of a startup. The

most important thing is to offer a solution to the right users, since their feedback will be the most relevant for further development. As noted by Mansoori et al. (2019), an interesting finding from research on the impact of feedback on company development is that relationships between entrepreneurs and mentors sometimes deviate from the fundamental principles of Lean startup methodology, which emphasize testing hypotheses through empirical data obtained from interaction with potential customers rather than relying on theory or authority as primary guidelines. When observing organizational units, recent years have seen the adoption of new models, with companies increasingly applying Lean approaches as well as trial-and-error methods. Ries believed that startups fail at the very beginning due to the fear of making mistakes. This fear prevents them from testing solutions early and obtaining feedback, while at the same time such an approach costs them large amounts of money and places the company in a survival state. In service processes, it is essential to analyze assumptions and hypotheses while systematically planning future steps (Gonzalez-Tamayo et al., 2025 according to Zellweger and Zenger, 2023). The application of Lean principles enables production with fewer human resources, less space, fewer financial resources, and fewer materials. To achieve this, Jones and Womack defined five fundamental Lean principles (Masai et al., 2015): value, value stream, flow, pull, perfection. As previously mentioned, feedback from the target audience is of crucial importance. Noted by Laptev et al. (2025), startups continuously conduct iterations of customer discovery and analysis of customer needs. The development process begins with identifying customer needs, conducting market research, and analyzing collected data, followed by the development phase. This process enables continuous growth without risky moves for the company, while also ensuring movement in the right direction. Such a systematic approach is particularly important for innovative service activities. This core mode of operation has evolved over time in accordance with environmental demands. Since the initial application of Lean methodology in 2000, fundamental changes have occurred in the approach. Through its evolution, Lean startup methodology has gradually shifted from continuous testing toward experimentation and innovation, including the creation of innovative business models. Mahdavi et al. (2026) emphasize the importance of implementing an integrated system of Lean production with automation and optimization of existing business processes, which can be applied not only in everyday operations but also in the adoption of new technologies in response to modern construction and business demands. By combining traditional and new business models, continuous progress has been enabled in accordance with innovations and environmental demands. Tang et al. (2023) highlight the importance of business model innovation processes being divided into clearly defined stages and maintaining constant interaction with Lean startup activities. When innovating business models, it is very important to take into account socially relevant issues, that is, the problems society faces. One of these is certainly environmental protection. Lean principles and the concept of zero-defect production in the modern digital environment (Powell et al., 2021). A strategy that has gained importance in recent years is so-called growth hacking, which involves focusing on rapid experimentation and testing of solutions in order to obtain feedback, reduce development time, and minimize resource loss (Foggetti et al., 2025, according to Bargoni et al., 2024). Moiana et al. (2025) refer to the phases through which the Lean approach evolves, emphasizing the search phase, which includes identifying stakeholders and validating them. It is important to emphasize that the Lean startup approach originated from startup practice itself and therefore does not have a strictly defined scientific framework or a single model that others must follow; rather, it is based on the accumulation and refinement of practical experience. Innovative initiatives undertaken by companies are always aligned with their business strategies; therefore, innovation within an ecosystem can be considered heterogeneous across organizations (Adner, 2017). Both in the manufacturing and service sectors, innovation is inevitable in order to keep pace with the rest of the world. When observing innovation in service delivery and business improvement, an effective way of creating value involves reducing resource usage while developing products and services with real utility, thereby contributing to both social and economic value creation (Frare et al., 2026). This approach has very broad applicability and a flexible way of adapting to different requirements; therefore, it saves both time and money while enabling new insights and the acquisition of experience that are crucial for further development. The Lean startup approach has a wide range of applications in the field of innovation. What is important to emphasize is that the Lean startup methodology evolves in accordance with market demands, representing a flexible model that can be applied in any sphere of business. This is one of the main advantages of this approach, especially when considering the speed of technological change. Furthermore, by reducing costs and saving time, this approach enables sustainable business operations while keeping pace with the development of innovative business models, as well as with the validation and optimization of business processes.

Mehrjoo et al. (2026 according to Rana, 2025) highlight the importance of research in integrating agile and Lean approaches, as Lean principles can provide a stable structure for innovation processes in early development stages. In order to establish a stable structure, as noted by Hogeboom et al. (2025 according to Hopkins et al., 2013), the Lean startup approach plays a role in defining the level of risk associated with developing products or services without practical value, while also reducing organizational bias and eliminating subjectivity in order to better identify weaknesses and opportunities for improvement. Lean methodology also faces certain limitations. According to Samson et al. (2026), companies encounter

difficulties in implementing Lean approaches for three main reasons: Insufficient understanding of processes, Resistance of management to major changes, Lack of understanding of a systemic approach to problem-solving. Precisely for these reasons, the fundamental idea of this approach is circular learning, that is, networking startups into communities, since it is believed that the experience startups can exchange is of invaluable importance for successful and long-term business operations. Today, particular attention is given to developing models that address global challenges such as inequality and environmental degradation (Bocken et al., 2020).

3. CONCLUSION

Throughout the entirety of the manuscript, it can be concluded that the Lean Startup methodology has grown and evolved from 2000 to the present day. Many approaches and perspectives have shaped its development, but the core principles of its functioning have remained unchanged. The main contribution of this paper can be observed through the analysis of the comprehensive Lean startup methodology approach and the way it has evolved from its inception to the present day. Today, this approach is focused on adapting to market needs and contemporary business methods. When examining its foundation—namely, the mindset from which innovators operate—this methodology proves to be highly flexible and adaptable, allowing startups to tailor the way it functions to suit their needs. Additionally, the management of production and services has enabled innovators, through interaction with their target audience, to obtain information that is crucial for the development and growth of the company itself. The advantages this approach brings to companies include: fast and easy testing of ideas and solutions, readily accessible user feedback, the ability to pivot in development direction, and savings in both time and money. In the future, this approach can be further improved through the exchange of experiences among startups, as well as by testing new variations within the Lean Startup methodology. Lean startup methodology today occupies one of the leading positions in the field of contemporary innovation development; however, this approach encounters certain limitations when applied across different startups. First and foremost, the scalability of Lean methodology is questionable, considering that solution development methods must continuously adapt to user demands, which means that the initial principles may not necessarily remain applicable in the long term. Furthermore, the success of this approach largely depends on the way a startup responds to customer feedback. If a startup collects feedback from users but fails to implement it properly, the development of the solution may be abruptly interrupted or may continue in the wrong direction. Within this approach, the most important factor is keeping pace with innovation and maintaining the ability to rapidly adapt to environmental and market demands. This process is based on continuous improvement and reassessment of solutions, meaning there is always room for progress and development.

REFERENCES

[1] Adner, R. (2017). *Ecosystem as structure: An actionable construct for strategy*. Journal of Management, 43(1), 39–58. https://doi.org/10.1177/0149206316678451

[2] Ayamga, M., Kassahun, A., & Tekinerdogan, B. (2026). Exploring the role of innovation hubs on agritech startups' innovation and success. Journal of Engineering and Technology Management – JET-M, 79, Article 101942. https://doi.org/10.1016/j.jengtecman.2026.101942

[3] Bargoni, A., Santoro, G., Messeni Petruzzelli, A., & Ferraris, A. (2024). *Growth hacking: A critical review to clarify its meaning and guide its practical application*. Technological Forecasting and Social Change, 200, 123111. https://doi.org/10.1016/j.techfore.2023.123111

[4] Bocken, N. M. P., & Snihur, Y. (2020). *Lean startup and the business model: Experimenting for novelty and impact*. Long Range Planning, 53(4), 101953. https://doi.org/10.1016/j.lrp.2019.101953

[5] Demir, F., & Lukeš, M. (2025). Corporate-startup collaboration: A managerial decision-making framework based on a systematic literature review. Review of Managerial Science, 19(8), 2423–2454. https://doi.org/10.1007/s11846-024-00818-1

[6] Faria, V. F. de, Santos, V. P., & Zaidan, F. H. (2021). *The business model innovation and lean startup process supporting startup sustainability*. Procedia Computer Science, 181, 93–100. https://doi.org/10.1016/j.procs.2021.01.104

[7] Foggetti, C., Natalicchio, A., Ardito, L., & Albino, V. (2025). Opening the black box of growth hacking: Insights into the microfoundations of Lean Startup capabilities. Journal of Business Research, 191, Article 115272. https://doi.org/10.1016/j.jbusres.2025.115272

[8] Font-Cot, F., Lara-Navarra, P., Sánchez Arnau, C., & Serradell-Lopez, E. (2025). *Scaling wheel framework: A holistic approach to startup scalability, governance models, and AI-driven innovation ecosystem competitiveness*. Journal of Open Innovation: Technology, Market, and Complexity, 11(3), 100619. https://doi.org/10.1016/j.joitmc.2025.100619

[9] Frare, A. B., & Beuren, I. M. (2026). Management control systems, bricolage, and frugal innovation: Effects on value creation in startups with a social mission. Knowledge and Process Management, 33(1), Article e70014. https://doi.org/10.1002/kpm.70014
[10] Ghezzi, A., & Cavallo, A. (2020). *Agile business model innovation in digital entrepreneurship: Lean startup approaches*. Journal of Business Research, 110, 519–537. https://doi.org/10.1016/j.jbusres.2018.06.013
[11] Gonzalez-Tamayo, L. A., Maheshwari, G., Krauss-Delorme, C., & Bonomo-Odizzio, A. (2025). Startup decision-making: Integrating trial-and-error with the entrepreneur-as-scientist approach. Journal of Small Business and Enterprise Development, 32(8), 169–196. https://doi.org/10.1108/JSBED-03-2025-0165
[12] Hogeboom, O., Hoveskog, M., Upward, A., Jones, P. H., & Fath-Kolmes, E. (2025). Lean startup supporting sustainability-as-flourishing during the early stages of enterprise development. Journal of Business Models, 13(1), 47–88, Article 1. https://doi.org/10.54337/jobm.v13i1.8222
[13] Jasemann, M., Schlögl, S., & Moser, R. (2021). *Lean startup as a driver for digital innovation: The role of customer feedback and iterative development*. International Journal of Innovation Management, 25(6), 2150058. https://doi.org/10.1142/S1363919621500583
[14] Laptev, G., & Shaytan, D. (2025). *University-based makerspace for entrepreneurial startup teams*. Procedia Computer Science, 263, 900–907. https://doi.org/10.1016/j.procs.2025.07.108
[15] Mahdavi, P., Rahimi, N., Mehranrad, P., Shams, N., Rasool, A., & Ahmad, R. (2026). Improving off-site roof construction through lean principles and automation. Procedia Computer Science, 277, 1423–1432. https://doi.org/10.1016/j.procs.2026.02.181
[16] Mansoori, Y., Karlsson, T., & Lundqvist, M. (2019). *The influence of the lean startup methodology on entrepreneur–coach relationships in the context of a startup accelerator*. Technovation, 84–85, 37–47. https://doi.org/10.1016/j.technovation.2019.03.001
[17] Masai, P., Parrend, P., & Zanni-Merk, C. (2015). *Towards a formal model of the lean enterprise*. Procedia Computer Science, 60, 226–235. https://doi.org/10.1016/j.procs.2015.08.122
[18] Mehrjoo, V., & Shahbazi, L. (2026). Evaluating and prioritizing factors affecting lean management in small and medium-sized enterprises (SMEs) within the digital entrepreneurship sector: A case study. Cogent Business & Management, 13(1), 2635684. https://doi.org/10.1080/23311975.2026.2635684
[19] Mohammadi, N., & Abarghouezade, Z. (2026). Startup mentoring as a catalyst for entrepreneurial success: Insights from a global trend analysis. Social Sciences & Humanities Open, 13, 102638. https://doi.org/10.1016/j.ssaho.2026.102638
[20] Moiana, D., Ghezzi, A., Rangone, A., & Balocco, R. (2025). When lean startup meets complex societal problems: Challenges in startup early-stage experimentation. Proceedings of the European Conference on Innovation and Entrepreneurship (ECIE), 20(1), 1–8. https://doi.org/10.34190/ecie.20.1.4144
[21] Olek, K. (2023). *Startups and lean startup approach in building innovative companies creating unique market values: Theoretical considerations*. Procedia Computer Science, 225, 3745–3753. https://doi.org/10.1016/j.procs.2023.10.370
[22] Peralta, A., & Pyka, A. (2025). Assessing Lean Startup for sustainable business models: Application of the SAFE framework. Environmental Innovation and Societal Transitions, 57, 101029. https://doi.org/10.1016/j.eist.2025.101029
[23] Pillai, S., & Ramakrishnan, R. (2025). Innovation culture: Fostering creativity in the TechStrat fusion era. In V. Jain, N. Gupta, A. K. Agarwal, G. Chetty, & R. Kannan (Eds.), Innovate to integrate: Data-driven management and TechStrat fusion unveiled (Chap. 14). Emerald Publishing Limited. https://doi.org/10.1108/978-1-83708-460-920251014
[24] Powell, D., & Oliveira, M. (2021). *Insights from a digital lean startup: Co-creating digital tools for cognitive augmentation of the worker*. Procedia CIRP, 104, 1046–1051. https://doi.org/10.1016/j.procir.2021.11.176
[25] Samson, D., Ho, W., Seyedghorban, Z., Bevington, D., & Kelly, S. (2026). *Solving the paradox of lean management's low success rate*. Australian Journal of Management. https://doi.org/10.1177/03128962241312725
[26] Tang, X., Du, S., & Deng, W. (2025). *Business innovation in digital startups: A case study of an AI startup*. International Review of Economics & Finance, 98, 103898. https://doi.org/10.1016/j.iref.2025.103898

INTEGRATING STRATEGIC MANAGEMENT AND CORPORATE ENTREPRENEURSHIP IN PUBLIC ENTERPRISES: EVIDENCE FROM SERBIA

Sara Dimitrijević*[1][0009-0000-2278-9952], Dejan Petrović[2][0000-0001-6900-8842]
[1]Teaching Assistant, University of Belgrade, Faculty of Organizational Sciences
[2]PhD, Full Professor, University of Belgrade, Faculty of Organizational Sciences
*Corresponding author, e-mail: sara.dimitrijevic@fon.bg.ac.rs

Abstract: *This paper analyzes innovation activities, digitalization, organizational change and business modernization in public enterprises in the Republic of Serbia as dimensions of corporate entrepreneurship in the period from 2020 to 2025. The research is based on a qualitative approach, combining a review of scientific and professional literature with an analysis of publicly available reports and work programs of three public enterprises: PE "Pošta Srbije", PE "Putevi Srbije" and PE "Srbijagas". The findings indicate that the selected public enterprises demonstrate a strategic orientation toward entrepreneurial and innovation-driven activities, depending on the specific characteristics of each sector. PE "Pošta Srbije" is primarily oriented toward users and the development of digital services. PE "Putevi Srbije" implements entrepreneurial activities mainly through technological and smart infrastructure systems. PE "Srbijagas" emphasizes strategic development, infrastructure expansion and energy system stability. The common trends of modernization and digitalization indicate a gradual shift of public enterprises in the Republic of Serbia toward more efficient, productive and innovation-oriented business models.*

Keywords*: corporate entrepreneurship, public enterprises, public sector, innovation*

1. INTRODUCTION

Changes in business operations and organizational structures have contributed to the growing importance of corporate entrepreneurship in academic research over the past 50 years (Funko et al., 2023; Glinyanova et al., 2021). In addition, various events at the global level have led to changes in policies and modes of governance in the public sector over the past few decades (Grimm & Bock, 2022). According to Verdenhofa (2025:176), “as global practice shows, the fundamental mechanism for optimizing the activities of public authorities is strategic management, which involves the development and introduction of the latest technology to optimize their functioning and contributes to the formation of positive trends in the innovative development of the entire public administration system”. A different approach has been required. Accordingly, the concept of corporate entrepreneurship in the public sector has also developed.

The aim of this paper is to analyse the presence and forms of corporate entrepreneurship in public enterprises in the Republic of Serbia, through the examination of innovation activities, digitalization, business modernization, and organizational changes in the period from 2020 to 2025.

The entrepreneurial approach in the public sector implies the use of innovative methods, proactive problem-solving, process digitalization, and the introduction of strategies that have traditionally belonged to the private sector (Ojiako et al., 2024). In the Republic of Serbia, this concept is still in the development phase, but examples of successful implementation can already be observed, such as the digitalization of public services, the improvement of the operations of state-owned enterprises, and public-private sector partnerships.

2. LITERATURE REVIEW

According to Bayiley and Behaylu (2022), supported by the views of other authors as well (Covin & Miles, 1999; Ireland et al., 2001), contemporary ways of doing business require a flexible and proactive managerial approach that implies the integration of strategic management and corporate entrepreneurship, which enables organizations to recognize their entrepreneurial capabilities and then transform them into competitive advantages.

“Strategic management is defined as the process of evaluation, planning, and implementation designed to maintain or improve competitive advantage” according to Sammut-Bonnici (2015:1791). On the other hand, corporate entrepreneurship is defined as “formal and informal activities aimed at creating new businesses in established companies through product and process innovations and market development” (Zahra, 1991, p.

262). Also, according to the Law on Public Enterprises (2019), a “public enterprise is an enterprise that performs an activity of general interest and is established by the Republic of Serbia, an autonomous province, or a local self-government unit.”
In that context, entrepreneurship in the public sector represents an important mechanism for encouraging innovation and modernizing the work of public institutions (Arundel, 2017; Nicholson-Crotty et al., 2017). Through the application of new ideas, process improvement, and the development of modern solutions, the public sector can increase its efficiency, flexibility, and ability to respond to the growing demands of society.
Public entrepreneurship should be conceptualized through its innovative activities, the transformation of the status quo, and inherent uncertainty (Hayter et al., 2018). For this reason, observing public and economic activity from a sector-specific perspective is of particular importance, as it enables not only a better understanding of the way entrepreneurship manifests itself in the public sector, but also an understanding of the influence that the public sector has on directing and developing entrepreneurial activities (Vivona et al., 2025), which means that the relationship between the public sector and entrepreneurship is bidirectional, with each shaping and reshaping the other (De Massis et al., 2018).

3. METHODOLOGY

This research is based on a qualitative approach, using the method of reviewing professional and scientific literature, with the aim of systematizing and analysing existing knowledge on the relationship between entrepreneurial and innovation activities and the modernization of the operations of the public enterprises Roads of Serbia, Post of Serbia, and Srbijagas.
To identify the key theoretical frameworks, definitions, and concepts of entrepreneurship and innovation in the public sector, a structured qualitative review of professional and scientific literature was applied, using academic databases such as Google Scholar and Web of Science, as well as books, reports, and annual work programs of public enterprises. In addition, the descriptive method was used through the tabular presentation of data.
Within the analysis, a comparative thematic analysis was applied, which includes the identification and comparison of key concepts of entrepreneurship and innovation, as well as approaches in the different activities performed by public enterprises.
The synthesis of findings, based on the collected data, was used in the research with the aim of identifying patterns, similarities, and differences in approaches, as well as formulating a conceptual framework that connects the analysed concepts.

4. RESULTS

Table 1 presents selected innovation and entrepreneurial activities of the analysed public enterprises in the Republic of Serbia in the period from 2020 to 2025.

Table 1: Selected innovation and entrepreneurial activities of public enterprises in the Republic of Serbia (2020–2025)

PE “Putevi Srbije” (2020–2025)	PE “Pošta Srbije” (2020–2025)	PE “Srbijagas” (2020–2025)
Full digitalization of toll collection (ETC/TAG system)	PostexpressBOX	Connection to the gas distribution network
Expansion of electronic toll collection (TAG devices)	Implementation of e-services	Introduction of household gas connections
Electronic toll collection (ETC)	Online and mobile payments of PostNet and international money orders	Construction of distribution gas pipelines in several regions
Installation of smart traffic signs	User identification and receipt of requests related to bank products	Gasification of certain municipalities
Introduction of the ISO 50001 standard	Introduction of OTT media content distribution	Noise measurement at above-ground facilities (legal requirement)
Development of the communication network topology (optical and wireless)	Establishment of the e-Notar information system for secure archiving of notarial documentation	Introduction of renewable energy sources into operation
Implementation of video surveillance systems at road weather stations	Development of new postal services in domestic and international postal traffic for shipment standardization purposes	Introduction of ESG procedures
Upgrading the meteorological data collection system	Packaging standardization and development of new postal services	Restructuring and organizational changes

PE “Putevi Srbije” (2020–2025)	PE “Pošta Srbije” (2020–2025)	PE “Srbijagas” (2020–2025)
Development of GIS applications and digitalization of work processes	Development of e-postal services	Establishment of new subsidiaries / business expansion
Implementation of monitoring systems	Introduction of e/m Telegram services	Modernization of equipment, gas infrastructure, and systems
Automatic traffic counting	WEB/m money orders – payment via the internet and mobile application	Integration into international energy systems
B-WIM axle load measurement	Banking-related services: user identification and receipt of requests for bank products	
Geospatial databases	Development of ICT projects	
Integration of data from multiple sources (Ministry of the Interior, Road Traffic Safety Agency)	Improvement of the receipt and delivery process for Post Express shipments and parcels through automation and enhanced application functionality	
Cybersecurity and modernization (penetration testing, firewall filters, IT data protection, disaster recovery)	Development of the Post Express receipt process through WEB services and integration with users’ software solutions	
	Improvement of the process of receiving and delivering Post Express	

5. DISCUSSION

Through the analysis of the innovation and entrepreneurial activities of the public enterprises PE "Putevi Srbije", PE "Pošta Srbije" and PE "Srbijagas" in the period from 2020 to 2025, presented in Table 1, significant differences in their approach to modernization can be identified. At the same time, trends in digitalization and business improvement were observed as important characteristics common to all three public enterprises.

PE "Pošta Srbije" had its focus on introducing innovations through the digitalization of services that were important for users. The listed activities, presented in Table 1, illustrate the change in the postal industry from traditional ways of providing services to modern and digitalized approaches to services that are important for users. PE "Putevi Srbije" upgraded the transport infrastructure sector with innovative technological solutions. Electronic toll collection and the introduction of GIS systems speak of a high level of innovation and a move toward “smart” roads, which implies greater efficiency in infrastructure management and traffic safety. PE "Srbijagas" has implemented its entrepreneurial activities through the introduction of innovations both in the direction of strategic activities and, also, infrastructure activities. The example of gasification of municipalities reflects infrastructure orientations, while the example of the introduction of ESG procedures or international integration with the energy systems of other countries demonstrates a long-term strategic orientation.

Over the past five years, the listed public enterprises have clearly focused their strategic directions through entrepreneurial activities and innovations. Accordingly, development goals and priorities have moved toward business modernization, digitalization, and the development of new services, which are part of a systematic and planned strategic approach. The common trends in entrepreneurial activities that accompany all three enterprises relate to business modernization, digitalization, and increased market orientation, even though they belong to different sectors of the economy. The analysis of the results supports previous research (Arundel, 2017; Catala et al., 2023; Ključnikov et al., 2025), which link entrepreneurship, innovation, digitalization, and organizational change.

6. CONCLUSION

Based on the discussion of the research findings, it is possible to conclude that all three enterprises are developing in accordance with Industry 4.0, i.e., they show an orientation toward the integration of business operations and new technologies, with differences in their approaches depending on the sectors in which they operate. PE "Putevi Srbije" prioritizes technology and "smart" infrastructure systems, while PE "Pošta Srbije" focuses on digital services and user experience. PE "Srbijagas" prioritizes strategic development, energy security, and infrastructure expansion.

Based on the research, it is concluded that, in the period from 2020 to 2025, the operations of all three public enterprises were improved through various entrepreneurial activities and innovations.
The results show a clear direction towards strategic development, which is based on digitalization and other innovations, which represents a transformation compared to the traditional way of doing business. This transition indicates a development toward more efficient, modern and market-oriented approaches to business operations, which implies the improvement of the public sector and the economy.
Corporate entrepreneurship contributes to the modernization of public enterprises, improves organizational efficiency and effectiveness and supports digitalization. The results also indicate that there are conditions for further research into market-oriented approaches and the application of innovative solutions in the public sector.
The limitations of the paper relate to the number of companies analysed, as well as the fact that the analysis is based on publicly available documents, as secondary data sources. Future research could include a larger number of companies and the use of other research methods, such as interviews, surveys and other quantitative approaches.

REFERENCES

[1] Arundel, A. (2017). Rethinking the effect of risk aversion on the benefits of service innovations in public administration agencies. *Research Policy*, 46(5), 900–910. https://doi.org/10.1016/j.respol.2017.03.009

[2] Bayiley, Y. T., & Hailegiyorgis Behaylu, R. (2022). Linking strategic management and corporate entrepreneurship for firm value creation: A developing country perspective. *Journal of African Business*, *23*(1), 79–103. https://doi.org/10.1080/15228916.2020.1812979

[3] Catala, B., Savall, T., & Chaves-Avila, R. (2023). From entrepreneurial and innovation ecosystems to the social economy ecosystem. *Journal of Business Research*, 163, 113932. 10.1016/j.jbusres.2023.113932

[4] Covin, J. G., & Miles, M. P. (1999). Corporate entrepreneurship and the pursuit of competitive advantage. *Entrepreneurship Theory and Practice*, 23(3), 47–63. https://doi.org/10.1177/10422587990230

[5] De Massis, A., Kotlar, J., Wright, M., & Kellermanns, F. W. (2018). Sector-based entrepreneurial capabilities and the promise of sector studies in entrepreneurship. *Entrepreneurship Theory and Practice*, 42(1), 3–23. https://doi.org/10.1177/1042258717740548

[6] Glinyanova, M., Bouncken, R. B., Tiberius, V., & Cuenca Ballester, A. C. (2021). Five decades of corporate entrepreneurship research: Measuring and mapping the field. *International Entrepreneurship and Management Journal*, 17(4), 1731–1757. https://doi.org/10.1007/s11365-020-00711-9

[7] Grimm, H. M., & Bock, C. L. (2022). Entrepreneurship in public administration and public policy programs in Germany and the United States. *Teaching Public Administration*, *40*(3), 322–353. https://doi.org/10.1177/01447394211021636

[8] Funko, I. S., Vlačić, B., & Dabić, M. (2023). Corporate entrepreneurship in public sector: A systematic literature review and research agenda. *Journal of Innovation and Knowledge*, *8*(2), Article 100343. https://doi.org/10.1016/j.jik.2023.100343

[9] Ireland, R. D., Hitt, M. A., Camp, S. M., & Sexton, D. L. (2001). Integrating entrepreneurship and strategic management actions to create firm wealth. *Academy of Management Perspectives*, 15(1), 49–63. https://doi.org/10.5465/ame.2001.4251393

[10] Ključnikov, A., Civelek, M., Kupec, V., & Badie, N. B. (2025). Breaking down entrepreneurial barriers: exploring the nexus of entrepreneurial behavior, innovation, and bank credit access through the lens of the theory of planned behavior. I*nternational Entrepreneurship and Management Journal*, 21(1), 13. DOI: 10.1007/s11365-024-01025-w

[11] Nicholson-Crotty, S., Nicholson-Crotty, J., & Fernandez, S. (2017). Performance and management in the public sector: Testing a model of relative risk aversion. *Public Administration Review*, 77(4), 603–614. https://doi.org/10.1111/puar.12619

[12] Ojiako, U., AlRaeesi, E. J. H., Chipulu, M., Marshall, A., & Bashir, H. (2024). Innovation readiness in public sector service delivery: An exploration. *Production Planning & Control,* 35(5), 437–460. https://doi.org/10.1080/09537287.2022.2089266

[13] Sammut-Bonnici, T. (2015). Strategic management. In C. L. Cooper, M. Vodosek, D. N. Den Hartog, & J. M. McNett (Eds.), *Wiley Encyclopedia of Management*. Wiley. https://doi.org/10.1002/9781118785317.weom060194

[14] Vivona, R., Clausen, T. H., Gullmark, P., Cinar, E., & Demircioglu, M. A. (2025). Public sector entrepreneurship: An integrative review. *Small Business Economics*, *64*(4), 1791–1815. https://doi.org/10.1007/s11187-024-00965-5

[15] Verdenhofa, O., Dzenis, O., Dehtjare, J., Djakona, V., & Gavkalova, N. (2024). Strategic management as a tool for the development of public sphere entities. *Public Policy and Administration*, 23(2), 175–196. https://doi.org/10.13165/VPA-24-23-2-05

Original Research Article

FROM QUARTERLY ACHIEVEMENT TO RENEWED ENGAGEMENT: BEHAVIORAL AND MOTIVATIONAL PATTERNS DURING STRATEGY IMPLEMENTATION IN A LARGE PUBLIC UTILITY

Goran Ličanin [1][0009-0009-3043-3468], Mladen Čudanov[1][0000-0001-6895-6525], Zorica Mitrović[1][0000-0001-9185-2652]

[1]University of Belgrade, Faculty of Organizational Sciences, Serbia
*Goran Licanin, e-mail:gl20255005@student.fon.bg.ac.rs

Abstract: *This exploratory interpretive paper examines a recurring behavioral pattern in a strategy implementation project at a large public utility, focusing on quarterly execution cycles. It explores why success in one cycle may not automatically translate into motivation for the next and how team engagement can fluctuate around temporal boundaries. Based on a longitudinal embedded case study with action research elements, the paper identifies challenges that arise during transitions between cycles. The findings suggest that the end of a cycle may temporarily exhaust alignment and the short-term personal logic of success of influential leaders, weakening momentum into the next cycle. Re-engagement is interpreted as occurring through renewed mission focus, psychological ownership, visibility to senior management, and accountability. The paper argues that strategy implementation requires not only planning and reporting, but also deliberate renewal of engagement at quarterly boundaries.*

Keywords*: Strategy implementation, quarterly cycles, motivational reframing, personal logic of success, influential leaders, goal commitment, public sector*

1. INTRODUCTION

Application of strategy is often structured in time-bound cycles that translate broad strategic priorities into concrete goals, expected results, control points, and defined responsibilities. Periods can be adjusted to days, weeks, or months, but quarterly segmentation is widespread in practice. This exploratory study examines a recurring behavioral pattern that occurs immediately after a quarterly segment is closed, focusing on continuity and motivational challenges across segments. Most practical research focuses on planning, monitoring, success indicators, and formal mechanisms of project accountability. Less attention is paid to the possibility that formal success in the previous quarter may be associated with a temporary weakening of engagement, especially among actors whose behavior has broader secondary consequences for the team. The paper, therefore, identifies and interprets a plausible mechanism of execution discontinuity rather than testing a general explanatory model.

The foundation of this paper emerged from a strategy implementation project in a large public enterprise, in which quarterly cycles served as the basic framework for translating strategic priorities into operational goals, facilitating cross-functional coordination, and enabling milestone-based reporting. Repeated observations across several successive quarterly transitions indicated that an even decline did not follow completion of one segment in engagement across the whole team, but rather by visible changes in the behavior of leaders perceived in practice as having significant influence on the flow and rhythm of work.

2. THEORETICAL FRAMEWORK AND LITERATURE REVIEW

In the context of this research, quarterly project cycles served not only as an administrative framework but also as a method of temporal structuring that influenced how individuals involved in the process perceived progress,

completion, and the transition to the next phase. This perspective is based on research regarding time-paced evolution, punctuated transitions, and temporal structuring within organizations. (Brown & Eisenhardt, 1997; Gersick, 1988; Orlikowski & Yates, 2002).
As each quarter was correlated with a specific goal or set of goals, it can be related to goal-setting theory, which shows that specific and challenging goals mobilize attention, effort, and persistence (Locke & Latham, 1990, 2002). Quarterly defined goals can therefore serve as strong mechanisms for focusing behavior. However, literature on goal progress and goal revision is crucial here. Perceived progress and experiences of success can change later goal orientation, reduce immediate investment in the same action pattern, and alter whether the next goal feels near or distant (Fishbach & Dhar, 2005; Huang, Zhang, & Broniarczyk, 2012). Therefore, quarterly achievement is treated not only as planning-period completion, but as a transition point where motivation, ownership, and execution intensity must be renewed for the next sub-goal within the same strategic trajectory.

A quarterly goal in strategy implementation projects may be understood not only as a formal organizational requirement but also as a temporary point of alignment between organizational priorities and personally relevant outcomes for key actors. Goal congruence and person–organization fit explain alignment between individual and organizational goals (Vancouver & Schmitt, 1991; Kristof, 1996). Self-concordance explains why goals consistent with internalized values are pursued more persistently (Sheldon & Elliot, 1999), while goal commitment explains attachment to specified goals (Klein, Wesson, Hollenbeck & Alge, 1999). In contrast, personal logic of success is used here as a narrower interpretive construct: the cluster of outcomes that makes a specific quarterly goal personally worth investing in, such as meaning, ownership, reputation, visibility, confirmation of influence, or career advancement. It is therefore not equivalent to commitment, self-concordance, or psychological ownership, but rather connects these lenses to the quarterly outcome that is experienced as valuable.

Self-determination theory distinguishes more autonomous forms of motivation, based on identification with the meaning of the goal, from more controlled forms of motivation, based on external pressure, rewards, accountability, or the avoidance of negative consequences (Deci & Ryan, 2000; Ryan & Deci, 2000). For this paper, what matters is that both types of motivational regulation may produce renewed engagement, but not necessarily with the same quality or sustainability. This opens space for understanding reactivation as a process that may be triggered by either meaning-and-ownership or more instrumental, control-based pathways.

The transition between quarters is not just a logistical process, but also involves interpretation. Organizational members do not simply react to events. They respond to the meanings they assign to those events (Weick, 1995). In times of change, certain individuals play a crucial role in shaping how others perceive the situation (Gioia & Chittipeddi, 1991). Therefore, the upcoming quarter needs to be actively redefined as relevant and deserving of attention. Once the previous quarter is symbolically concluded, the next phase cannot automatically carry the same level of engagement without intentional interpretive efforts. Psychological ownership refers to the personal feeling that a particular project, task, or area of work belongs to an individual or a group (Pierce, Kostova, & Dirks, 2001; Pierce & Jussila, 2011). The term execution multipliers is used here as a descriptive analytical label for leaders whose ownership, visibility, coordination position, and informal influence make their engagement consequential for the pace of the wider team. It does not describe a fixed personality type, but an observed role in this project context.

The argument that quarterly boundaries matter is grounded in literature on time, temporality, and temporal structuring in organizations and projects. Work groups often shift around salient temporal markers and deadlines (Gersick, 1988). Actors also produce and reproduce temporal rhythms that influence action (Orlikowski & Yates, 2002), whereas entrainment emphasizes pace, cycles, and synchronization (Ancona & Chong, 1996). Project research similarly treats temporal structures as active elements of how projects are understood and performed (Winch & Sergeeva, 2022). The quarterly boundary is therefore interpreted here as a psychological and organizational event rather than a neutral calendar division. This explains why structures that improve manageability may also require deliberate renewal of engagement. Similar problems are observed in education, where motivation can be sustained through continuous, yet varied, communication (Hegyi, 2023).

Research questions
The exploratory research questions derived from theory and empirical qualitative insights are:

1. Does completing a quarterly segment in a strategy implementation project appear to be associated with a temporary decrease in execution intensity?
2. Is this decline manifested evenly across the whole team, especially among leaders with perceived influence?
3. Through which mechanisms are these leaders re-engaged for the next quarterly cycle?

4. How do various reactivation pathways impact the restoration of execution continuity?
5. How may this behavioral pattern relate to continuity in temporally segmented strategy implementation projects, particularly in the public sector?

3. METHODS AND DATA

This paper presents a longitudinal embedded case study with elements of action research, focusing on a dynamic observed over time within a large public utility enterprise. Data were collected during a project led by one of the authors, involving four middle management members (B-2 level) and two heads of function (B-1 level), as part of a broader strategy implementation effort. The purpose is interpretive pattern identification, not formal explanatory model testing.

The first author's role as lead implementation consultant enabled access to behavioral changes that might not have been visible retrospectively. To reduce interpretive bias, the analysis used concrete indicators, repetition across four quarterly episodes, project-documentation checks, and discussion within the three-consultant implementation team. Alternative explanations, such as workload, role change, and uncertainty, were considered before interpreting changes as declining or renewed engagement.

Data sources included project notes, meeting dynamics, communication patterns, project plans, and indicators such as postponed meetings, reduced initiative, slower responses, missed meetings, and changes in task engagement. Analysis moved from indicators to cross-quarter comparison and then to interpretation through goal commitment, sensemaking, and psychological ownership. Interpretations were treated as plausible and validated through cross-episode consistency, evidence triangulation, and agreement among the consulting team.

4. RESULTS AND DISCUSSION

Table 1 summarizes recurring indicators and interpretations. Reduced responsiveness is read through goal commitment; reframing of new priorities through sensemaking and sensegiving; and renewed ownership, visibility, and accountability through psychological ownership. This supports an exploratory cross-quarter interpretation rather than a causal test.

Table 1 Behavioral and Reactivation Patterns Across Quarterly Episodes

Episode	Dominant context	Leader with perceived influence	How declining engagement was manifested	Reactivation path	Outcome
Quarter 1	New organizational structure and initial consolidation	A leader whose first wave of consolidation projects was successful and whose status was reinforced by a new organizational position	Postponed meetings, distance in communication, slower replies, and contesting of new projects	Visibility/reputation risk; sector remains in implementation	Partial stabilization and continuation of work
Quarter 2	Stabilization of the new structure and arrival of the new head of the function	The same leader with a large operational scope and field centralization responsibilities	Lower proactivity, missed meetings, overload, and the need for additional pressure	Formal support from the new leadership and justification of the trust placed in the leader	Renewed continuity, but with lingering distance
Quarter 3	First stabilization of the function and a new cross-functional proof of concept	A second leader who became the focal actor because of visible success	Passivity toward scaling the success, slower communication, and postponement of the next initiative	New project focus and strategic positioning	Reactivation through new ambition and a new political-organizational logic
Quarter 4	Replacement of the head of the function and a shift in power relations	Leaders in transition to new roles, including the new executive director of the function	Decline in engagement due to role changes, uncertainty, and the exit of the consultant	Established processes and attachment to earlier successes	Continuity preserved through a new distribution of roles

5. CONCLUSIONS, LIMITATIONS, AND FUTURE RESEARCH

Qualitative findings suggest that maintaining momentum in quarterly strategy implementation is not ensured by formal planning and monitoring alone. In this case, each successful segment was followed by a transitional phase in which execution continuity had to be renewed. The contribution should therefore be read as an empirically grounded interpretive insight from one embedded longitudinal case, not as a broadly established explanatory model.

The findings are consistent with prior research on goal progress, which shows that perceived achievement can change subsequent goal orientation, and with goal commitment theory, because renewed effort depended on whether the next quarter again became personally worth pursuing. They also extend sensemaking research by showing how leaders had to reinterpret the next cycle as relevant, and psychological ownership research by indicating why some leaders had a multiplier effect on team continuity.

The main exploratory finding is that completing a quarterly cycle was repeatedly associated with a temporary decline in engagement, especially among leaders whose personal logic of success had been satisfied by the previous quarter. Reactivation appeared to occur when a new valued-outcome logic was created through mission focus, visibility, accountability, or ownership. These are interpreted as plausible reactivation pathways, not definitive causal drivers.

The study has limited generalizability because it is based on a single public utility context and the first author's embedded role. As an exploratory study, these results do not call for determining future regulations or strong practice guidelines, but rather for setting the direction for future research. Bias was mitigated through concrete indicators, cross-quarter comparison, documentation checks, consultant-team validation, and consideration of alternative explanations, but it cannot be eliminated. Future research should test the proposed pattern quantitatively, compare cases across sectors and organizational sizes, and develop measures of personal logic of success and execution multiplier effects.

REFERENCES

[1] Ancona, D., & Chong, C.-L. (1996). Entrainment: Pace, cycle, and rhythm in organizational behavior. In Research in organizational behavior (Vol. 18, pp. 251–284). JAI Press.

[2] Brown, S. L., & Eisenhardt, K. M. (1997). The art of continuous change: Linking complexity theory and time-paced evolution in relentlessly shifting organizations. Administrative Science Quarterly, 42(1), 1–34.

[3] Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. Psychological Inquiry, 11(4), 227–268.

[4] Fishbach, A., & Dhar, R. (2005). Goals as excuses or guides: The liberating effect of perceived goal progress on choice. Journal of Consumer Research, 32(3), 370–377.

[5] Gersick, C. J. G. (1988). Time and transition in work teams: Toward a new model of group development. Academy of Management Journal, 31(1), 9–41.

[6] Gioia, D. A., & Chittipeddi, K. (1991). Sensemaking and sensegiving in strategic change initiation. Strategic Management Journal, 12(6), 433–448.

[7] Hegyi, B. (2023). Experimental Conclusions of the Online Inter-University Creativity, Responsibility, and Entrepreneurship Course Implementation. Education Sciences, 13(3), 240. DOI:10.3390/educsci13030240.

[8] Huang, S.-C., Zhang, Y., & Broniarczyk, S. M. (2012). So near and yet so far: The mental representation of goal progress. Journal of Personality and Social Psychology, 103(2), 225–241.

[9] Kristof, A. L. (1996). Person–organization fit: An integrative review of its conceptualizations, measurement, and implications. Personnel Psychology, 49(1), 1–49.

[10] Klein, H. J., Wesson, M. J., Hollenbeck, J. R., & Alge, B. J. (1999). Goal commitment and the goal-setting process: Conceptual clarification and empirical synthesis. Journal of Applied Psychology, 84(6), 885–896.

[11] Locke, E. A., & Latham, G. P. (1990). A theory of goal setting and task performance. Prentice Hall.

[12] Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. American Psychologist, 57(9), 705–717.

[13] Orlikowski, W. J., & Yates, J. (2002). It’s about time: Temporal structuring in organizations. Organization Science, 13(6), 684–700.

[14] Pierce, J. L., Kostova, T., & Dirks, K. T. (2001). Toward a theory of psychological ownership in organizations. Academy of Management Review, 26(2), 298–310.

[15] Pierce, J. L., & Jussila, I. (2011). Psychological ownership and the organizational context: Theory, research evidence, and application. Edward Elgar Publishing. DOI: 10.4337/9780857934451

[16] Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation. social development, and well-being. American Psychologist, 55(1), 68–78.

[17] Sheldon, K. M., & Elliot, A. J. (1999). Goal striving, need satisfaction, and longitudinal well-being: The self-concordance model. Journal of Personality and Social Psychology, 76(3), 482–497.

[18] Vancouver, J. B., & Schmitt, N. W. (1991). An exploratory examination of person–organization fit: Organizational goal congruence. Personnel Psychology, 44(2), 333–352.

[19] Weick, K. E. (1995). Sensemaking in organizations. Sage.

[20] Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (2005). Organizing and the process of sensemaking. Organization Science, 16(4), 409–421.

[21] Winch, G. M., & Sergeeva, N. (2022). Temporal structuring in project organizing: A narrative perspective. International Journal of Project Management, 40(1), 40–51.

Original Research Article *https://doi.org/10.5281/zenodo.23398811*

STUDENTS' ENVIRONMENTAL AWARENESS: LINKING POLLUTANTS, EDUCATION, SCIENCE, AND ISO 14001 STANDARD

Dušan V. Trajković*[1,2][0009-0007-7830-0192], Filip Gaus[3][0009-0008-7272-7395], Teodora Križanić[2][0009-0007-0893-3154], Jelena Ruso[2][0000-0003-4120-8759]

[1]Innovation Centre of the Faculty of Technology and Metallurgy, Belgrade, Serbia
[2]Faculty of Organizational Sciences, University of Belgrade, Belgrade, Serbia
[3]Faculty of Technology and Metallurgy, University of Belgrade, Belgrade, Serbia

*Corresponding author, e-mail: dtrajkovic@tmf.bg.ac.rs

Abstract: *Environmental pollution is one of the major challenges of today. As future professionals and decision-makers, students' environmental knowledge, habits, and perceptions are highly important for society. This study aimed to examine students' perception of polluting materials in the environment, the role of education and science in solving environmental problems, and the importance of environmental management systems as defined by ISO 14001 standard. A survey was conducted among active higher education students in the Republic of Serbia from different genders, study levels, and fields of study. Relatively high values were obtained for all examined constructs. In addition, respondents were grouped using non-hierarchical k-means cluster analysis. Three clusters were identified, with the smallest cluster including 5.7% of respondents and characterized by low perception of polluting materials, low trust in science and education, and low confidence in environmental management systems. The results indicate positive student attitudes, but also highlight the need to further strengthen education, science, and environmental management practices.*

Keywords*: emerging pollutants; science; students; environmental education; ISO 14001.*

1. INTRODUCTION

Constant industrial development and rising living standards have inevitably led to an increasing number of environmental problems. At the same time, issues related to environmental protection are not limited to specific geographic areas or regions, but rather represent a global challenge affecting even the most developed countries in the world (Trajković & Glogovac, 2025). Improper waste disposal, the use of pesticides and fertilizers in agricultural activities, greenhouse gas emissions, and the use of hazardous chemicals are just some of the everyday activities that can cause serious environmental consequences. The development of analytical methods and techniques has led to the emergence of new polluting materials such as pharmaceutical residues and hormones, UV filters, and other emerging contaminants. Their monitoring in environmental samples has become one of the leading challenges.(Khan et al., 2022).

Due to the complexity of the modern world and numerous environmental challenges, people are often required to make decisions based on their attitudes and knowledge in order to address various problems (Righi et al., 2024). As future professionals and decision-makers, students' attitudes and awareness regarding environmental issues represent one of the key factors that may influence future environmental protection policies and changes in individual behavior. In this context, preventive action and effective solutions to environmental problems do not depend solely on legal regulations, but are also strongly influenced by education and individual awareness. More precisely, environmental education can be considered one of the most important tools for sustainable development (Dagiliūtė & Liobikienė, 2015). On the other hand, scientific advancements in the field of environmental pollution are developing in two directions: (1) the development of adequate analytical methods for detecting trace levels of emerging pollutants (Bujagić et al., 2020), and (2) the development of appropriate methods for their removal from the environment (Trajković et al., 2025; Vukčević et al., 2024). In this regard, students' trust in science and scientific achievements represents an important prerequisite for their future active involvement in scientific processes and the application of science-based solutions in environmental protection.

In addition to education and scientific advancements, standardized environmental management systems also play a significant role in pollution prevention and control. The ISO 14000 series addresses environmental protection issues, while ISO 14001:2015, Environmental Management Systems – Requirements with Guidance for Use, is one of the most important frameworks for implementing an environmental management

system (Loste et al., 2019). This standard defines the requirements for environmental management systems in organizations, with the aim of identifying processes and activities that may affect the environment, preventing pollution, and ensuring the continual improvement of environmental performance (Arocena et al., 2023). Understanding the ISO 14001 standard provides students with knowledge about institutional responsibility and the ways in which environmental policies are translated into concrete results. However, when students lack awareness of environmental management systems, their ability to participate in these processes and initiate useful environmental initiatives is reduced.

2. METHODOLOGY AND MEASUREMENT

The main objective of this study was to assess the level of environmental awareness among students in the Republic of Serbia by examining their perceptions of polluting materials in the environment, trust in education and science, and familiarity with the standards of the International Organization for Standardization (ISO). The research was conducted through an online survey (Microsoft Forms), administered in Serbian, during the period from March to April 2026. A total of 88 respondents participated in the study, differing in gender, level of study, and field of study. The questionnaire was distributed through electronic channels (including official and informal student group pages on social networks). Participation was voluntary and anonymous. The structure of the respondents is presented in Table 1.

Table 1: Description of the student sample (N=88)

Attributes	Category	Frequency	Percentage (%)
Gender	Male	37	42.0
	Female	51	58.0
Level of study	Bachelor's studies	43	48.9
	Master's studies	19	21.6
	Doctoral studies	26	29.5
Field of study	Natural Sciences and Mathematics	17	19.3
	Engineering and Technology	49	55.7
	Medical Sciences	7	8.0
	Social Sciences and Humanities	15	17.0

The questionnaire was structured to include introductory questions on the respondents' socio-demographic characteristics and three constructs related to Perception of Polluting Materials (PPM), Environmental Education and Science (EES), and ISO 14001 standards (ISO). Respondents indicated their level of agreement with the given statements using a five-point Likert scale, where 1 indicated "strongly disagree" and 5 indicated "strongly agree." The PPM and EES constructs consisted of 5 items, while ISO consisted of 3 items, and composite variables were formed by calculating the mean value of the items. The internal consistency of the examined constructs was assessed using Cronbach's alpha coefficient, with all constructs rated as acceptable (> 0.7). The questionnaire structure and Cronbach's alpha values are presented in Table 2.

Table 2: Description of the measured constructs and reliability estimates

Construct	Number of items	Measurement scale	Cronbach's α
Perception of polluting materials (PPM)	5	5-point Likert scale	0.88
Environmental education and science (EES)	5	5-point Likert scale	0.76
ISO 14001 standards (ISO)	3	5-point Likert scale	0.74

The collected data were analyzed using IBM SPSS Statistics, version 27. Statistical analysis included descriptive statistics and non-hierarchical cluster analysis. K-means cluster analysis was performed using three composite variables: PPM, EES and ISO. Clustering was performed from 3 to 5 clusters, with optimal results obtained when the number of clusters was three, thus identifying different student profiles according to their level of environmental awareness. This method allowed grouping respondents based on the similarity of their responses. The interpretation of the resulting clusters was based on the mean values of the composite variables within each cluster.

3. RESULTS

The obtained results indicate a relatively high level of student awareness regarding the perception of polluting materials, Environmental Education and Science, and the importance of the ISO 14001 standard. Descriptive statistics of the examined constructs are presented in Table 3. The construct Perception of polluting materials

(PPM) showed the highest value (4.80 ± 0.68), which indicates that students adequately recognize the presence of emerging pollutants in the environment. The low standard deviation value indicates small differences among respondents. The construct EES also had a relatively high value (4.00 ± 0.77), suggesting that students recognize the important role of education and science in solving environmental problems. The lowest mean value was recorded for the ISO 14001 construct (3.75 ± 0.77), indicating somewhat lower awareness among students about the role of international standards and their importance. However, the fact that the mean value is in the upper half of the scale indicates a moderately positive attitude toward this issue.

Table 3: Description of the measured constructions and reliability estimates

Construct	Median	Mean	Standard Deviation
Perception of polluting materials (PPM)	4.57	4.80	0.68
Environmental education and science (EES)	3.90	4.00	0.77
ISO 14001 standards (ISO)	3.67	3.75	0.77

The non-hierarchical cluster analysis grouped the respondents into three clusters. The mean values of the composite variables (PPM, EES, and ISO) are shown in Figure 1.

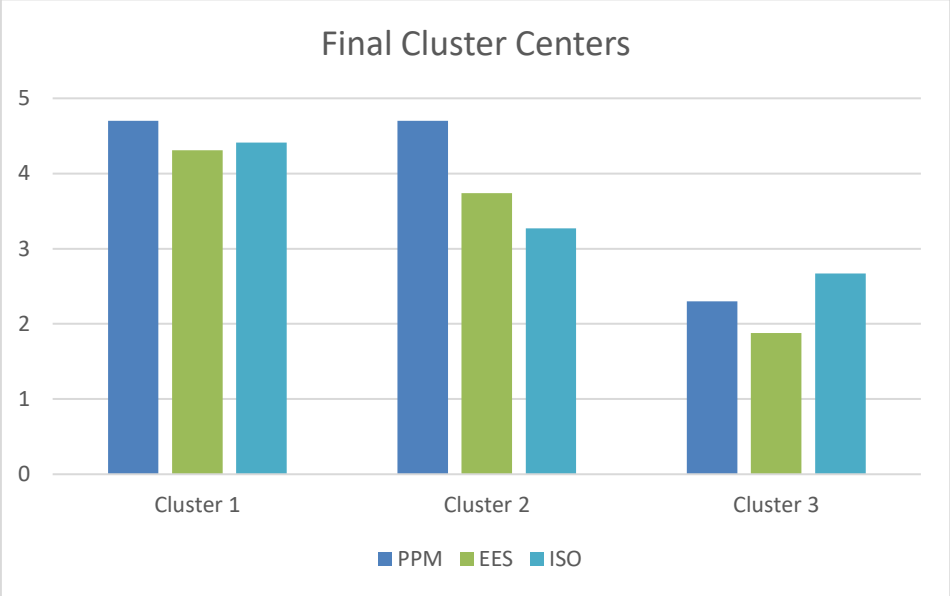


Figure 1: Mean construct values across the identified clusters

The first cluster consisted of respondents who had high values of all examined composite variants (PPM = 4.7, EES = 4.31 and ISO = 4.41). Almost half of the respondents (n = 40, 45.4%) are in this cluster. The second cluster contains 43 respondents, or 48.9%. This cluster is characterized by slightly lower values of the composite variants PPM, EES and ISO, and one is 4.7, 3.74 and 3.27, respectively. This cluster contains a respondent with slightly lower values of the EES and ISO parameters, while the PPM parameter is identical to that of the group in the first cluster. In other words, if this group has a good perception of polluting materials, it is characterized by a slightly less positive attitude towards the importance of ISO standards, as well as lower trust in science and education. The third cluster contained only 5 respondents, or 5.7%. This cluster represents a group with a low level of perception of polluting materials (2.3), as well as low trust in ISO standards (2.67), science and education (1.88).

4. DISCUSSION

The obtained results show that students in Serbia have a highly developed awareness and positive attitudes regarding environmental protection, which is reflected in the high values of all three constructs. The high value of the perception of pollutants indicates a starting point in the development of pro-environmental behavior, activating an individual's sense of responsibility (Song et al., 2019). Chan et al. (2014) found that environmental knowledge can be a positive predictor of the development of pro-environmental behavior; on the other hand, Vainio & Paloniemi (2014) showed in their research that excessively high trust in science can lead to the transfer of responsibility from the individual level to the level of scientific institutions. The lowest value of the construct was recorded for the ISO construct, indicating that only a portion of students had the opportunity to

study the environmental management system during their formal education. As students people are the carriers of future social changes and one of the most important actors in the development of society, such results are encouraging. However, although this type of research has yielded positive results, the further development of educational programs, stronger promotion of science in this field, and the promotion of environmentally responsible behavior can further improve the current environmental situation in Serbia.

5. CONCLUSION

High level of awareness of the presence of polluting materials in the environment, as well as positive attitudes toward education, science and standardization in this field indicate a relatively favorable basis for the development of pro-environmental behavior among students in Serbia. The highest level of agreement was recorded within the construct of perception of polluting materials, confirming that students recognize the importance of contemporary environmental challenges and the risks they pose to human health and ecosystems. Cluster analysis grouped students based on the values of the examined constructs into three clusters. The smallest number of respondents belonged to the group with the lowest values of all three constructs. The obtained results indicate that there is a need for further improvement of educational profiles and strengthening awareness of the importance of standardization in the field of environmental management. Future research should also include various other factors that may influence the attitudes of students regarding the environment, such as politics, media and family, as well as their role in shaping the environmental behavior of the individual.

ACKNOWLEDGMENTS

This work was supported by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia (Contract No. 451-03-136/2025-03/200287).

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES

During the preparation of this work, the authors used Claude (Anthropic) for language editing purposes, including grammar and spelling correction, as well as improving the clarity and readability of the manuscript. The authors reviewed and edited all AI-assisted content and take full responsibility for the published article.

REFERENCES

[1.] Arocena, P., Orcos, R., & Zouaghi, F. (2023). The scope of implementation of ISO 14001 by multinational enterprises: The role of liabilities of origin. *Journal of Environmental Management*, *327*, 116844.
[2.] Matić Bujagić, I., Grujić, S., Đurkić, T., & Laušević, M. (2020). Methods for trace analysis of pharmaceuticals and pesticides in water and sediment samples. IWA Publishing.
[3.] Chan, E. S. W., Hon, A. H. Y., Chan, W., & Okumus, F. (2014). What drives employees' intentions to implement green practices in hotels? The role of knowledge, awareness, concern and ecological behaviour. *International Journal of Hospitality Management*, *40*, 20–28. https://doi.org/10.1016/j.ijhm.2014.03.001
[4.] Dagiliūtė, R., & Liobikienė, G. (2015). University contributions to environmental sustainability: Challenges and opportunities from the Lithuanian case. *Journal of Cleaner Production*, *108*, 891–899.
[5.] Khan, S., Naushad, M., Govarthanan, M., Iqbal, J., & Alfadul, S. M. (2022). Emerging contaminants of high concern for the environment: Current trends and future research. *Environmental Research*, *207*, 112609.
[6.] Loste, N., Roldán, E., Lomba, L., & Giner, B. (2019). Green Chemistry and Environmental Management Systems: Relationships, Synergies, Advantages and Barriers of Joint Implementation at Universities. *Environmental Management*, *64*(6), 783–793.
[7.] Righi, S., Bergami, E., Simonini, R., & Prevedelli, D. (2024). Level of Awareness and Attitudes towards Plastic Contamination by Students of an Italian University. *Sustainability*, *16*(11), 4637.
[8.] Song, Y., Zhao, C., & Zhang, M. (2019). Does haze pollution promote the consumption of energy-saving appliances in China? An empirical study based on norm activation model. *Resources, Conservation and Recycling*, *145*, 220–229.
[9.] Trajković, D., & Glogovac, M. (2025). Procena ekološke svesti stanovnika Republike Srbije u kontekstu primene ISO 14001 standarda. *Procesna tehnika*, *37*(2), 36–41.
[10.] Trajković, D. V., Maletić, M. M., Vukčević, M. M., Veljović, Đ. N., Perić Grujić, A. A., & Živojinović, D. Z. (2025). Geopolymers Based on Fly Ash for Organic Dye Removal from Water. *Separations*, *12*(11), 299.
[11.] Vainio, A., & Paloniemi, R. (2014). The complex role of attitudes toward science in pro-environmental consumption in the Nordic countries. *Ecological Economics*, *108*, 18–27.
[12.] Vukčević, M., Trajković, D., Maletić, M., Mirković, M., Perić Grujić, A., & Živojinović, D. (2024). Modified Fly Ash as an Adsorbent for the Removal of Pharmaceutical Residues from Water. *Separations*, *11*(12), 337.

APPENDIX

Table A2: Questionnaire

Construct		
Perception of polluting materials (PPM)	PPM1	I believe that the accumulation of heavy metals in waterways poses a critical long-term threat to public health.
	PPM2	I believe that particulate matter (PM2.5 and PM10) is the most significant cause of air quality deterioration in our cities.
	PPM3	I am concerned about the emission of polluting gases from energy and industrial facilities.
	PPM4	I am very concerned about the presence of microplastics in drinking water and the risks they pose to human health.
	PPM5	I believe that pharmaceutical residues in water represent a serious, yet insufficiently recognized environmental hazard.
Environmental education and science (EES)	EES1	I believe that formal education is essential for improving environmental awareness.
	EES2	I believe that scientific research can offer key solutions to current environmental problems.
	EES3	I consider scientific knowledge to be the most reliable source of information about the actual state of the environment.
	EES4	I believe that the educational system in Serbia adequately contributes to the understanding of environmental issues.
	EES5	I believe that technological innovations derived from science can drastically reduce pollution levels.
ISO 14001 standards (ISO)	ISO1	Organizations that implement an environmental management system (ISO 14001) have a lesser negative impact on the environment.
	ISO2	The implementation of an environmental management system contributes to better control and reduction of industrial pollution.
	ISO3	I believe that the ISO 14001 standard makes it easier for companies to operate in compliance with environmental regulations and legislation.

Original Research Article https://doi.org/10.5281/zenodo.23398867

APPLIED AI FOR CARDIOVASCULAR DISEASE

Nenad Filipovic[1,2*] [0000-0001-9964-5615], Djordje Jakovljevic[3] [0000-0003-2686-6542], Zoran Bosnic[4] [0000-0003-0501-7146] and Dimitrios Boucharas[5] [0000-0002-8613-2746]
[1] BioIRC doo Kragujevac, Andre Marinkovića 26, 34000 Kragujevac, Serbia
[2] Faculty of Engineering, University of Kragujevac, Sestre Janjic 6, 34000 Kragujevac, Serbia
[3] University Hospitals Coventry and Warwickshire NHS Trust, Coventry, UK
[4] University of Ljubljana, Faculty of Computer and Information Science, Ljubljana, Slovenia
[5] Department of Biomedical Research, Foundation for Research and Technology Hellas, Greece
*Corresponding author, e-mail: fica@kg.ac.rs

Abstract: *Heart failure (HF) affects up to 15 million individuals across Europe and continues to represent a major burden in terms of hospitalizations, mortality, and healthcare costs. Given its complex and heterogeneous nature, HF demands more advanced approaches for early detection, accurate risk stratification, and personalized clinical management. The STRATIFYHF project addresses these challenges through the development of an advanced artificial intelligence (AI)–driven Decision Support System (DSS) that integrates multiscale computational modeling with large-scale real-world data. The platform leverages diverse patient-specific inputs—including demographic, clinical, genetic, lifestyle, and socio-economic factors—combined with state-of-the-art machine learning algorithms to enable improved risk prediction, early diagnosis, and forecasting of disease progression. By facilitating earlier clinical intervention and enabling personalized treatment strategies, STRATIFYHF aims to reduce hospital admissions, slow disease progression, and contribute to more efficient, sustainable, and patient-centered healthcare systems.*

Keywords: *Heart Failure, Artificial Intelligence, Decision Support System, Cardiac Modeling, Finite Element Modeling*

1 INTRODUCTION

Heart failure (HF) represents a growing global public health challenge, affecting more than 15 million individuals across Europe and over 60 million people worldwide [1],[2]. It is a complex, multifactorial clinical syndrome characterized by the heart's inability to pump sufficient blood to meet the body's metabolic demands, resulting in symptoms such as dyspnea, fatigue, and fluid retention. Despite advances in pharmacological and device-based therapies, HF remains one of the leading causes of hospitalization among older adults and is associated with high mortality rates, frequent readmissions, reduced quality of life, and substantial socioeconomic burden [3],[4]. The economic impact is considerable, accounting for approximately 1–2% of total healthcare expenditure in developed countries, largely driven by hospital admissions and long-term disease management [5].

The clinical heterogeneity of HF—arising from diverse etiologies, comorbidities, and patient-specific factors—poses significant challenges for timely diagnosis, accurate risk stratification, and optimal therapeutic decision-making [6]. Current clinical guidelines rely on a combination of imaging, biomarkers, and clinician judgment; however, these approaches are often insufficient to capture the dynamic and nonlinear progression of the disease, particularly in its early stages [2]. Furthermore, the increasing availability of large-scale healthcare data, including electronic health records, wearable sensor data, and genomic information, has outpaced the capacity of traditional analytical methods to extract clinically actionable insights [7].

In this context, artificial intelligence (AI) and machine learning (ML) have emerged as transformative technologies with the potential to revolutionize HF management. AI-driven models can integrate heterogeneous, high-dimensional datasets to uncover hidden patterns, predict disease trajectories, and support personalized treatment strategies [8],[9]. Recent studies have demonstrated the utility of AI in improving HF risk prediction, early detection, phenotyping, and outcome forecasting, often outperforming conventional statistical approaches [10],[11]. However, most existing solutions remain fragmented, lack clinical integration, or fail to incorporate mechanistic insights from physiological and biomechanical modeling.

The STRATIFYHF [12] project addresses this critical gap by developing, validating, and implementing a comprehensive, AI-based Decision Support System (DSS) for HF. The project aims to deliver the first integrated platform capable of assessing individual risk, enabling early diagnosis, and predicting disease progression through the fusion of multiscale data and advanced computational methods (Figure 1).

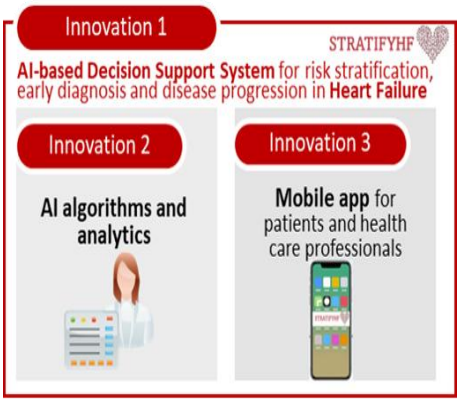


Figure 1: Three main innovations in the STRATIFYHF project

2 METHODS

2.1 Clinical Study

Eight European Clinical Centre of Excellence in diagnosis and management of heart failure, who are part of the STRATIFYHF project, have delivered retrospective and prospective clinical studies since the beginning of the project. In the retrospective phase, Centres managed to collect demographic and clinical data for 5,624 patients with confirmed diagnosis of heart failure, and 4,465 patients with suspected heart failure. In additional, access to Clinical Practice Research Data link allowed access to 11 million patient records who are at risk of developing heart failure, and 680,000 patients with confirmed heart failure. These data have been used to develop risk stratification, diagnostic and prognostic models. We have also analysed and disseminated findings on multimorbidity in heart failure using retrospective data (Figure 2).

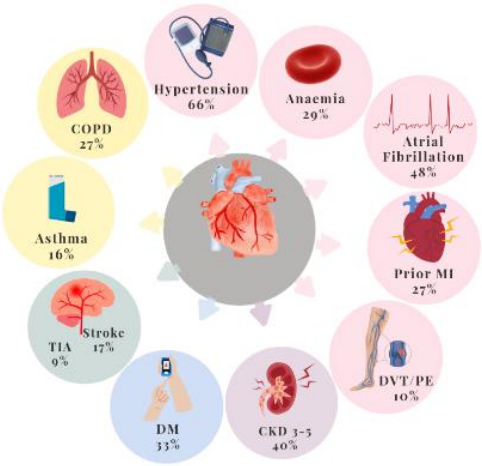


Figure 2: Co-morbidities in heart failure population. Abbreviations: MI, myocardial infarction; DVT, deep venous thrombosis; CKD, chronic kidney disease; DM, diabetes mellites; TIA, transient ischaemic attack; COPD, chronic obstructive pulmonary disease.

Prospective phase of the clinical study recruited 350 patients. All patients underwent a standard heart failure diagnostic clinical investigation including medical history, physical examination, blood test for NTproBNP, electrocardiography, echocardiography. In addition, investigations included the two novel technologies such as cardiac output response to stress test, and voice recognition and voice biomarkers assessment. The prospective phase of the clinical study is used to validate and further refine the accuracy of the risk stratification, diagnostic and prognostic models developed using data from the retrospective phase of the clinical study.

2.2 Decision Support System

STRATIFYHF cloud-based Decision Support System (DSS) is a platform designed to assist healthcare professionals in the risk stratification, early diagnosis, and management of heart failure. A milestone has been reached with the finalization of the formal Software Architectural Design, which provides the structured roadmap necessary for the system's technical realization. The STRATIFYHF Conceptual Multi-layer Hierarchical Framework is shown in the figure below. As illustrated in the Conceptual Multi-layer Hierarchical Framework below, we have defined a five-layer architecture—spanning hardware, security, workflow, back-end, and front-end layers. The functionality of the STRATIFYHF DSS is categorized into specialized modules: The Data

Management Module (DMM) serves as the gateway, utilizing a Data Quality Control Engine to clean, normalize, and assess the reliability of incoming patient records. Central to the system's operation is the Workflow Manager Module (WMM), which acts as an orchestrator, employing a Docker Engine to deploy and manage high-computation tasks like the 3D Computer Modeling (PAK) tool for biomechanical heart simulations. Analytical power is provided by the Risk Stratification and Early Diagnosis modules, which apply machine learning algorithms to identify high-risk patients and detect heart failure markers that might be missed during routine exams. Finally, the Visual Analytics Module (VAM) and the User Access Management Module (UAMM) ensure that these complex outputs are translated into intuitive, secure, and interactive dashboards, allowing clinicians to visualize disease progression and voice biomarker trends directly within their workflow.

3 RESULTS AND DISCUSSION

Here we focused on the design and specification of the STRATIFYHF mobile application, entitled mAIHeart, laying the foundations for a user-centered, secure, and scalable digital solution for heart failure management and end-user adherence to prevention and treatment strategies. The work conducted spans from the systematic elicitation of end-user requirements to the definition of the core architectural components that enable the application's implementation and integration with the wider STRATIFYHF ecosystem.

The first phase of work, concentrated on understanding the needs of the two primary end-user groups: heart failure patients or individuals at high risk of developing heart failure, and healthcare professionals operating in primary and secondary care settings. This process combined (i) a literature review on heart failure risk factors and (ii) management strategies with a structured analysis of the existing mobile health software landscape. In parallel, dedicated questionnaires and interviews were conducted with patients, carers, and healthcare professionals, ensuring early and continuous user involvement (Co-Design/Co-Create). The collected insights were translated into functional and non-functional requirements, addressing usability, security, performance, and regulatory considerations, and were further articulated through detailed user stories capturing real-world clinical and self-management scenarios.

Building on the validated requirements, we focused on the definition of the mAIHeart reference architecture and its core architectural components. A modular client–server architecture was designed to support scalability, interoperability, and secure data handling. The architecture clearly separates patient-facing and healthcare professional-facing functionalities, while introducing dedicated layers for data management, business logic, communication, and security. Emphasis was placed on the integration of wearable devices, external storage systems, and Decision Support Systems, enabling continuous data collection, AI-driven analytics, and real-time risk assessment. Detailed data flow diagrams and interface specifications were produced to describe key operational scenarios, including sensor data ingestion, manual data entry, reporting, goal management, alerts, and asynchronous communication. Robust security and compliance mechanisms, such as role-based access control, OAuth-based authentication, AES encryption, and GDPR-aligned data governance, were embedded in the architecture by design [13-15].

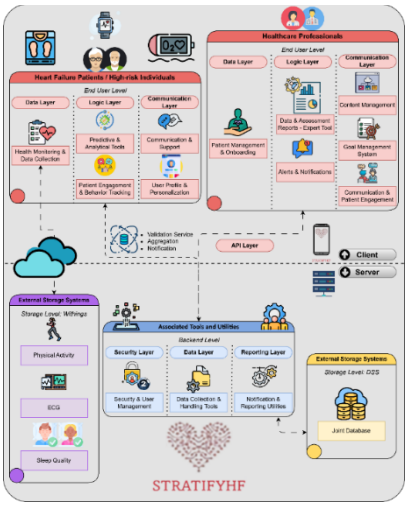


Figure 3: Reference Architecture of mAIHeart Mobile Application

4 CONCLUSION

The STRATIFYHF consortium is a well-balanced combination of partners with complementary knowledge and skills. The consortium consists of 15 partners, which comprises Universities and research institutions, including 8 clinical centres, as well as 3 SMEs and two not-for-profit organisations, from 11 different countries across Europe.

STRATIFYHF will change the ways in which HF is managed today, thereby improving the quality and length of patients' lives. Solution in stratifyhf will lead towards an efficient and sustainable healthcare systems by reducing the number of HF-related hospital admissions and deaths, and unnecessary referrals from primary to secondary care

Acknowledgements: This study has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101080905 (STRATIFYHF). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or granting authority – European Health and digital executive agency (HADEA). Neither the European Union nor the granting authority can be held responsible for them. This research was also supported by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia, contract number [451-03-33/2026-03/200378 (Institute of Information Technologies, University of Kragujevac) and 451-03-34/2026-03/200107 (Faculty of Engineering, University of Kragujevac)].

REFERENCES

[1] Savarese, G., & Lund, L. H. (2017). Global public health burden of heart failure. Cardiac Failure Review, 3(1), 7–11.

[2] McDonagh, T. A., Metra, M., Adamo, M., et al. (2021). 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. European Heart Journal, 42(36), 3599–3726.

[3] Ponikowski, P., Voors, A. A., Anker, S. D., et al. (2016). ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. European Heart Journal, 37(27), 2129–2200.

[4] Virani, S. S., Alonso, A., Aparicio, H. J., et al. (2023). Heart disease and stroke statistics—2023 update. Circulation, 147(8), e93–e621.

[5] Cook, C., Cole, G., Asaria, P., Jabbour, R., & Francis, D. P. (2014). The annual global economic burden of heart failure. International Journal of Cardiology, 171(3), 368–376.

[6] Esteva, A., Robicquet, A., Ramsundar, B., et al. (2019). A guide to deep learning in healthcare. Nature Medicine, 25, 24–29.

[7] Groenewegen, A., Rutten, F. H., Mosterd, A., & Hoes, A. W. (2020). Epidemiology of heart failure. European Journal of Heart Failure, 22(8), 1342–1356.

[8] Johnson, K. W., Soto, J. T., Glicksberg, B. S., et al. (2018). Artificial intelligence in cardiology. Journal of the American College of Cardiology, 71(23), 2668–2679.

[9] Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. Nature Medicine, 25(1), 44–56.

[10] Krittanawong, C., Zhang, H., Wang, Z., Aydar, M., & Kitai, T. (2017). Artificial intelligence in precision cardiovascular medicine. Journal of the American College of Cardiology, 69(21), 2657–2664.

[11] Attia, Z. I., Kapa, S., Lopez-Jimenez, F., McKie, P. M., Ladewig, D. J., Satam, G., ... & Friedman, P. A. (2019). Screening for cardiac contractile dysfunction using an artificial intelligence–enabled electrocardiogram. Nature Medicine, 25(1), 70–74.

[12] Choi, E., Schuetz, A., Stewart, W. F., & Sun, J. (2020). Using recurrent neural network models for early detection of heart failure onset. Journal of the American Medical Informatics Association, 27(4), 543–551.

[13] StratifyHF EU project, www.stratifyhf.eu, 2026

[14] Stojanovic, B., Milosevic, M., Kojic, M., & Filipovic, N. (2021). Multiscale finite element simulation using the Huxley muscle model. Journal of the Serbian Society for Computational Mechanics, 15(1), 1–23.

[15] Milićević, B., Ivanović, M., Stojanović, B., Milošević, M., Kojić, M., & Filipović, N. (2022). Huxley muscle model surrogates for high-speed multi-scale simulations of cardiac contraction. Computers in Biology and Medicine, 149, Article 105963. https://doi.org/10.1016/j.compbiomed.2022.105963

[16] Kojic, M., Milosevic, M., Ziemys, A., & Filipovic, N. (2022). Computational models in biomedical engineering: Finite element models based on smeared physical fields: Theory, solutions, and software. Elsevier.

CAN AI REPLACE HUMAN EXPERTS IN CROWD VOTING? A COMPARATIVE STUDY USING BARGAINING-BASED AGGREGATION

Ana Vukicevic[2][0000-0003-3014-9116], Milan Vukicevic[1][0000-0002-1631-6531], Sandro Radovanovic[1][0000-0002-9975-8844], Boris Delibasic[1][0000-0002-6153-5119], Milija Suknovic[1][0009-0003-0449-5258]

[1] University of Belgrade, Faculty of Organizational Sciences, Jove Ilica 154, 11000 Belgrade, Serbia
[2] Union University, School of Computing, Kneza Mihaila 6/VI, 11000 Belgrade, Serbia

Abstract: *Crowd voting exploits the "wisdom of crowds" at low cost, but is fragile under sparse, heterogeneous or multimodal votes; augmenting it with human-expert votes improves quality at the price of recruitment cost. Our prior BargCrEx system combines a small number of expert votes with crowd votes through matrix factorization for vote estimation and Kalai–Smorodinsky / Nash bargaining for aggregation. In this paper we replace the human-expert layer with a panel of three current AI systems (Claude Sonnet 4.6, DeepSeek v3.2, GPT 5.4) prompted with the same instructions used for human experts. Reusing the Credibility Coalition data and the same pipeline, we compare AI- and human-expert configurations for both Scientist and Journalist groups. AI experts match human experts on grade prediction and stay competitive on cumulative satisfaction although human experts retain an edge there, most clearly for journalists, while being substantially cheaper and instantly available. This positions AI experts as a strong, low-cost complement to, and under tight expert budgets a partial substitute for, human experts, rather than a full replacement.*

Keywords: *Crowd voting, Expert opinion, Large language models, AI experts, Bargaining models, Vote aggregation.*

1. INTRODUCTION

Crowdsourcing and crowd voting are widely used to solve tasks that need many judgments at low cost, from labelling and ranking to participatory budgeting and contest scoring. Their value rests on the well-known "wisdom of crowds" effect, but in practice crowd-voting data are sparse (often 1–3 votes per alternative), heterogeneous and noisy, which makes central-tendency aggregation susceptible to outliers and to bias under multimodal or near-uniform vote distributions.

A common remedy is to enrich crowd votes with a small number of expert votes that anchor the aggregation. Our prior BargCrEx system [1] combines crowd and expert votes through matrix factorization (ALS) plus a regression layer for vote estimation, and through Kalai–Smorodinsky / Nash bargaining for aggregation. BargCrEx outperforms classical rules in cumulative satisfaction, but still relies on expensive, slow-to-recruit human-expert votes, the binding cost in many real applications.

Recent large language models (LLMs) can read, summarise and grade textual content with consistency that rivals trained human raters, scale at marginal cost and respond on demand. This raises a natural question: can AI systems play the role of expert voters in BargCrEx, while preserving aggregated-decision quality and substantially reducing cost? We keep the BargCrEx pipeline unchanged and replace the human-expert layer with an AI-expert panel, allowing a direct comparison on the same task. Section 2 summarises BargCrEx, Section 3 describes the experimental setup, and Section 4 reports the results and conclusions.

2. THE BARGCREX SYSTEM

Since the same aggregation system is reused here, we briefly summarize BargCrEx [1]. The system targets crowd-voting problems with many alternatives, sparse votes and a strict expert budget, and addresses two coupled problems: estimating missing votes under the limited-budget assumption and aggregating (true and estimated) crowd and expert votes into a single final grade per alternative.

Vote estimation. BargCrEx fills the voter–alternative matrix in two steps. Alternating Least Squares (ALS) matrix factorization is first applied jointly to crowd and expert votes to produce low-dimensional voter and alternative embeddings. A regression layer (Random Forest, Gradient Boosting Trees) is then trained on these embeddings to predict every voter–alternative grade. The procedure is unsupervised, needs no "gold" data, and once trained yields expert grades for every alternative from a small fraction of collected expert votes, the main source of cost reduction in BargCrEx.

Bargaining-based aggregation. Given expert and crowd grade vectors per alternative, BargCrEx aggregates them by solving max $[\lambda \cdot ES(x) + (1-\lambda) \cdot CS(x)]$, where $ES(x)$ and $CS(x)$ are the maximum grade minus the average absolute deviation of votes within each group from the candidate grade x, and λ encodes the relative importance of the two groups, set automatically from their internal level of agreement via the Kalai–Smorodinsky or Nash bargaining solution. More homogeneous groups obtain greater bargaining power, which mitigates

the failures of central-tendency aggregation under multimodal vote distributions. Crucially, BargCrEx treats experts as an undifferentiated group of agents, so human experts can be directly replaced with AI experts without modifying the framework.

3. EXPERIMENTAL SETUP

Data used for experiments are the same as in [1]. Briefly, they come from the Credibility Coalition (https://credibilitycoalition.org/) and consist of 2,472 environment and climate related articles from 2019, each rated on a 5-point Likert credibility scale by a set of crowd raters and by two expert groups, scientists and journalists. The two expert groups are kept separate, so the dataset defines two paired tasks (Scientist and Journalist), each evaluated by the crowd and by the corresponding expert group; the data are publicly available at https://data.world/credibilitycoalition/credibility-factors2020/.

To simulate expert opinions for both Scientist and Journalist groups we used three current AI systems – Claude Sonnet 4.6, DeepSeek v3.2 and GPT 5.4 as AI experts. Each model graded the same articles as the original human experts, using a role-conditioned prompt to act as a scientist or journalist and rate each article on the five-level human credibility scale (mapped to the 5-point scale). The exact prompt and full code are available at https://github.com/ana-kovacevic/Expert_Crowd_Grade_Optimization.

All other aspects of the pipeline match the original paper [1]: ALS produces voter and alternative embeddings from the joint crowd + AI-expert vote matrix; Linear Regression (LR), Random Forest (RF) and Gradient Boosted Trees (GBT) regress true grades on embeddings under the same hyper-parameter grid as in [1]; and the BargCrEx aggregation methods are run over both true and estimated grades. Supervised models are evaluated by MAE on a held-out test set and aggregation methods by the cumulative satisfaction sum and area introduced in [1].

4. RESULTS AND FUTURE WORK

Table 1 reports the best regression model per expert source. Train and test MAE are very close between AI and human experts; in fact, models trained on AI-expert grades reach slightly lower test MAE in both Scientist and Journalist groups (the gap is below 0.02 MAE either way), so the regression layer recovers AI-expert preferences at least as accurately as human-expert preferences.

Table 1: Best grade-estimation model per expert source (lower MAE is better).

Data Set	Model	Best parameters	MAE (Train)	MAE (Test)
Scientist	RF	{'max_depth': 15, 'n_estimators': 500}	0.549	0.574
AI Scientist	GBT	{'learning_rate': 0.1, 'max_depth': 10, 'n_estimators': 500}	0.538	0.561
Journalist	RF	{'max_depth': 15, 'n_estimators': 200}	0.542	0.576
AI Journalist	GBT	{'learning_rate': 0.1, 'max_depth': 10, 'n_estimators': 500}	0.553	0.570

Cumulative satisfaction represented as sum of crowd and expert satisfaction across the BargCrEx aggregation methods is shown in Fig. 1 for the four expert sources. Satisfaction sum is slightly higher with human-expert grades, with the largest gain for Human Journalist, which dominates across all methods. For AI experts the Scientist votes give more satisfaction in most cases, except under crowd-only methods, where the Journalist votes are better.

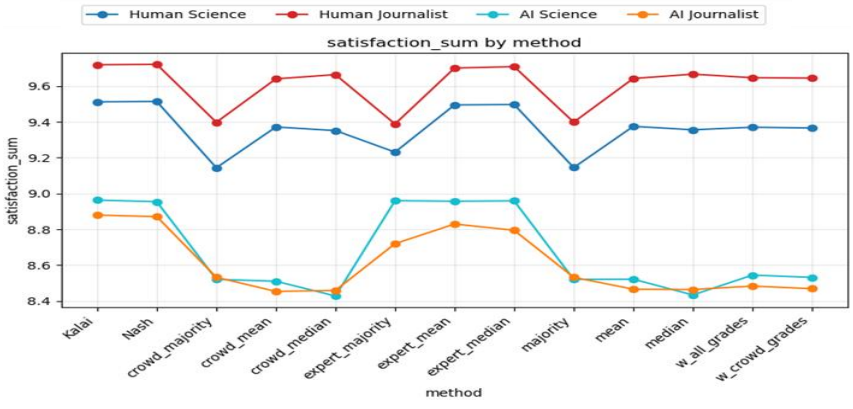


Figure 1: Cumulative satisfaction (sum) by aggregation method, for the four experts

Overall, human experts retain a small satisfaction edge for Scientists and a larger one for Journalists, but AI experts match them on grade prediction and stay competitive across most aggregation methods while being inexpensive, on-demand and reproducible, clear advantages over slow-to-refresh human panels. They thus serve as a complement to, and under tight budgets (rating about 40 articles cost $0.84) a partial substitute for human experts in BargCrEx-style crowd voting, rather than a full replacement; further investigation is warranted. Limitations include reliance on a single dataset, possible model-specific bias, and unverified domain expertise. Future work will explore prompting strategies and AI-expert personas, grade all documents with AI experts for denser estimation, use AI-expert disagreement as a bargaining signal, add domain-specialized and fine-tuned models, and test contest-scoring and participatory-budgeting tasks where human-expert cost is prohibitive.

REFERENCES

[1] Vukicevic, A., Vukicevic, M., Radovanovic, S., Delibasic, B.: BargCrEx: A System for Bargaining Based Aggregation of Crowd and Expert Opinions in Crowdsourcing. Group Decis Negot 31, 789–818 (2022). https://doi.org/10.1007/s10726-022-09783-0

[2] Kovacevic, A., Vukicevic, M., Radovanovic, S., Delibasic, B. CrEx-Wisdom Framework for Fusion of Crowd and Experts in Crowd Voting Environment – Machine Learning Approach. In: Bellatreche, L., et al. ADBIS, TPDL and EDA 2020 Common Workshops and Doctoral Consortium. TPDL ADBIS 2020. Communications in Computer and Information Science, vol 1260. Springer, Cham. https://doi.org/10.1007/978-3-030-55814-7_11

[3] Keuschnigg, M., Ganser, C.: Crowd Wisdom Relies on Agents' Ability in Small Groups With a Voting Aggregation Rule. Management Science 63(3), 818–828 (2017). https://doi.org/10.1287/mnsc.2015.2364

[4] Kalai, E., Smorodinsky, M.: Other Solutions to Nash's Bargaining Problem. Econometrica 43(3), 513–518 (1975). https://doi.org/10.2307/1914280

[5] Snow, R., O'Connor, B., Jurafsky, D., Ng, A.: Cheap and Fast – But Is It Good? Evaluating Non-Expert Annotations for Natural Language Tasks. In: Proceedings of EMNLP 2008, pp. 254–263 (2008).