ITIS 25
11111101001

Fakulteta za informacijske študije
Faculty of information studies

ITIS 2025 BOOK OF PROCEEDINGS

"Building a Sustainable Future with AI and Human-Centric Digital Innovation"

16TH INTERNATIONAL CONFERENCE ON INFORMATION TECHNOLOGIES AND INFORMATION SOCIETY

DECEMBER 2025, NOVO MESTO, SLOVENIA

16th International Conference on Information Technologies and Information Society

ITIS 2025

Conference proceedings

November 12-13, 2025
Otočec, Slovenia

Organized by

Fakulteta za informacijske študije
Faculty of information studies

RUDOLFOVO Science and Technology Centre Novo mesto

GRM Visoka šola za upravljanje podeželja Grm Novo mesto
Landscape Governance College Grm Novo mesto

FIN Fakulteta za industrijski inženiring
Faculty of Industrial Engineering

FOS Fakulteta za organizacijske študije
Faculty of organisation studies

Supported by

ARIS Slovenian Research and Innovation Agency

2025 science month

Slovenian AI Factory (SLAIF)

CONFERENCE COMMITTEES:

Organizing committee:

Maruša Gorišek, *chair*, Faculty of Information Studies in Novo mesto, Slovenia
Tea Golob, *chair*, Faculty of Information Studies in Novo mesto, Slovenia
Biljana Mileva Boshkoska, *chair*, Faculty of Information Studies in Novo mesto, Slovenia
Blaž Rodič, *chair*, Faculty of Information Studies in Novo mesto, Slovenia
Danica Kodela, Faculty of Information Studies in Novo mesto, Slovenia
Neža Repanšek, Faculty of Information Studies in Novo mesto, Slovenia
Aljaž Blatnik, Faculty of Information Studies in Novo mesto, Slovenia
Nika Robida, Faculty of Information Studies in Novo mesto, Slovenia
Kseniia Gromova, Faculty of Information Studies in Novo mesto, Slovenia
Mateja Lesar, Faculty of Information Studies in Novo mesto, Slovenia
Katja Peterlin, Faculty of Information Studies in Novo mesto, Slovenia
Alex Zorko, Faculty of Information Studies in Novo mesto, Slovenia
Teja Štrempfel, Faculty of Information Studies in Novo mesto, Slovenia
Lea-Marija Colarič Jakše, Landscape Governance College Grm in Novo mesto, Slovenia
Annmarie Gorenc Zoran, Faculty of Organization Studies in Novo mesto, Slovenia
Simon Muhič, Faculty of Industrial Engineering in Novo mesto, Slovenia
Janez Povh, Rudolfovo – Science and Technology Centre in Novo mesto, Slovenia

Program committee:

Zoran Levnajić, *chair*, Faculty of Information Studies in Novo mesto, Slovenia
Nuša Erman, Faculty of Information Studies in Novo mesto, Slovenia
Filipo Sharevski, DePaul University, Chicago, USA
Vesna Andova, "SS. Cyril and Methodius" University in Skopje, North Macedonia
Galia Marinova, Technical University of Sofia, Bulgaria
Jana Suklan, Newcastle University, UK
Małgorzata Pańkowska, University of Economics in Katowice, Poland
Marija Mitrović Dankulov, Institute of Physics Belgrade, Serbia
Dolores Modic, Nord University, Norway
Vida Vukašinović, Jožef Stefan Institute, Slovenia
Mateja Rek, School of Advanced Social Studies, Slovenia
Srđan Škrbić, Faculty of Information Studies in Novo mesto, Slovenia
Mirjana Mikalački, University of Novi Sad, Faculty of Science and Mathematics, Serbia
Ana Bezić, Institute for Contemporary Art (IZK), TU Graz (Graz University of Technology)
Ana Meštrović, University of Rijeka, Croatia

Published by: Faculty of information studies in Novo mesto
Ljubljanska cesta 31 a
8000 Novo mesto, Slovenia

Published in: 2025

Edited by: Maruša Gorišek, Tea Golob, Teja Štrempfel

Kataložni zapis o publikaciji (CIP) pripravili v Narodni in univerzitetni knjižnici v Ljubljani /
Cataloguing record of the publication (CIP) prepared by the National and University Library in Ljubljana

COBISS.SI-ID: 263628291

ISBN: 978-961-96549-2-7

Editors' note: *Participants were asked to submit papers written in English. Keywords were chosen by participants, and their number has not been reduced. Authors of submissions are responsible for the reliability of contents and other statements made in their work. Submissions are not proofread.*

PROGRAM

DAY 1: 12. 11. 2025 – Hotel Šport Otočec

9.00 – 9.30: Opening of the conference

- prof. dr. Matej Makarovič, Dean of the Faculty of Information Studies in Novo mesto

9.30 – 11.00: Session 1: SLAIF (Chair: Biljana Mileva Boshkoska)

- Sagnika Sen (Penn State Great Valley) – From Insight to Impact: AI-Powered Process Improvement in Healthcare and Curriculum Design (Keynote speech)
- Pavle Boškoski, Martin Brešar – Tracking Physiological System Integration in Cardiac Rehabilitation Using the Cross-Vector Approach
- Albert Zorko – Advancing AI-Based Depression Detection: A Preliminary Study on Feature Optimization and Model Robustness
- Borut Lužar; Nika Robida – Evolution of topics in Slovenian science

11.30 – 13.14: Session 2: Cybersecurity and Data-Driven Innovation (Chair: Blaž Rodič)

- Matjaž Drev – Extending the Privacy by Design Model to Address NIS 2 Cybersecurity Requirements
- Boštjan Delak; Matjaž Drev – Cybersecurity Auditing
- Andrej Bregar – Collaborative Decision-Making for Cybersecurity
- Martin Žnidaršič – Multi-criteria model for evaluation of sustainable transportation initiatives
- Srdjan Škrbić – Impact of Filtering Policy Changes on Wikipedia Pageview Metrics
- Matija Klančar – Climate Cards: Battle of the Weather Stations!

14.30 – 16.00: Session 3: AI4VET4AI (Chair: Alenka Pandiloska Jurak)

- Dan Podjed (ZRC – SAZU) – Staying Human in the Age of AI (Keynote speech)
- Tea Golob; Matej Makarovič; Romina Gurashi – Loneliness, Reflexivity, and AI in Youth
- Ana Hafner – Can artificial intelligence invent in Slovenia?
- Urša Lamut – Social Infrastructure for Digitalization

16.30 – 17.30: Session 4: AI and Heritage (Chair: Vesna Pungerčar)

- Ines Vodopivec – Reliable and Trustworthy Use of AI in Cultural Heritage Digital Transformation
- Lucijano Berus; Vesna Pungerčar – Cultural Heritage analysis with YOLO based object detection
- Branimir Kolarek; Davor Davidović – A Two-Phase AI Framework for Fresco Analysis and Digital Restoration

17.45 – 18.30: Session 5: AI in Education – AI2MED (Chair: Biljana Mileva Boshkoska)

- Katarina Rojko – The Use of Artificial Intelligence in Education

- Branka Klarić – Promoting Digital Inclusion through Digital Tools: The Role of eAsistent in the Work of Secondary School Teachers in Slovenia
- Nejra Arnautović; Denis Kantić – A Digital Transformation for Student Living in Slovenia: Addressing Key Challenges

Day 2: 13. 11. 2025 – Hotel Šport Otočec

9.00 – 10.30: Session 1: SLAIF (Chair: Borut Lužar)

- Tatjana Petrov (University of Trieste) – To sting or not to sting: Unraveling Collective Behavior in Honeybees via Stochastic Modeling (Keynote speech)
- Miloš Ivanović; Milan Matijević; Lazar Krstić – Inverse Modeling of Flexible Rotational Systems via Physics-Informed Deep Learning
- Andrej Furlan – PINN: machine learning on complex systems described by small or limited quality datasets
- Robi Pritržnik – How should we benchmark community detection algorithms in complex networks?

11.00 – 12.45: Session 2: Smart Industry and Infrastructure (Chair: Aljaž Blatnik)

- Nika Brili – Smart AI-Based System for Turning Tool Condition Monitoring
- Jelena Topić Božič; Andreja Dobrovoljc; Simon Muhič – Time-resolved Life Cycle Assessment for Sustainable Industry: Integrating Hourly Analysis into Smart Infrastructure and Energy Management
- Valerij Grašič; Biljana Mileva Boshkoska – Global Electric Circuit as a Driver of Space Weather Impacts: Cross-Sectoral Risks for Energy and Digital Infrastructures with a Spain Blackout Case Study
- Mare Srbinovska; Zivko Kokolanski; Vladimir Dimcev; Dimitar Taskovski; Marija Markovska Dimitrovska; Maja Celeska Krstevska – Digital Twin and Machine Learning for Industrial Measurement and Diagnostics in Industry 4.0
- Tomaž Podobnikar – Perceived vs. Actual Spatial Data Quality: Challenges, Consequences, and an Innovative Management Framework
- Andrej Dobrovoljc – Establishing a Data-Driven Feedback Loop for the Optimization of Production Processes

14.00 – 15.30: Session 3: AI and Machine Learning Applications (Chair: Zoran Levnajić)

- Boštjan Gabrovšek (Rudolfovo – Science and Technology Centre Novo mesto) – Classifying knotted proteins in the age of AlphaFold and machine learning (Keynote speech)
- Maja Cerjan; Leo Mršić; Kornelije Rabuzin; Biljana Mileva Boshkoska – Comparative Analysis of Machine Learning Models for Telecommunications Churn Prediction
- Tomaž Aljaž, Uroš Kosanovič – Transforming Generic Flyers into Tailored Promotions: A Case Study in AI-Powered Grocery Retail
- Peter Zupančič; Panče Panov – User Evaluation as Part of Doctoral Research on Intelligent HR Decision-Support Systems

16.00 – 17.15: Session 4: Technology, Innovation, and Society (Chair: Maruša Gorišek)

- Tamara Besednjak Valič; Karin Dobravc Škof – Following Quantum Innovation Flows: The Feedback Loop Between Strategic Timing and Patent Activity (2014–2023)
- Kseniia Gromova – Methodological Framework for Studying Industrial Path Development: Social Fields Analysis
- Milica Stankovic; Gordana Mrdak; Jovana Džoljić; Stevan Simic – Sustainable Practices and Digital Innovation in Serbia and Slovenia: Comparative Analysis and Economic Implications
- Matevž Mandl; Martina Plantak – From Tool to Co-Author: Legal and Ethical Thresholds of Authorship in the Age of Generative AI
- Nadia Molek; Annmarie Gorenc Zoran; Lejla Imamović Lerić, Njerman Ljevo – Exploring the Adoption of Generative AI in Higher Education – A Cross-National Comparison

17.30 – 18.30: Session 5: Jean Monnet Centre of Excellence: Technology and Innovations for Agenda 2030 – EU Global Leadership (TIA2030) (Chair: Kseniia Gromova)

- Kseniia Gromova, Victor Cepoi, Borut Rončević – Round table discussion: The European Union and public diplomacy for Agenda 2030

KEYNOTE SPEAKERS

Lectures by keynote speakers

From Insight to Impact: AI-Powered Process Improvement in Healthcare and Curriculum Design

by Sagnika Sen, PhD, Penn State Great Valley, USA

Staying Human in the Age of AI

by Dan Podjed PhD, ZRC-SAZU, Research Centre of the Slovenian Academy of Sciences and Arts

To sting or not to sting: Unravelling Collective Behaviour in Honeybees via Stochastic Modelling

by Tatjana Petrov, PhD, University of Trieste, Italy

Classifying knotted proteins in the age of AlphaFold and machine learning

by Boštjan Gabrovšek, PhD, Rudolfovo – Science and Technology Centre Novo mesto

CONFERENCE PAPERS

Advancing AI-Based Depression Detection: A Preliminary Study on Feature Optimization and Model Robustness

Albert Zorko
Faculty of Information Studies, Complex Systems and Data Science Lab, Ljubljanska cesta 31A, 8000 Novo mesto, Slovenia
albert.zorko@fis.unm.si

Abstract: *This study constitutes the second part of our investigation presented at ITIS 2023, which explores the search for objective physiological biomarkers for major depressive disorder (MDD). Moving beyond the established role of Heart Rate Variability (HRV), this preliminary research focuses on Pulse-Respiratory Coupling (PRC) – the coordination between cardiac and respiratory rhythms. We hypothesize that depression, characterized by autonomic nervous system (ANS) dysregulation, disrupts this coupling. A group of 73 subjects (healthy controls, untreated depressed patients, and patients treated with tricyclic antidepressants) were submitted to simultaneous electrocardiogram (EKG) and respiratory recording. Analysis revealed a distinct degradation of PRC in the depressed group, manifesting as a loss of synchronous patterns observed in healthy subjects. Machine learning models were trained on features derived from PRC timing. The k-Nearest Neighbors algorithm achieved a promising classification accuracy of 97.3% in distinguishing depressed from healthy individuals, outperforming other classifiers like Random Forest (95.9%) and Support Vector Machine (95.9%). While these results are preliminary and require validation in larger cohorts, they strongly suggest that PRC is a sensitive, non-invasive marker of ANS dysfunction in depression. This work underscores the potential of integrating multi-system physiological analysis with artificial intelligence to create objective aids for psychiatric diagnosis.*

Key Words: *major depressive disorder, physiological biomarkers, pulse-respiratory coupling, heart rate variability, autonomic nervous system, machine learning*

1 Introduction

Major Depressive Disorder (MDD) represents a significant global health challenge, contributing profoundly to disability and reduced quality of life [1]. Contemporary diagnostic practices, as codified in the DSM-5 questionnaire
and ICD-11 [1], rely primarily on subjective self-reporting of symptoms and clinical interviews [2]. This approach, while essential, introduces variability and can lead to misdiagnosis or delayed intervention, particularly in cases of treatment-resistant

[1] The International Classification of Diseases, published by the World Health Organization

depression [3]. Consequently, the pursuit of objective, biologically grounded biomarkers has become a paramount objective in psychiatric research [4].

As established in the first part of this work presented at ITIS 2023 [5], dysregulation of the Autonomic Nervous System (ANS) is a well-documented feature of MDD. This dysregulation is often quantified through Heart Rate Variability (HRV), a measure of the beat-to-beat alterations in heart rate. Reduced HRV, indicating a loss of autonomic flexibility, has been consistently reported in individuals with depression [6, 7]. However, the ANS does not operate in isolation; its function is intimately linked with other physiological systems, particularly respiration. The rhythmic activity of breathing exerts a powerful influence on heart rate, a phenomenon known as respiratory sinus arrhythmia.

This interaction points to a more complex biomarker: the coordination between the cardiac and respiratory cycles, termed Pulse-Respiratory Coupling (PRC) or cardiorespiratory coupling. In healthy individuals at rest, the onset of inspiration is often phase-locked to specific points in the cardiac cycle, creating a stable, synchronous pattern [8]. This coupling is thought to reflect efficient central autonomic control and optimal energy utilization [9]. We hypothesize that the ANS instability inherent to MDD disrupts this fine-tuned coordination, leading to a more random and less coupled state between the pulse and respiration.

Building upon our initial findings with HRV [5], this follow-up study aims to investigate PRC as a novel physiological marker for depression. Our specific objectives are twofold: first, to characterize PRC patterns qualitatively and quantitatively in healthy, depressed, and antidepressant-treated individuals; and second, to evaluate the efficacy of various machine learning algorithms in classifying depression status based solely on PRC-derived features.

2 Related Work

The intersection of physiological signal analysis and artificial intelligence has become a focal point in contemporary mental health research. Numerous studies have explored the potential of bio signals such as heart rate variability (HRV), electrodermal activity (EDA), electroencephalography (EEG), and electrocardiography (EKG) in detecting emotional and cognitive states, particularly those associated with depression and anxiety.

HRV has been extensively studied as a biomarker for autonomic nervous system regulation. Thayer and Lane proposed the neurovisceral integration model, linking HRV to emotional regulation and cognitive flexibility [10]. Inoue has demonstrated in study [11] that individuals with major depressive disorder exhibit significantly reduced HRV parameters, suggesting impaired parasympathetic activity. Similarly, Kreibig reviewed autonomic nervous system activity across emotional states and emphasized HRV's diagnostic relevance [12].

EDA, which reflects sympathetic nervous system arousal, has also been used to assess emotional reactivity. Sano and Picard utilized wearable sensors to monitor stress levels in

real-world settings, finding strong correlations between EDA fluctuations and self-reported mood [13]. Gjoreski extended this approach by combining EDA with accelerometer data to detect stress continuously in daily life [14].

EEG-based emotion recognition has gained traction due to its direct measurement of cortical activity. Al-Shargie applied deep learning models to EEG signals and achieved high accuracy in classifying emotional states [15]. Delorme and Makeig developed EEGLAB, an open-source toolbox that has facilitated widespread EEG analysis in affective computing [9]. However, EEG remains less practical for everyday use due to its complexity and sensitivity to noise.

EKG, while traditionally used for cardiac diagnostics, has proven valuable in extracting HRV features. Two studies have incorporated EKG-derived HRV into multimodal emotion recognition systems, demonstrating its robustness and reliability [16, 17].

From a machine learning perspective, ensemble methods have shown superior performance in biosignal classification. Random Forests, as introduced in study [18], offer high accuracy and resistance to overfitting. XGBoost, developed by Chen and Guestrin, has become a standard in structured data classification due to its scalability and precision. Zhang emphasized the importance of feature selection in improving model interpretability and generalization, particularly in biomedical applications [20].

Compared to these studies, our approach builds on the foundation laid in our ITIS 2023 paper by integrating multiple physiological modalities and applying advanced feature selection techniques. While previous work has often focused on single-signal analysis or deep learning models with limited transparency, our study prioritizes interpretability and clinical relevance. By comparing classical classifiers and optimizing feature sets, we aim to contribute a robust and scalable framework for depression detection that can be adapted to real-world scenarios.

3 Methods

3.1 Participant Recruitment and Characteristics

The study was conducted with approval from the relevant institutional ethics committee. A total of 73 participants were recruited and provided informed consent. The group was divided into three distinct groups to facilitate comparison: control, depressed and treatment group. First control group (n=38) comprised 14 males (mean age ± SD: 36.6 ± 10.4 years) and 24 females (32.8 ± 11.0 years). All participants in this group were confirmed to be free of any medical or psychiatric conditions, as verified by medical history and the Structured Clinical Interview for DSM Disorders (SCID). The second group, the depression group (n=17, included 6 males (24.8 ± 6.3 years) and 11 females (25.5 ± 7.5 years) diagnosed with MDD, melancholic subtype, without clinical agitation. Diagnosis was confirmed using the Beck Depression Inventory-II (BDI-II; mean score 23 ± 12) and the State-Trait Anxiety Inventory (STAI). Participants in this group were not undergoing pharmacological treatment at the time of measurement. The third and final

group was the treatment group (n=18) consisting of 6 males (39.3 ± 9.0 years) and 12 females (39.0 ± 11.0 years), this group included patients diagnosed with MDD who had been undergoing stable pharmacotherapy with tricyclic antidepressants (equivalent to ≥150 mg amitriptyline daily) for a minimum of 14 days. Their mean BDI-II[2] score was 17 ± 13.

3.2 Data Acquisition and Preprocessing

Physiological data were acquired in a controlled laboratory setting. Participants rested in a supine position for a period of acclimatization prior to measurement to minimize initial arousal effects. In order to ensure data for our study, we have to measure two signals: cardiac and respiratory signal.

Electrocardiogram was recorded continuously using a chest-mounted sensor unit at a sampling rate of 200 Hz and a resolution of 10 bits. Second respiratory signal was captured using a nasal thermistor, which provides a sensitive measure of airflow and allows for precise detection of the onset of inspiration.

The raw data were processed to extract key time point markers. The peak of the R-wave in the EKG (`Rpeak`) was detected with a temporal resolution of 1 ms, providing a precise marker of each heartbeat. Similarly, the onset of each inspiration was identified from the respiratory signal with a resolution of 5 ms.

3.3 Feature Extraction: Pulse-Respiratory Coupling (PRC)

The core feature for analysis was the time delay from the most recent `Rpeak` to the immediate onset of inspiration. To normalize for individual differences in heart rate and to create a scale-invariant metric, this absolute delay was divided by the total R-R interval (the time between two consecutive `Rpeaks`) in which the inspiration occurred.

This calculation yielded a normalized PRC time delay value between 0 and 1 for every breath. This process was repeated for all breaths over a standardized nine-minute recording period for each subject. A histogram of these normalized delay values was then constructed for each individual, visualizing the distribution of inspiratory onsets relative to the cardiac cycle.

3.4 Machine Learning and Statistical Analysis

A structured data model was created for machine learning, incorporating the following features for each participant: gender, age, diagnostic label (healthy/depressed), treatment status, and the full set of normalized PRC time-delay values.

This dataset was analyzed using two primary platforms: first we use Weka Toolkit. We employed five distinct classification algorithms known for their varied approaches to learning: J48 Decision Tree, Random Forest, Logistic Regression, k-Nearest Neighbors (k-NN), and a Support Vector Machine (SVM) with a polynomial kernel. Models were

[2] Beck Depression Inventory-II

evaluated based on accuracy, precision, recall, and F1-score. To generate an interpretable model and gain deeper insight into the feature interactions, we also utilized the predictive clustering trees method within the ClusPlus tool. To complement the Weka analysis, we utilized the predictive clustering trees method within the ClusPlus tool, accessible via the ClowdFlows web-based platform [21]. This allowed for additional validation and the generation of a comprehensive decision tree.

Dimensionality reduction and visualization were performed using Principal Component Analysis (PCA) to inspect linear separability, and t-Distributed Stochastic Neighbor Embedding (t-SNE) to explore non-linear cluster structures within the high-dimensional PRC data.

4 Results

4.1 Qualitative Analysis of PRC Patterns

Visual inspection of the PRC histograms revealed striking differences between the groups, as illustrated in Figure 1. The control group consistently displayed a histogram with three distinct peaks, indicating that inspirations were not randomly distributed but preferentially occurred at specific, coupled phases of the cardiac cycle. In stark contrast, the depression group exhibited a notably flatter, more uniform histogram, suggesting a desynchronization of respiratory and cardiac rhythms. The treatment group showed an intermediate pattern, with the reemergence of subtle peaks, hinting at a partial restoration of PRC following antidepressant therapy.

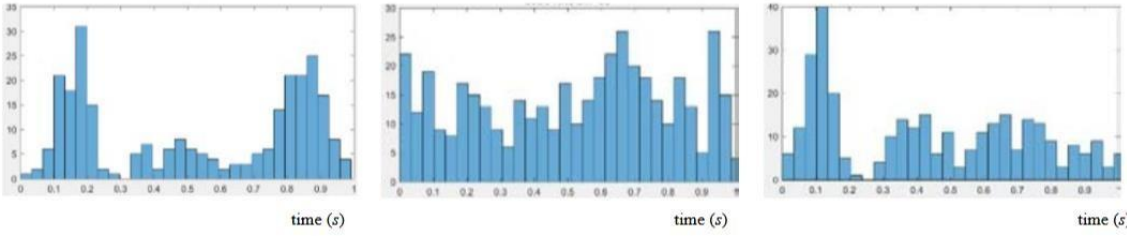

Figure 1. Representative histograms of normalized inspiratory delay times relative to the cardiac cycle. (Left) A healthy 28-year-old subject shows a clear triple-peak pattern. (Center) A 30-year-old depressed subject shows a flat, random distribution. (Right) A 27-year-old treated subject shows a partial return of peak structure.

4.2 Machine Learning Classification Performance

The performance of the various classifiers in distinguishing between healthy and depressed subjects is summarized in Table 1. The k-Nearest Neighbors (k-NN) algorithm demonstrated the highest performance, achieving an accuracy of 97.3% with

commensurately high precision, recall, and F1-score. The Random Forest and Support Vector Machine classifiers also performed exceptionally well, both attaining an accuracy of 95.9%. The J48 Decision tree achieved respectable results (89.0%), while Logistic Regression, a linear model, performed the poorest (80.0%), suggesting that the underlying relationships in the data are predominantly non-linear.

Table 1. Performance metrics of machine learning classifiers for depression detection.

Classifier	Accuracy	Precision	Recall	F1-Score
k-Nearest Neighbor (k-NN)	97.3%	0.97	0.97	0.97
Random Forest	95.9%	0.96	0.96	0.96
Support Vector Machine	95.9%	0.96	0.96	0.96
J48 Decision Tree	90.4%	0.91	0.90	0.90
Logistic Regression	80.0%	0.81	0.80	0.80

Further analysis using the ClusPlus tool indicated that a minimum of six minutes of recorded data was sufficient for the model to achieve consistent and correct classification, suggesting that prolonged measurements may not be necessary for a reliable assessment.

4.3 Decision Tree Analysis

To visualize the classification logic, a decision tree was generated using the ClusPlus tool (Figure 2). The tree reveals a hierarchical structure where age emerges as the primary splitting criterion, highlighting its significant role in differentiating physiological patterns. Notably, the model separates subjects older than 46, indicating that age-related changes in autonomic function are a dominant feature in the dataset. Gender appears as a factor only at lower levels of the tree, suggesting it plays a secondary, more nuanced role compared to age. The position of the 'medicated' feature low in the tree suggests that while treatment status is relevant, its effect on PRC patterns is less pronounced than the fundamental differences between healthy and depressed states.

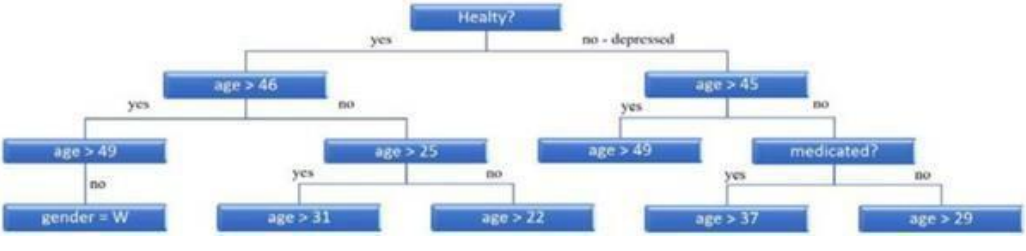

Figure 2. ClusPlus tool with CloudFlows user interface generated decision tree.

4.4 Feature Importance and Sufficiency of Data

Further analysis using the ClusPlus tool indicated that a minimum of six minutes of recorded data was sufficient for the model to achieve consistent and correct classification, suggesting that prolonged measurements may not be necessary for a reliable assessment.

4.5 Data Visualization and Dimensionality Reduction

The application of PCA for linear dimensionality reduction resulted in limited group separation. The first two principal components collectively explained only 10.6% of the total variance (PC1: 5.7%, PC2: 4.9%), and the confidence ellipses for the healthy and depressed groups showed substantial overlap (Figure 3a). This indicates that linear methods are insufficient to capture the key differences between the groups.

In contrast, the non-linear t-SNE visualization revealed a markedly different picture (Figure 3b). The t-SNE plot showed two distinct, well-separated clusters, with healthy individuals forming a tight, dense cluster and depressed individuals forming a more dispersed one. This clear separation confirms that potent discriminatory information exists within the PRC data, but it is encoded in complex, non-linear patterns that are effectively captured by algorithms like k-NN and Random Forest.

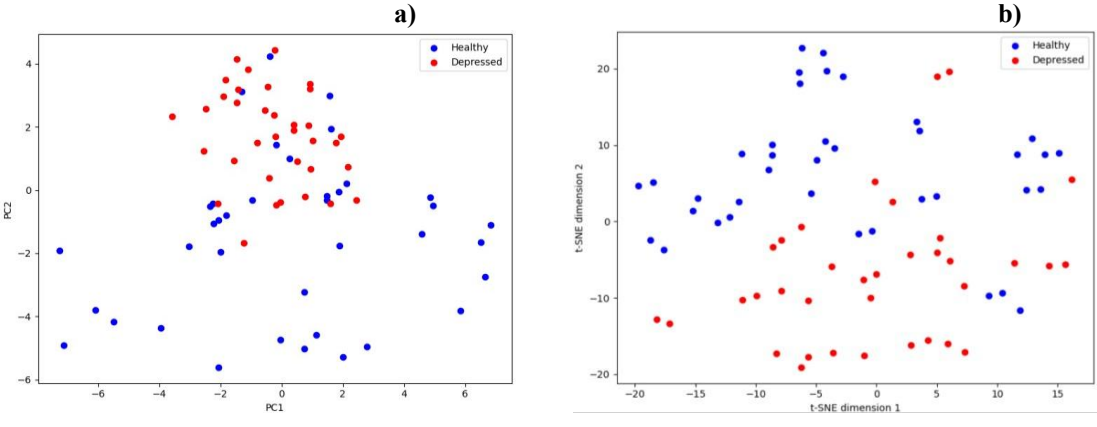

Figure 3. (a) PCA scatter plot showing limited linear separation and significant overlap between healthy (blue) and depressed (red) groups. (b) t-SNE visualization revealing clear, non-linear clustering of the two groups.

5 Discussion

The findings of this study provide compelling preliminary evidence that Pulse-Respiratory Coupling is a robust and sensitive biomarker for Major Depressive Disorder. The observed degradation of PRC in the depressed group aligns with the broader literature on autonomic dysfunction in depression [11, 12]. The loss of synchronous

cardiorespiratory rhythms suggests a state of autonomic incoordination and reduced adaptive capacity, which may underlie some of the somatic and cognitive symptoms of depression [13].

The clinical validity of our group stratification was supported by standardized psychological assessments. The high mean BDI-II score in the depressed group (23 ± 12) confirms a significant level of depressive symptomatology, while the intermediate score in the treatment group (17 ± 13) aligns with expectations of partial symptom remission under pharmacotherapy. These scores provide a clinical anchor to our physiological findings, reinforcing that the observed degradation in PRC corresponds to a clinically recognized depressive state.

The exceptionally high classification accuracy achieved by the k-NN algorithm (97.3%) is noteworthy. The success of k-NN, an instance-based learner, alongside other non-linear models like Random Forest, underscores the complex nature of the physiological disruption in MDD. The poor performance of Logistic Regression further reinforces that these patterns cannot be adequately described by simple linear models. The clear clustering in the t-SNE visualization, despite the low explained variance in PCA, is a powerful visual corroboration of this fact. It demonstrates that machine learning, particularly non-linear methods, can successfully decode these complex physiological signatures.

The partial restoration of PRC patterns in the treatment group is an intriguing finding that warrants further investigation. It suggests that PRC may not only serve as a diagnostic marker but also as a potential biomarker for monitoring treatment efficacy. This aligns with previous work showing that HRV parameters can normalize with successful antidepressant treatment [14]. The observation that age appeared to be a more significant factor in the decision tree models than gender is consistent with known age-related declines in autonomic function [15] and merits exploration in a larger, age-stratified cohort.

From a clinical perspective, the method presented here offers a truly objective and non-invasive approach to supporting depression diagnosis. Unlike subjective questionnaires, physiological measurements like PRC are not susceptible to response bias or varying interpretation. The finding that only six minutes of data are needed for accurate classification enhances the practical feasibility of this method for clinical screening applications.

5.1 Limitations and Future Directions

Despite these promising results, several limitations must be acknowledged. The sample size, while sufficient for this preliminary investigation, is modest. The diagnosis of depression, though supported by standardized scales, was not confirmed by a structured clinical interview in all cases, which is considered the gold standard. Furthermore, the study lacks an independent validation cohort, which is crucial for assessing the generalizability and real-world performance of the model.

Future research must therefore prioritize large-scale, multi-center validation studies that include participants with diverse depression subtypes and comorbidities. It would be highly informative to integrate PRC with other biomarkers, such as HRV and EEG, in a multimodal approach to create a more comprehensive diagnostic profile [16]. Finally, longitudinal studies tracking PRC changes throughout the course of treatment are essential to firmly establish its utility as a dynamic biomarker of therapeutic response.

6 Conclusion

This study demonstrates that Pulse-Respiratory Coupling (PRC) is a potent physiological biomarker for Major Depressive Disorder. Our findings reveal a significant degradation of PRC in depressed individuals, which is quantifiable through machine learning. The high classification accuracy (97.3%) achieved with non-linear algorithms like k-NN, based solely on PRC-derived features, underscores the complex, non-linear nature of autonomic dysregulation in depression.

By moving beyond single-system metrics like HRV to the interplay between cardiac and respiratory systems, we capture a more holistic and sensitive measure of physiological disruption. This work establishes PRC as a promising candidate for the development of accessible, non-invasive, and objective tools to aid in the diagnosis and monitoring of depression, setting the stage for future validation in larger, more diverse clinical cohorts.

7 References

[1] World Health Organization. (2022). World mental health report: transforming mental health for all.

[2] American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorders (5th ed.).

[3] [3] Fava, M. (2003). Diagnosis and definition of treatment-resistant depression. Biological Psychiatry, 53(8), pp. 649-659.

[4] Insel, T., et al. (2010). Research Domain Criteria (RDoC): Toward a new classification framework for research on mental disorders. American Journal of Psychiatry, 167(7), 748-751.

[5] Zorko, A., & Levnajić, Z. (2023). Diagnosing Depression from Physiological Data Using AI. In Proceedings of the ITIS Conference.

[6] Kemp, A. H., et al. (2010). Impact of depression and antidepressant treatment on heart rate variability: a review and meta-analysis. Biological Psychiatry, 67(11), 1067-1074.

[7] Ćukić, M., et al. (2023). When Heart Beats Differently in Depression: Review of Nonlinear Heart Rate Variability Measures. JMIR Mental Health, 10, e40342.

[8] Kralemann, B., et al. (2013). In vivo cardiac phase response curve elucidates human respiratory heart rate variability. Nature Communications, 4, 2418.

[9] Delorme, A. in Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, vol. 134, no. 1, pp. 9–21.

[10] Thayer, J. F. in Lane, R. D., 2000. A model of neurovisceral integration in emotion regulation and dysregulation. *Journal of Affective Disorders*, vol. 61, no. 3, pp. 201–216

[11] Inoue, T., et al., 2019. Heart rate variability in patients with major depressive disorder: a retrospective study.

International Journal of Psychiatry in Clinical Practice, vol. 23, no. 1, pp. 50–57.

[12] Kreibig, S. D., 2010. Autonomic nervous system activity in emotion: A review. Biological Psychology, vol. 84, no. 3, pp. 394–421.

[13] Sano, A. in Picard, R. W., 2013. Stress recognition from wearable sensors. In: *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, pp. 881–886.

[14] Gjoreski, M., et al., 2017. Continuous stress detection using a wrist device: In a day of a student's life. In: *2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*. IEEE, pp. 379–384.

[15] Al-Shargie, F., et al., 2020. Emotion recognition from EEG signals using a developed deep learning model: A pilot study. *IEEE Access*, vol. 8, pp. 116395–116405.

[16] Wang, Y., et al., 2020. A multimodal emotion recognition method based on physiological signals. *Biomedical Signal Processing and Control*, vol. 57, 101732.

[17] SCHMIDT, P., et al., 2018. Unobtrusive physiological sensing for emotion recognition in the wild. In: *Proceedings of the 2018 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. pp. 1237– 1246.

[18] Breiman, L., 2001. Random Forests. Machine Learning, vol. 45, no. 1, pp. 5–32.

[19] Chen, T. in Guestrin, C., 2016. XGBoost: A Scalable Tree Boosting System. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. pp. 785–794.

[20] Zhang, Y., et al., 2021. A review of feature selection methods for biomedical data. *Journal of Biomedical Informatics*, vol. 124, 103942.

[21] Jožef Stefan Institute. (2016). ClowdFlows Data mining workflows on the cloud. http://clowdflows.org/

Extending the Privacy by Design Model to Address NIS 2 Cybersecurity Requirements

Matjaž Drev,
National Institute of Public Health
Trubarjeva 2, 1000 Ljubljana, Slovenia
matjaz.drev@protonmail.com

Abstract: *Organizations are increasingly confronted with regulatory requirements that encompass both personal data protection and cybersecurity. While the GDPR establishes a clear framework for processing personal data, the NIS 2 Directive introduces additional obligations aimed at strengthening cybersecurity resilience. Addressing these combined demands represents a complex legal, organizational, and technical challenge. One way forward is the development of integrated audit frameworks that support systematic compliance assessment across both domains. Building on prior work in which the original privacy by design (PbD) model was developed and empirically tested, this paper proposes an extended model that incorporates NIS 2 requirements. The extended framework aspires to provide a robust and comprehensive instrument for identifying compliance gaps and supporting organizations in adapting more effectively to an increasingly demanding regulatory landscape.*

Key Words: *privacy by design, conceptual model, cybersecurity, NIS 2*

1 Introduction

In the contemporary information society, regulatory compliance has become one of the central challenges for organizations. Within the European Union (EU), the General Data Protection Regulation (GDPR) emphasizes the protection of personal data, not only as a legal obligation but also as an expression of the fundamental right to privacy [3]. One of GDPR's significant contributions was the introduction of the concept of Privacy by Design (PbD), first articulated in Cavoukian's [5] influential essay. As the original concept of PbD was broadly defined and somewhat abstract, Drev and Delak [9] sought to identify building blocks that could provide the basis for a structured framework. This effort culminated in the development of a conceptual PbD model, which was validated through several case studies in healthcare organizations [10].

Parallel to the regulatory emphasis on privacy, the EU legal environment has emphasized the importance of cybersecurity resilience, most notably through the adoption of the Directive (EU) 2022/2555 (NIS 2) [4]. While the GDPR and the PbD model primarily focused on personal data protection, NIS 2 extends the scope toward broader cybersecurity obligations of companies within critical sectors, termed essential and important entities. The directive establishes binding requirements concerning risk management, incident reporting, business continuity, and supply chain security, thereby obliging organizations to adopt more comprehensive and integrated compliance strategies. In this respect, NIS 2 introduces principles closely aligned with the notion of

Secure by Design (SbD), emphasizing that cybersecurity safeguards should be systematically embedded throughout organizational and technical processes [1], much like privacy safeguards in the context of PbD.

This article proposes an extension of the conceptual PbD model, which, in a new iteration, incorporates the requirements introduced by the NIS 2. While empirical validation of the original model has thus far been limited to GDPR related case studies, the extended framework could be similarly tested. However, new case studies should be limited to essential and important entities, as only those are formally required to comply with NIS 2 obligations.

2 Literature review

The original conceptual PbD model [10] was grounded in an extensive review of literature on privacy protection in information systems. Cavoukian's [5] seven foundational principles of PbD provided the initial conceptual framework, while earlier contributions such as Denning's [8] and Chaum's [6] work on Privacy-Enhancing Technologies (PETs) in the 1980s established important technical underpinnings. Subsequent developments, including the Fair Information Privacy Principles of the U.S. Federal Trade Commission, the institutionalization of Privacy Impact Assessments (PIAs) [3], and systematic approaches such as Hoepman's privacy design strategies [7] [13], Foukia et al.'s PISCES [11], Jensen et al.'s STRAP [15], and Kalloniatis et al.'s PriS method [16], were necessary for establishing a theoretical structure for the PbD model. At the European level, the PRIPARE project [24] and standards such as ISO/IEC 27701:2019 further consolidated and formalized best practices for integrating privacy and security into system design.

Alongside privacy protection, cybersecurity legislation started to gain increasing importance [22]. Specifically, the NIS 2 Directive introduced binding obligations for essential and important entities, which include risk analysis, incident reporting, business continuity, and supply chain security [4]. Fortunately, the field of cybersecurity is well developed [1], as many authors contributed to establishing theoretical foundations, sociological [18] [21] and political implications [19], and most importantly, for outlining structured conceptual tools. In this context, NIST with its Cybersecurity Framework [24] or ISO with the widely recognized ISO/IEC 27001:2022 standard should be emphasized. Within the wide body of cybersecurity theory, the concept of Secure by Design (SbD) has been established [1] [25]. It emphasized the importance of proactive integration of security requirements into system architecture [14], building on principles articulated in the software engineering literature [20].

The increasing importance of a systematic approach to cybersecurity challenges and the extensive legal requirements of the NIS 2 directive introduced an opportunity for revising and extending the original PbD model. By combining insights from both privacy and cybersecurity research, the extended PbD model seeks to establish a conceptual tool for addressing both privacy and cybersecurity obligations in a unified manner.

3 Extending the conceptual model

The original PbD model categorized key elements of personal data protection into three sets: legal elements, security elements, and privacy by design and by default elements. This structure reflected the orientation of the GDPR, where legal compliance forms the foundational layer, followed by the technical and organizational measures required to protect data, and finally by mechanisms that integrate privacy considerations directly into system design [25]. While this arrangement provided a robust and coherent model for assessing personal data protection compliance, it did not include the extensive cybersecurity requirements introduced by NIS 2 directive [4].

To address those issues, the conceptual model was revised, guided by two key objectives: (1) to integrate NIS 2 requirements, and (2) to enhance the coherence and usability of the model.

First, the integration of NIS 2 requirements facilitated an update of the model to address contemporary cybersecurity challenges. NIS 2 requires essential and important entities to incorporate policies for risk assessment and information system security, manage cybersecurity incidents, maintain business continuity, secure supply chains, implement cryptographic safeguards, control access rights, and ensure staff competence through training. Those requirements were mapped to the appropriate GDPR elements of the original model. For example, incident management was linked to data integrity which is a part of the security elements group, supply chain security was linked to contractual processing, and cryptography was merged with the broader concept of data confidentiality. Mapping of GDPR and NIS 2 elements can be seen in Table 1. In this way, the general structure of the original model was preserved, while new NIS 2 requirements were also introduced in an unobtrusive way.

Second, the original model was revised in a way that should enable broader usability while still maintaining compact structure. This was done by repositioning and merging existing elements. For example, cross border data transfer was merged with contractual obligations [17]. Encryption, previously separated element, is now part of data confidentiality. On the other hand, a new category of risk analysis and security policies was introduced, as this is a formal requirement of NIS 2, and due to its broad nature cannot be included under more specific security elements [4]. Also, emerging new technologies incited the introduction of privacy enhancing technologies (PET) [2].

The revised model should therefore represent a more comprehensive, yet still compact, conceptual framework that addresses both GDPR and NIS 2 compliance of personal data processing within organizations.

Tabel 1: Mapping of GDPR and NIS 2 elements within model structure

Category	Component / Subcomponent	Purpose / Content	Examples / Mechanisms	NIS 2 Requirement (Art. 21 NIS 2)
Legal elements	Legality of processing	Ensuring processing complies with legal requirements	GDPR legal basis for processing	–
	Individual Rights and Transparency	Enabling individuals to exercise data rights and ensuring transparency of processing	Access, rectification, erasure, restriction of processing, data portability, privacy policy, breach notifications	–

	Contractual Processing and Data Transfers from EU	Governing relationships between controller and processor, ensuring secure data transfer	Contracts, SLAs, subcontractor control, standard contractual clauses	(d) Supply chain security
	DPIA (Data Protection Impact Assessment)	Assessing privacy risks, complements legal framework	Formal template, risk analysis, mitigation recommendations	–
Security elements	Confidentiality (including encryption)	Protecting data from unauthorized access	RBAC, MFA, secure communications, AES/TLS, end-to-end encryption	(i) Human resource security, access management, asset management; (j) MFA, secure communications (h) Cryptography policies, encryption
	Integrity	Ensuring data is not altered without authorization	Checksums, hashes, audit trails	(b) Incident management (e) Security in development, vulnerability management (f) Control effectiveness policies
	Availability	Ensuring data access when needed	Backup, redundancy, disaster recovery	(c) Business continuity, crisis management
	Risk Analysis and Security Policies	Strategic framework for all security components: risk assessment, monitoring, and policies	Security policies, periodic risk assessments, controlling, audits	(a) IS risk and security policies (f) control effectiveness assessment (e) security in development
Privacy by Design & Default elements	PET (Privacy-Enhancing Technologies)	Technological solutions to protect privacy	Anonymization, pseudonymization, secure multiparty computation	–
	Minimization	Limiting data collection and reducing human errors	Filtering, selective processing, data TTL, standard default settings	(g) Basic hygiene practices and training

4 Revising the implementation procedure

The original PbD model, focused on GDPR compliance, was tested on 4 case studies within the Slovenian central health information system (eHealth) [10]. In all cases, the implementation followed the model's five-step procedure: information gathering, analysis of legal, security, and privacy by design and by default elements, and final reporting, which included gap analysis and recommendations for enhancing data processing practices (as seen in Table 2).

These applications demonstrated that the original procedure was robust and practical. Organizations could systematically identify and assess GDPR-related elements, score them effectively, and receive actionable recommendations. However, challenges arose, particularly in evaluating security and privacy by design elements, where GDPR's high-level guidance needed supplementation through ISO/IEC standards to achieve operational clarity [9].

The extended PbD model, incorporating the additional requirements of the NIS 2 Directive, calls for a refined implementation approach. While the five-step methodology remains, two key enhancements are introduced. First, the compliance matrix is updated to reflect NIS 2 obligations, integrating the directive's cybersecurity and risk management expectations (as seen in Table 3). Second, the special questionnaires should be revised to include both GDPR and NIS 2 elements, with clear scoring criteria

that address the combined regulatory framework. These adjustments should ensure the procedure remains systematic and actionable while encompassing the broader compliance landscape.

By retaining the sequential methodology and refining the scoring and questionnaire mechanisms, the revised procedure enables organizations to conduct structured and clear assessments of personal data processing operations.

Tabel 2: Implementation procedure

Sequence of steps	**Step description**
1 Information gathering	Collection of legal documents, internal policies, DPIAs, contracts, and structured interviews. Identification of personal data processing operations, asset registration, data flow mapping. Inclusion of NIS2-related elements such as risk management policies, supply chain security, access controls, and staff competence.
2 Analysis of legal elements	Assessment of legality of processing, transparency, data subject rights, contractual processing, cross-border transfers, and DPIAs. Integration of NIS2 obligations regarding contractual security and supply chain risk management.
3 Analysis of security elements	Evaluation of confidentiality, integrity, and availability of data. Assessment of encryption, access control, incident management, audit measures. Inclusion of NIS2-specific requirements for secure configuration, vulnerability management, cryptography, and staff competence.
4 Analysis of privacy by design and by default elements	Examination of encryption, data minimization, pseudonymization/anonymization, and DPIA alignment. Assessment of PETs and default privacy-preserving settings reflecting GDPR and NIS2 requirements.
5 Final Report with gap analysis and recommendations	Consolidation of findings, gap analysis against benchmarks, and preparation of prioritized recommendations to enhance GDPR and NIS2 compliance.

Table 3: Compliance matrix with revised and extended elements

Processing operation / presence of privacy by design model elements Legend: 1 – element not present 2 – element is present, mayor inadequacy 3 – element is present, minor inadequacy 4 – element is fully present	**Legal elements**				**Security elements**				**Data protection by design and by default**		**Compliance**	
	Legality of processing	Individual rights and transparency	Contractual Processing and Data Transfers from EU	DPIA	Confidentiality	Integrity	Accessibility	Risk Analysis and Security Policies	PET (Privacy-Enhancing Technologies)	Data minimization	Basic (GDPR & NIS 2)	Upgraded (PbD & SbD)
Processing operation 1												
Processing operation 2												
Processing operation 3												

5 Discussion

The extended PbD model provides a structured framework for addressing both privacy and cybersecurity challenges. By incorporating NIS 2 requirements, such as risk analysis, incident management, supply chain security, operational security controls, and

staff competence, among others, the revised PbD model represents a more comprehensive conceptual tool for assessing the compliance of personal data processing operations with both GDPR and NIS 2.

However, the revised model also introduces additional complexity, and due to the lack of new empirical data, it remains unclear whether the implementation and, consequently, its usefulness will prove satisfactory. At this point, a revision of the structured questionnaire used for the interviewing process with managers and related experts is necessary before attempting to test the model on new case studies.

Though not specifically addressed in this paper, there is a whole array of questions regarding the use of AI-based tools in relation to the proposed model. Those tools could be used in restructuring, optimization, and implementation of the model. More importantly, they imply the possibility for automated use of the model in assessing compliance with regulations within organizations.

6 Conclusion

The extended PbD model, integrating NIS 2 requirements, provides a comprehensive conceptual framework for simultaneously addressing privacy and cybersecurity challenges. By combining GDPR privacy requirements with NIS 2 cybersecurity demands, the model increases its usability as a tool for assessing and identifying potential gaps in both privacy and cybersecurity compliance.

From a practical perspective, the model offers organizations a structured approach to designing, assessing, and managing information systems that are resilient, secure, and regulation-compliant. While operationalizing NIS 2 obligations introduces additional complexity compared to GDPR alone, the model's clear mapping of legal, technical, and organizational elements facilitates implementation and provides a foundation for potential automation or semi-automation, for instance, using software or AI-based assessment tools. Careful attention to secure processing, audit trails, and transparency is required to mitigate the inherent risks of automated evaluations.

From a theoretical perspective, the extended PbD model contributes to the conceptual understanding of how privacy by design principles can be integrated with broader cybersecurity requirements. It establishes a unified framework that highlights the interdependencies between privacy, security, and risk management, offering a direction for future research and empirical validation across diverse organizational contexts.

Overall, the extended PbD model enables organizations to assess their level of compliance with both GDPR and NIS 2 requirements, which emphasizes the interconnectedness of data privacy and data security. Adopting a more holistic approach could benefit organizations not only by improving their level of compliance, but also by guiding them to build more resilient, secure, and privacy-friendly information systems.

6 References

[1]. Anderson, R. (2020). Security engineering: A guide to building dependable

distributed systems (3rd ed.). John Wiley & Sons. ISBN: 978-1-119-64278-7.
[2]. Balboni, P.; Bella, G.; Capparelli, F.; Taborda Barata, M. Chapter 54 - Privacy-Enhancing Technologies. In: Computer and Information Security Handbook (Fourth Edition), 891-905, 2025.
[3]. Bellamy, F. U.S. data privacy laws to enter new era in 2023. https://www.reuters.com/legal/legalindustry/us-data-privacy-laws-enter-new-era-2023-2023-01-12/, downloaded: May 16th 2023.
[4]. Castiglione, G.; Santamaria, D. F.; Bella, G.; Brisindi, L.; Puccia, G. Guiding cybersecurity compliance: An ontology for the NIS 2 directive. Computers & Security, 157:104617, 2025.
[5]. Cavoukian, A. Privacy by Design. The 7 Foundational Principles, https://www.ipc.on.ca/wp-content/uploads/Resources/7foundationalprinciples.pdf. 2009.
[6]. Chaum, D. Security without Identification Card Computers to make Big Brother Obsolete. Communications of the ACM 28(10): 1030-1044,1985.
[7]. Colesky, M.; Hoepman, J. H.; Hillen, C. A Critical Analysis of Privacy Design Strategies. In Proceedings - 2016 IEEE Symposium on Security and Privacy Workshops, pages 33–40, San Jose, California, 2016.
[8]. Denning, D.E. Cryptography and Data Security. Addison-Wesley Publishing Company, Boston, USA, 1982.
[9]. Drev, M.; Delak, B. Conceptual Model of Privacy by Design. Journal of Computer Information Systems, 62(5), 888-895, 2021.
[10]. Drev, M.; Stanimirovič, D.; Delak, B. Implementation of privacy by design model to an eHealth information system. Online Journal of Applied Knowledge Management, 10(1):77-87, 2022.
[11]. Foukia, N.; Billard, D.; Solana, E. PISCES: A Framework for Privacy by Design in IoT. In Proceedings of the 2016 14th Annual Conference on Privacy, Security and Trust (PST), Auckland, New Zealand, 2016.
[12]. Gurses, S., Troncoso, C., Diaz, C. Engineering Privacy by Design, Available from https://www.esat.kuleuven.be/cosic/publications/article-1542.pdf, downloaded: July 24th 2019.
[13]. Hoepman, J. H. Privacy Design Strategies. IFIP International Information Security Conference, pages 446–459, Berlin, Germany, 2014.
[14]. Huth, D.; Matthes, F. Appropriate Technical and Organizational Measures: Identifying Privacy Engineering Approaches to Meet GDPR Requirements. AMCIS 2019 Proceedings, pages 1790-1799, Cancun, Mexico, 2019.
[15]. Jensen, C.; Tullio, J.; Potts, C.; Mynatt, E. D. STRAP: A Structured Analysis Framework for Privacy, https://www.academia.edu/62138420/Strap_A_structured_analysis_framework_for_privacy. 2005.
[16]. Kalloniatis, C.; Kavakli, E.; Gritzalis, S. Addressing Privacy Requirements in System Design: The PriS Method. Requirements Engineering, 13(3):241-255, 2008.
[17]. Kurtz, C.; Semmann, M.; Böhmann, T. Privacy by Design to Comply with GDPR: A Review on Third-Party Data Processors, https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1401&context=amcis2018. 2018.
[18]. Laiz-Ibanez, H.; Mendaña-Cuervo, C.; Carus Candas, J. L. The metaverse: Privacy and information security risks. International Journal of Information Management Data Insights, 5(2):100373, 2025.

[19]. Lin, H. (2009). *Cybersecurity and national policy*. Washington, DC: Brookings Institution Press. ISBN: 978-0-8157-2549-7.
[20]. Mouratidis, H.; Giorgini, P.; Manson, G. When security meets software engineering: a case of modelling secure information systems. Information Systems, 30(8):609-629, 2005.
[21]. Nye, J. S. (2011). *The future of power*. New York, NY: PublicAffairs. ISBN: 978-1-58648-642-1.
[22]. Papakonstantinou, V. Cybersecurity as praxis and as a state: The EU law path towards acknowledgement of a new right to cybersecurity? Computer Law & Security Review, 44:105653, 2022.
[23]. Pascoe, C., Quinn, S., & Scarfone, K. (2024). *The NIST Cybersecurity Framework (CSF) 2.0*. National Institute of Standards and Technology. https://doi.org/10.6028/NIST.CSWP.29
[24]. PRIPRE Project. PRIPARE Handbook - Privacy and Security by Design Methodology, http://pripareproject.eu/wp-content/uploads/2013/11/PRIPARE-Methodology-Handbook-Final-Feb-24-2016.pdf. 2016.
[25]. Ruohonen, J. SoK: The design paradigm of safe and secure defaults. Journal of Information Security and Applications, 90:103989, 2025.
[26]. Semantha, F.H.; Azam, S.; Yeo, K.C.; Shanmugam, B. A Systematic Literature Review on Privacy by Design in the Healthcare Sector. Electronics, 9(3):452, 2020.
[27]. Spiekermann, S.; Cranor, L. F. Engineering privacy. IEEE Transactions on Software Engineering, 35(1), 67–82, 2009.

Cybersecurity Auditing

Boštjan Delak, Matjaž Drev
Faculty of Information Studies
Ljubljanska cesta 31A, 8000 Novo mesto, Slovenia
bostjan.delak@fis.unm.si, matjaz.drev@fis.unm.si

Abstract: *Cybersecurity is becoming increasingly important for any organization. Nowadays, most management is concerned about cybersecurity. It is especially a big concern in the European Union, as the NIS2 directive foresees their responsibility and effective risk management. Cybersecurity audits are essential for assessing the effectiveness of an organization's security measures, identifying vulnerabilities, and ensuring compliance with industry standards and regulations. By conducting regular cybersecurity audits, organizations can demonstrate to their customers that their security is being taken seriously. As cybersecurity audit reports are mostly classified as confidential, they are not easily accessible on the World Wide Web. Exceptions are audit reports carried out by the Courts of Audits of each country. The article presents new approaches for auditing with the help of artificial intelligence and auditing cyber risks. Based on some cybersecurity audit reports that are publicly available online, it verifies the application of these approaches.*

Key Words*: Cybersecurity, cybersecurity audit, auditors, audit reports*

1 Introduction

Cybersecurity and digital autonomy have become a subject of strategic importance for the EU and its Member States and, as the threat level rises, we must step up our efforts to protect our critical information systems and digital infrastructures against cyber-attacks. Cybersecurity not only concerns our utilities, defense, or health systems, it is also about protecting our personal data, business models and intellectual property [1]. Cybersecurity is the practice of ensuring protection within the cyber domain—specifically, safeguarding the confidentiality, integrity, and availability of individual and organizational data against malicious activities carried out by unauthorized threat actors.

Organizations, on daily basis, need an independent assessment of their cyber-threat preparedness, identification of vulnerabilities and information security deficiencies, and recommendations for mitigating these risks. This can be done by information systems auditors as part of a cybersecurity audit. Information systems auditors use various methodologies and standards in these audits (such as NIST CSF 2.0, ISO/27001:2022, ISO/IEC 27032:2023, COBIT 2019 and others). The European Union's (EU) NIS2 Directive is set to transform the cybersecurity landscape across Europe. This landmark legislation aims to unify the cybersecurity practices among the member states of the EU, thus increasing their ability to prevent incidents and minimize their impact when they

occur. For the public, this means a more resilient digital infrastructure, safeguarding essential services across an extended list of industries [2].

The article's main objective is to present additional approaches at cybersecurity audits. The article is organized as follows: In the second section, information about audits and auditing is given. In the third section information about cybersecurity and cybersecurity auditing is presented. The fourth section presents motivation of our research and research question. The fifth section presents literature review about cybersecurity audits. The core of the article outlines review of some cybersecurity audit reports in the sixth section. The seventh section forms a discussion. And the eighth section presents the conclusion, research limitations, and further work.

2 Auditing

The word "audit" comes from the Latin word *audire*, meaning "to hear". Flint [3] explains that the audit function has evolved in response to a perceived need of individuals or groups in society who seek information or reassurance about the conduct or performance of others in which they have an acknowledged and legitimate interest. Auditing in the form of basic checking activities was found in the ancient civilizations of China, Egypt, and Greece [4]. In mid-1980 was developed risk based auditing [4], which is an audit approach where an auditor is focusing on those areas which are more likely to contain threats (human errors, inadequate and insufficient controls, vulnerabilities).

The assessment of information systems (IS) by auditors has generated the term "IS auditing", which is the evaluation of IS practices and operations to assure the integrity of an entity's information. Such evaluation can include assessment of efficiency, effectiveness, and economy of computer-based practices [5]. Recently, more and more IS auditing is based on information security and/or cybersecurity. There are three different audit types: internal auditing, provided by internal employees; external auditing, provided by external companies; internal auditing, in cooperation with external (outsourced) auditors.

3 Cybersecurity and cybersecurity auditing

There are several definitions of cybersecurity. One of them is: *Cybersecurity means the activities necessary to protect network and information systems, the users of such systems, and other persons affected by cyber threats* [6].

EU by its regulation created European agency ENISA –*European Union Agency for Cybersecurity*. ENISA is dedicated to achieving a high common level of cybersecurity across Europe and contributes to EU cyber policy, enhances the trustworthiness of IT products, services and processes with cybersecurity certification schemes, cooperates with Member States and EU bodies, and helps Europe prepare for the cyber challenges of tomorrow [7].

ENISA identified seven prime cybersecurity threats landscape in 2024, which are: Ransomware, Malware, Social Engineering, Threats against data, Threats against availability: Denial of Service, Information manipulation and interference, and Supply chain attacks [8]. All of the above threats, if realized, have major financial impacts on organizations. Direct costs associated with cyber incidents encompass the tangible financial losses, damages, and hardships endured by victims following such events. Indirect costs from cyber incidents are difficult to quantify and are often ignored; although, they can have important long-term impacts on organizations and the economy [9].

Cybersecurity audits ensure that organizations are well prepared to defend themselves against internal and external cyberthreats. The objective of a cybersecurity audit is to provide management with an evaluation of the effectiveness of cybersecurity processes, policies, procedures, governance and other controls. The review focuses on cybersecurity standards, guidelines and procedures as well as the implementation of these controls. The audit/assurance review relies upon other operational audits of the incident management process, configuration management and security of networks and servers, security management and awareness, business continuity management, information security management, governance and management practices of both IT and the business units, and relationships with third parties [10].

Cybersecurity auditing is vital for any organization to achieve six business objectives: identify and mitigate the risk, protect sensitive information, comply with regulations, improve security posture, gain customer confidence, and maintain business continuity [11].

4 Motivation for our research

The research question was: How are new cyber security auditing approaches implemented in practice?

The aim of this study is to examine whether national audit organizations have applied new approaches in their cybersecurity audits – the use of big language models and artificial intelligence in cybersecurity audits, as well as cybersecurity risk audits.

5 Literature review

Within the last ten years there have been published a lot of articles exploring cybersecurity. ISACA[1] auditor guidelines [12] and a detailed audit program based on the NIST cybersecurity framework [13] for IS auditors. Cooke [14] presented the ISACA audit process and give a brief review of cybersecurity vulnerabilities, threats and risks; he also explained cybersecurity goals and related audit objectives. Clark & Charles [15] presented key steps in Cybersecurity Audit preparation and requested documentation. They also presented common challenges at such an audit.

[1] URL: https://www.isaca.org/.

Vuko et al. [16] analyzed and presented which factors explain the effectiveness of internal audit in providing assurance about cybersecurity risk management. Ferrerira et al. others [17] identified ways in which internal audits could contribute to ensuring the effectiveness and improvement of cybersecurity controls. They explained that internal audit should not focus only on high-level management and governance aspects, but also conduct detailed assessments of operational activities.

Munster [18] explained that in the Voice of CISO report some 60% of CISO respondents claimed staff are their greatest corporate cybersecurity risk. He also mentioned that nearly a third of UK organizations (30%) still lack dedicated insider risk resources.

CSAM – Cybersecurity Audit Model with the main goal to improve Cybersecurity assurance was presented by Sabillon et al. [19]. CSAM has been designed to address the limitations and inexistence of cybersecurity controls to conduct comprehensive cybersecurity or domain-specific cybersecurity audits. CSAM was validated also at three Canadian higher education. The key findings indicate that tailored cybersecurity strategies, when supported by continuous and comprehensive auditing, significantly mitigate cyber risks in higher education environments [20].

Ilori et al. [21] critically analyze emerging technologies for cyber security auditing and different standards, frameworks, and methodologies. As manual audit processes and static compliance checks can no longer keep pace with the dynamic nature of today's digital environments. To address this gap, organizations are turning to the combined power of Artificial Intelligence (AI) and blockchain technology to transform how cybersecurity audits and compliance management are executed [22]. Miracle & Grace [23] explored how machine learning can transform cybersecurity audits, focusing on its role in threat detection, risk management, and compliance monitoring.

Recently a new field appeared in cybersecurity – cybersecurity risk audits. The cybersecurity risk audit, as part of the monitoring and review of a cybersecurity risk management system, refers to the countermeasures established during the risk treatment stage [24] presented guidelines to perform such an audit, which consist of seven steps and 28 activities. Doris & Rosinski explained that the frequency and sophistication of cyberattacks continue to increase, so organizations must adapt and innovate their audit processes to remain resilient in the face of emerging threats. The findings underscore the growing importance of cybersecurity audits, not only for ensuring compliance but also for proactively identifying vulnerabilities and addressing threats before they escalate into serious breaches [25].

6 Cybersecurity audits reports review

As part of short research, we attempted to examine whether new approaches were being used in cybersecurity audits - the use of large language models and AI, and auditing cybersecurity risks related to our research question.

Most audits, information security audit and also cybersecurity audits commissioned by organizations and conducted by internal or external auditors are marked confidential and

not publicly released, as these releases could expose their vulnerabilities and allow attackers to launch a successful cyberattack. Exception are audits carried out by Courts of Audits –CoA (in some countries these are also named National Audit Organizations – NAO, or Offices of Auditor General – OoAG), these reports are mostly not marked confidential. **Napaka! Vira sklicevanja ni bilo mogoče najti.** presents some audit reports made by CoA regarding cybersecurity.

Table 1: Some audit reports

#	Audit title	Issued by	Published
A	Government cyber resilience	NAO United Kingdom	2025
B	Combatting Cybercrime	OoAG Canada	2024
C	The strengths and weaknesses of NAFIN, the digital armed forces network	CoA Netherlands	2024
D	Management of Cyber Security Incidents	NAO Australia	2024
E	Government control of national information and cyber security	NAO Sweden	2023
F	Information security at higher education institutions	NAO Sweden	2023
G	Audit of Cybersecurity Resiliency at the Colorado Governor's Office of Information Technology	Colorado State Auditor USA	2023
H	Effectiveness of cybersecurity management for critical infrastructure in ELES	CoA Slovenia[2]	2021
I	The effectiveness of ensuring cybersecurity in the Republic of Slovenia	CoA Slovenia[10]	2021
J	Cyber security of border controls operated by Dutch border guards at Amsterdam Schiphol Airport	CoA Netherlands	2020
K	Cyber security and critical water structures	CoA Netherlands	2019
L	Progress of the 2016–2021 National Cyber Security Program	NAO United Kingdom	2019
M	Information Technology Audit: Cyber Security across Government Entities	NAO Malta	2017

Summary of audit reports in alphabetical order by country

Below is a brief description of the revision and at the end of the description in parentheses is a link to the revision in the table 1.

NAO Australia's objective of this audit was to assess the effectiveness of the selected entities' implementation of arrangements for managing cyber security incidents in accordance with the Protective Security Policy Framework (PSPF) and relevant ASD Cyber Security Guidelines (**D**). OoAG Canada's audit focused on whether the Royal

[2] In Slovene language

Canadian Mounted Police and selected federal entities had the capacity and capability to effectively enforce laws against cybercrime and to ensure the safety and security of Canadians (**B**). NAO Malta embarked on a horizontal audit to compare the level of adoption of selected cybersecurity controls across selected auditee sites. The horizontal audit was conducted across 10 different Government entities (**M**). CoA Netherlands audit in 2019 describes the audit the way in which the Minister of Infrastructure and Water Management has prepared to deal with cyber attacks mounted against the critical water structures managed on her behalf by the Directorate-General for Public Works and Water Management (**K**). In 2020 they audited the measures taken by the Minister of Defence and the Minister of Justice and Security to protect the IT systems supporting the border controls operated by the border guards at Amsterdam Schiphol Airport against cyber attacks (**J**) and in 2024 they audited concerns a communication network of vital importance to central government in general and the Ministry of Defence in particular: the Netherlands Armed Forces Integrated Network (NAFIN) (**C**). NAO Sweden published two cybersecurity audits: audit regarding whether the Government's efforts to strengthen Sweden's national information and cyber security have been efficient (**E**), and audit of Information security at higher education institutions – management of research data requiring protection, the topic of information security surrounding research data has been in the spotlight (**F**). CaO Slovenia published in 2021 two audit reports: an audit of the effectiveness of the Government of the Republic of Slovenia, the Office of the Government of the Republic of Slovenia for the Protection of Classified Information, and the Ministry of Public Administration in ensuring the cybersecurity of the Republic of Slovenia (**H**), and audit to assess the cyberthreat preparedness of an organization that operates critical infrastructure for electric power transmission (**I**). OoAG USA in state Colorado audited how is cybersecurity resilience among government bodies (**G**).

Above-mentioned Court of Audit reports covered specific domains of cybersecurity objectives. The results of our research showed that auditors have not used machine learning or AI as a tool to support their audits, which has been recommended by several authors. Also, none of the reports concentrate on cybersecurity risk audits, which have recently been recommended by Sanchez-Garcia [24].

7 Discussion

AI has significantly accelerated the pace of change across all segments of the business technology landscape, including the audit and assurance world. For IT auditors, and for auditors in general, AI and related technologies such as Large Language Models offers tremendous value. This includes areas such as automation of audit processes, risk identification, controls testing, continuous auditing, and more [26]. Similar conclusions were presented by Hanry and Iqbal [22]. Although AI was not used in the audit reports reviewed, it offers potential value that will be demonstrated in future cybersecurity audits.

The audit areas presented by ISACA within the framework of the Cybersecurity Information Systems Audit/Assurance Program [10] and already explained in this article will be characterized in the future by an additional and independent audit of cybersecurity risks, as presented by Sanchez-Garcia et al. [24].

There are some limitations of this paper. First, the sample of Courts of Audit reports was small and covered some of the English written reports published on the Court of Audit websites. Second, only Courts of Audit reports are publicly accessible. Third Court of Audit in Slovenia has made until now only 2 cybersecurity audits.

There is still much space for further research in this area**.** At the Faculty of Information Studies, a master's student is working on her final master's degree thesis, where she will survey all certified IS auditors in Slovenia about their experiences in the field of cybersecurity audits. The results will be presented at the ITIS 2026 conference. Another possibility for further research is to examine all the reports of the Courts of Auditors in EU or even on a wider scale and conduct an analysis of the objectives, criteria, and findings of individual reports. Such research would require a lot of resources. A further research opportunity is to identify cybersecurity audits that will use AI or ML in their implementation and analyze the results, advantages, and disadvantages. As it can be seen, there is still a lot of potential for further research in this area.

8 Conclusion

Cybersecurity audit could help organizations' owners and top management to identify vulnerabilities and recommend measures on improving their security. Regular penetration tests and cybersecurity audits with specific objectives can mitigate some cybersecurity threats and risks, but they are not enough to avoid them completely. In the future AI and large language models show potential to help IS auditors to deliver cybersecurity audits more effectively and efficiently.

9 References

[1]. EUROSAI, Cybersecurity in the EU and its Member States, EUROSAI, 2020.

[2]. Dimitriadis, C. NIS2 Directive: Strengthening Public Cybersecurity and Infrastructure Across Europe, https://www.isaca.org/resources/news-and-trends/isaca-now-blog/2024/nis2-directive-strengthening-public-cybersecurity-and-infrastructure-across-europe?gad_source=1&gad_campaignid=21147302543&gclid=EAIaIQobChMIoNz7o6yrkQMVrJCDBx3dCSh9EAAYAiAAEgLDV_D_BwE, accessed on: September 3rd 2025.

[3]. Flint, D. Philosophy and principles of auditing. Hampshire: Macmillan Education Ltd. 1988.

[4]. Teck-Heang, L., Ali, M.A. The evolution of auditing： An analysis of the historical development, Journal of Modern Accounting and Auditing, 4, 12, 2008.

[5]. Senft, S., Gallegos, F., Davis Aleksandra. Information Technology Control and Audit, Fourth edition, CRC Press, Boca Raton, USA, 2013.

[6]. EU, Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of

cybersecurity across the Union, amending Regulation (EU) No 910/2014 and Directive (EU) 2018/1972, and repealing Directive (EU) 2016/1148 (NIS 2 Directive), (NIS2), EU, 2022.
[7]. ENISA, NIS2 Technical Implementation Guidance, ENISA, Greece, 2025.
[8]. ENISA, ENISA Threat Landscape 2024, ENISA, 2024.
[9]. Vergara Cobos, E. &Cakir, S. A Review of the Economic Costs of Cyber Incidents. Washington, DC: World Bank, 2024.
[10]. ISACA. IS Audit/Assurance Program Cybersecurity: Based on the NIST Cybersecurity Framework, ISACA, 2016.
[11]. Azab, O. Six Benefits of a Cybersecurity Audit (and 6 Steps to Perform One), https://www.isaca.org/resources/news-and-trends/industry-news/2024/six-benefits-of-a-cybersecurity-audit, downloaded: September 3rd 2025.
[12]. ISACA, CSX, European Cybersecurity Audit/Assurance Program, ISACA, 2014.
[13]. ISACA, ISACA Cybersecurity Audit Program: Based on the NIST Cybersecurity Framework 2.0, ISACA; 2024.
[14]. Cooke, I. IS Audit basic, Auditing Cybersecurity, ISACA Journal, Vol. 2, 2019.
[15]. Clarke, A. & Charles, J. Cybersecurity Audits Preparing for Success, https://www.researchgate.net/publication/388459058_Cybersecurity_Audits_Preparing_for_Success , downloaded: September 3rd 2025.
[16]. Vuko, T., Slapničar, S., Čular, M. & Drašček, M. Key Drivers of Cybersecurity Audit Effectiveness: a neo-institutional perspective, International Journal of Auditing, 29:188-206, 2025.
[17]. Ferreira, L.V.A.; Alves, C.A.d.M.; Peotta de Melo, L.; Nunes, R.R. Internal Audit Strategies for Assessing Cybersecurity Controls in the Brazilian Financial Institutions. Appl. Sci. 2025, 15, 5715.
[18]. Munster, P. Insider Threats Surge: What CISOs Must Know to Protect Their Organizations, https://www.infosecurity-magazine.com/news-features/insider-threats-what-cisos-must/, downloaded: September 8th 2025.
[19]. Sabillon, R., Serra-Ruiz, J., Cavaller, V & Cano, Jeimy, J.M. A Comprehensive Cybersecurity Audit Model to Improve Cybersecurity Assurance: The Cybersecurity Audit Model (CSAM), 2nd International Conference on Information Systems and Computer Science - INCISCOS 2017.
[20]. Sabillon, R.; Higuera, J.R.B.; Cano, J.; Higuera, J.B.; Montalvo, J.A.S. Assessing the Effectiveness of Cyber Domain Controls When Conducting Cybersecurity Audits: Insights from Higher Education Institutions in Canada. Electronics 2024, 13, 3257.
[21]. Ilori, O., Lawal, C.I., Friday, S.C., Isibor, N.J & Chukwuma-Eke, .C. Cybersecurity Auditing in the Digital Age: A Review of Methodologies and Regulatory Implications, Journal of Frontiers in Multidisciplinary Research, 3(1): 174-187, 2022.
[22]. Henry, G. & Iqbal, J, Harnessing AI and Blockchain to Automate Cybersecurity Audits and Strengthen SOC Compliance, December 2023, https://www.researchgate.net/publication/393591152_Harnessing_AI_and_Blockchain_to_Automate_Cybersecurity_Audits_and_Strengthen_SOC_Compliance , downloaded September 3rd 2025.
[23]. Miracle, A. & Grace, W. Transforming Cybersecurity Audits: The Role of

Machine Learning in Enhancing Compliance and Threat Mitigation, Journal of Machine Learning, 8(7):12-18, 2023.

[24]. Sanchez-Garcia, I. D., Gilabart, T.S.F. & Calvo-Manzano, J.A CRAG: A Guideline to Perform a Cybersecurity Risk Audits, WITCOM 2023, CCIS 1906, 517–532, 2023.

[25]. Doris, L. & Rosinski, J. Reviewing the Processes for Conducting Effective Cybersecurity Audits, Including Threat Identification and Response, https://www.researchgate.net/publication/386342884_REVIEWING_THE_PROCESSES_FOR_CONDUCTING_EFFECTIVE_CYBERSECURITY_AUDITS_INCLUDING_THREAT_IDENTIFICATION_AND_RESPONSE#fullTextFileContent, downloaded: September 2nd 2025.

[26]. Jaleel, A. Five Ways that IT Auditors Can Put AI to Good Use, https://www.isaca.org/resources/news-and-trends/isaca-now-blog/2025/five-ways-that-it-auditors-can-put-ai-to-good-use , accessed on: September 3rd 2025.

Collaborative Decision-Making for Cybersecurity

Andrej Bregar
Informatika d.o.o., Vetrinjska ulica 2, 2000 Maribor, Slovenia
andrej.bregar@informatika.si

Abstract: *This paper provides a systematic analysis of different aspects, methodologies, and technologies for stakeholder collaboration in cybersecurity. These aspects include collective situational awareness, CTI (Cyber Threat Intelligence) sharing and utilization, standardization and exchange of incident response playbooks, community building, and group decision-making involving multiple organizational levels. The presented research study incorporates collaborative mechanisms and multi-criteria decision analysis into an advanced decision process, facilitating the proactive selection of mitigation measures against cyber threats and attacks. The efficiency of the process is demonstrated through its application to an electricity distribution network with a SCADA (Supervisory Control and Data Acquisition) layer. The proposed approach allows organizations to improve their proactive and reactive cybersecurity strategies, enhance operational and strategic decision-making, and leverage the cybersecurity posture of corporate ecosystems.*

Key Words: *Group Decision Processes, Collaboration Technologies, Cybersecurity.*

1 Introduction

The growing importance of cybersecurity in IT (Information Technology) systems, OT (Operational Technology) networks, critical infrastructures, enterprises, society, and everyday life is driven by intensifying digitalization, sophisticated cyber-attacks, and geopolitical risks. According to the World Economic Forum, cybersecurity is an essential resource for information societies that fosters development, accelerates progress, and harnesses the potential of digital technologies to deliver outcomes beneficial to society [35]. It enables critical infrastructures and services to work, allows people to perform their daily routines, and can favor socially acceptable outcomes of digital technologies.

Yet, cybersecurity introduces many challenges. A big issue in corporate environments relates to the cascading effects of cyber-attacks that can compromise dependent assets operated by several stakeholders. Moreover, to efficiently manage cyber threats and facilitate strategic decision-making, experts must be provided with rich information and encouraged to share knowledge and procedures for cyber-attack detection and response. Therefore, the focus moves increasingly to the collaborative aspects of cybersecurity, which include CTI (Cyber Threat Intelligence) sharing, adoption of group platforms, standardization of incident response procedures, establishment of communities, and (cross)organizational group decision-making. Several past studies about organizational decision-making for cybersecurity confirm this assumption [1, 20, 28, 33, 36].

This paper systematically addresses the methodologies for collaboration and group decision-making in cybersecurity. It discusses their application and presents some results of the Horizon 2020 CyberSEAS research project [10] dealing with CTI exchange and cooperative incident response management. It also elaborates on the group multi-criteria

decision-making process for cyber-attack mitigation, which we introduced in the scope of our preceding research work [8]. The paper extends this process and applies it in real-life settings to demonstrate its usability and efficiency.

2 Organizational Decision-Making

In cybersecurity, collaboration between different levels of the organization is essential to ensure decisions are made with input from those affected and to build support for the chosen course of action. NIST (National Institute of Standards and Technology) suggests a multi-tiered risk management approach, encompassing organization, business process, and information system levels [31]. This approach is adopted by the group decision-making process for cyber-attack mitigation presented in Section 4. It is generalized in Figure 1, which shows the competencies at different levels and the decision-making roles that participate. All organizational levels must collaborate coherently in the decision-making process to cover complementary aspects. In this way, cybersecurity analysts evaluate technical requirements at the information system level. Their evaluations are combined further with the assessments at the business process level, where business analysts and executive staff consider financial, reputational, operational, and risk tolerance requirements, as well as organizational metrics, including RTO (Recovery Time Objective) and RPO (Recovery Point Objective).

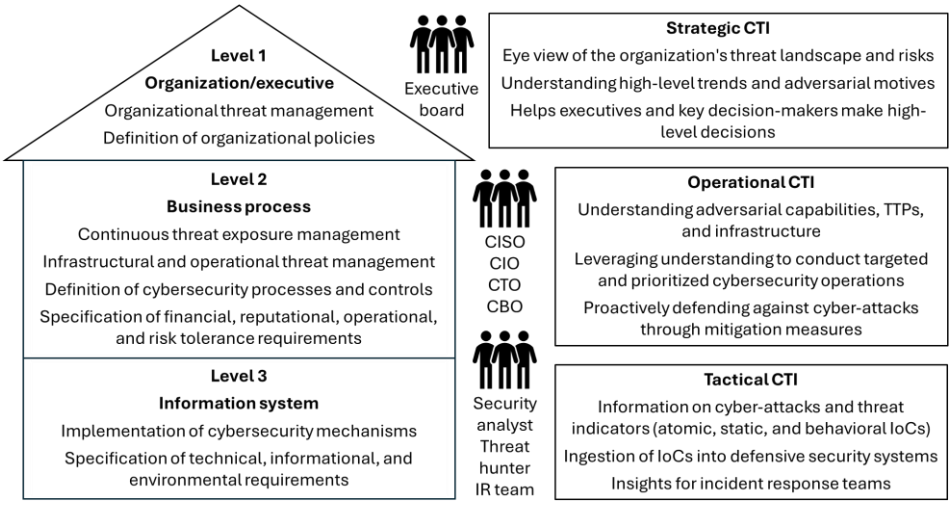

Figure 1: A multi-tiered approach to organizational decision-making leveraging CTI.

The organization and business process levels support cybersecurity risk management and continuous threat exposure management, two of Gartner's top cybersecurity trends [15]. The first focuses proactively on resilience investments. The latter allows organizations to continuously evaluate accessibility, exposure, and exploitability of digital and physical assets to highlight vulnerabilities and threats by aligning assessment and remediation with threat vectors. Organizations can utilize CTI to achieve this. Several categories of CTI are employed to provide insights at various abstraction levels, as presented in Figure 1. Individual CTI categories are therefore suitable for different types of decision-makers in alignment with corresponding tiers of the organization.

ISO/IEC 27002:2022 [19] and NIST [21] consider CTI an important collaborative effort. CTI shared within communities can be beneficial because member organizations often face actors that use common TTPs (Tactics, Techniques, and Procedures) targeting the same or similar types of systems and information. Hence, cyber defense is most effective when organizations cooperate to defend against threat actors, involving the interaction of stakeholders possessing diverse knowledge, skills, and tools [2, 4]. Figure 2 shows the collaborative process of decision-making and CTI sharing and utilization.

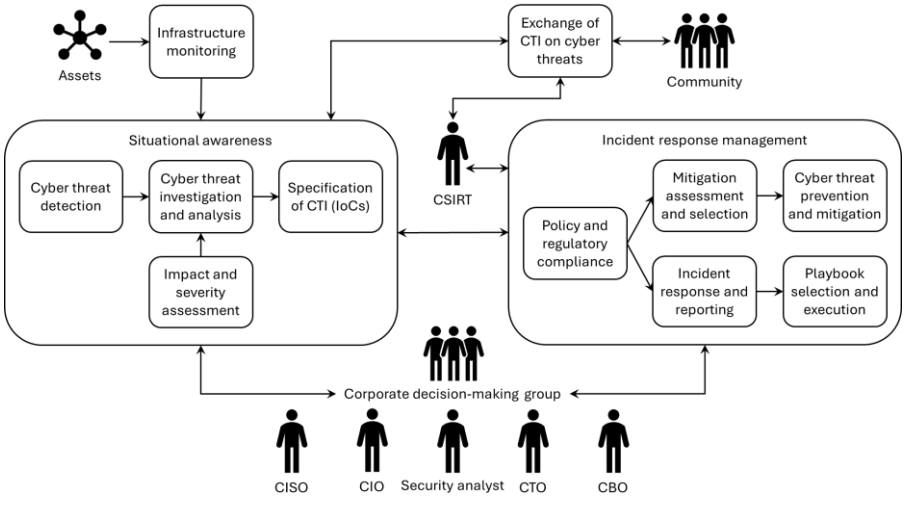

Figure 2: A collaborative process for decision-making, CTI sharing, and CTI utilization.

The idea behind CTI exchange is to create situational awareness among stakeholders regarding the newest threats and vulnerabilities, which allows for swift and proactive implementation of remediation strategies and actions [34]. Situational awareness helps organizations obtain and integrate relevant cybersecurity information [18], enabling them to leverage the analytic capabilities, knowledge, and experience of sharing partners within a community of interest, thereby enhancing the defensive capabilities of multiple parties [17]. Franke et al. [14] argue that situational awareness is best understood by combining three complementary perspectives: socio-cognitive, organizational, and technological. This aligns with the three organizational levels we propose in Figure 1.

3 Coordination and Cooperation for Incident Response

Collaborative activities in cybersecurity are recommended by several standards and required by national and international legislation. In the European Union, the NIS 2 Directive [12], CER (Critical Entities Resilience) Directive [13], and NCCS (Network Code on Cybersecurity) [11] are of particular significance, especially in essential services and critical infrastructures, such as banking, healthcare, and the energy sector. To adhere to the regulations, it is necessary to introduce standardized collaborative mechanisms targeting 1) required coordination of SOCs (Security Operation Centers) and CERTs (Computer Emergency Response Teams), 2) common incident response procedures and rules, 3) communication and information-sharing strategies, and 4) unified data formats, structures, and tools for collaboration and reporting. These mechanisms should be applied in response to detected cyber incidents.

Figure 3 gives a generic representation of coordination and cooperation to facilitate incident response management and proactive cyber-attack mitigation. Incident response adheres to the regulatory requirements and is operationalized with playbooks, which are collectively modeled and shared in the community. Situational awareness and CTI exchange between SOCs and CERTs allow for continuous enhancements of incident response based on operational feedback and newly detected cyber threats. CTI also enables proactive cyber-attack mitigation and protection of corporate assets. An important activity is multi-criteria group decision-making. Its purpose is twofold. Firstly, it allows SOCs to evaluate the efficiency of countermeasures. Secondly, it is applied for incident impact assessments to direct incident response and reporting. According to legislation, reporting rules depend on the severity of incidents [29, 30]. Incident response actions are also triggered according to the scale of incidents and the damage they cause. Only high-impact incidents require swift response, extensive coordination, and strict reporting.

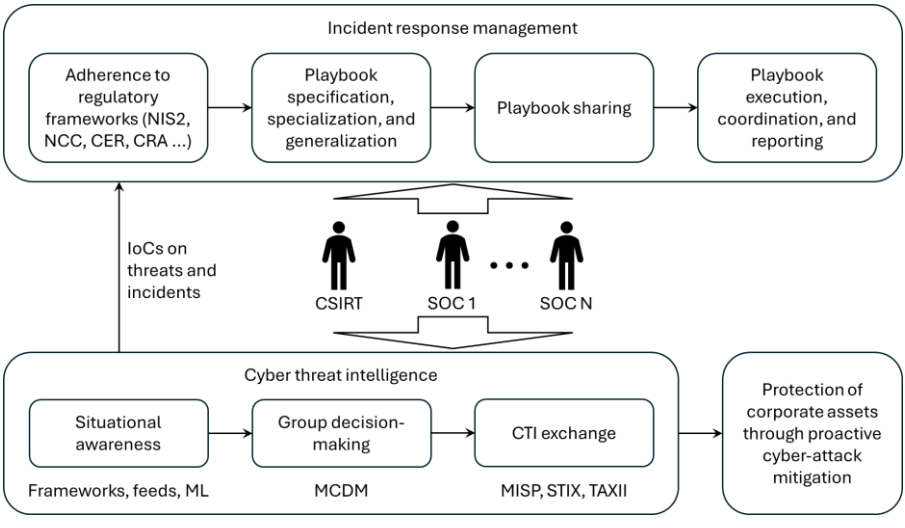

Figure 3: Cooperation for incident response and cyber-attack mitigation.

Reporting and coordination activities should be integrated into standardized incident response procedures represented as playbooks and aligned with the incident response life cycle [26]. The structured approach to playbook management is a collective effort, encouraging organizations in the community to learn from the best practices, enhance response, or open the window for collaborative incident response and automation [16]. A major benefit is a common repository of standardized playbooks, which can be uniformly used by organizations from various sectors dealing with similar security issues.

The collaboration between SOCs and CERTs represents a solid basis for a unified cyberspace. It is supported by CTI-sharing platforms, particularly MISP (Malware Information Sharing Platform) and STIX/TAXII, where STIX is a common language for describing threat indicators and incidents, and TAXII is a transport protocol [27]. A prerequisite for collaboration is that organizations, sectoral SOCs, and national CERTs connect in communities and establish trusted information flows. There are several ways to engage in communities. Table 1 summarizes collaborative and group decision-making scenarios underlying the common cross-border and cross-sectoral cyberspace.

Table 1: Collaborative and group decision-making scenarios.

Scenario	Key steps and benefits
CTI sharing for the SOC	The SOC receives CTI information from the community to proactively mitigate known cyber threats; the CERT acts as an intermediary and the single point of contact.
CTI exchange from the SOC	The SOC shares CTI information about the detected cyber incident with the community to enhance the resilience of the ecosystem (sectoral, cross-sectoral, cross-border); the CERT acts as an intermediary and the single point of contact.
Standardized cooperation and reporting to the CERT	The SOC reports a security incident to the CERT through a standardized procedure; the CERT cooperates in incident response and spreads situational awareness to the community.
Management and sharing of playbooks	Standard playbooks are available for the community of SOCs and CERTs, increasing awareness and knowledge on incident response; SOCs and CERTs share and enhance playbooks to meet organizational and legislative requirements.

4 Delphi Cyber-Attack Mitigation Process

This section introduces the group decision process and the multi-criteria decision-making methodology for cyber-attack mitigation. We elaborate on our previous work [8, 10]. We also apply the process to a real-life setting to demonstrate its usability and effectiveness.

4.1 Group MCDM Methodology and Process

Figure 4 gives a compact representation of the group decision-making process. It includes the standard intelligence, design, choice, and implementation phases [32]. The first two allow the decision-making group to leverage situational awareness, CTI on the known TTPs of attackers, and the common generic mitigation measures against these TTPs. The process relies on the mitigation framework, providing two key capabilities. Firstly, it facilitates the development of mitigation strategies based on the identified threats and compromised assets. Secondly, it integrates the best practices and recommendations from established sources, including MITRE ATT&CK [24], CSC (Critical Security Controls) [9], NVD (National Vulnerability Database), and others.

The choice phase is based on the Multi-Criteria Decision-Making (MCDM) methodology and the Delphi group decision-making process [7]. It incorporates incident impact assessment, mitigation assessment, and mitigation selection steps. These steps are carried out in several consecutive Delphi rounds until a consensus or a sufficient compromise is reached. All three levels of the organization are involved in the decision, each focusing on specific and complementary criteria. The incident impact assessment model includes criteria on corporate impact, financial impact, technical severity, asset criticality, system scale, safety concerns, and ecological concerns. The mitigation selection model addresses technical requirements (expertise, equipment, cost, and time), IT and OT impact (CIA & CAIC – Confidentiality, Integrity, Availability, and Control), infrastructural complexity, asset dependencies, business impact, and business constraints (RTO and RPO).

The methodology uses the uniform Countermeasure Scoring System (CMSS). It unifies quantitative and qualitative assessments as defined in Table 2. It is based on the widely used Common Vulnerability Scoring System (CVSS) [25]. The boundaries of categories

are aligned for both systems, except that CMSS introduces the "Negligible" category to make the distribution of categories more balanced.

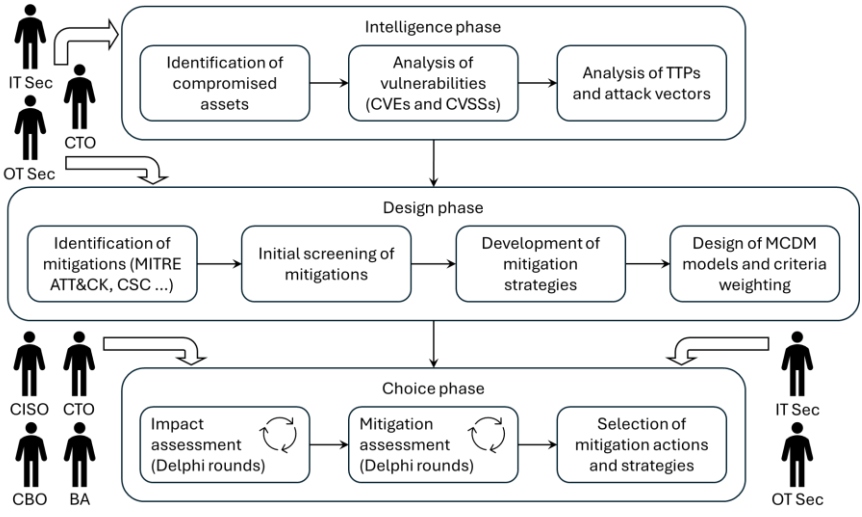

Figure 4: Group decision-making process for cyber-attack mitigation.

Table 2: Uniform countermeasure scoring system.

None	Negligible	Low	Medium	High	Critical
0.0	0.1 to 1.0	1.1 to 4.0	4.1 to 7.0	7.1 to 9.0	9.1 to 10.0

The quantitative approach is based on MAVT (Multi-Attribute Value Theory) [22]. It uses the additive value function as a simple aggregation method to obtain the incident impact score and the mitigation efficiency score, respectively:

$$s^I(A_l) = \sum_{j=1}^{n} w_j s_j^I(A_l) \quad (1)$$

$$s^E(M_k) = \sum_{j=1}^{n} w_j s_j^E(M_k) \quad (2)$$

The additive value function requires the independence of criteria. This may not always be valid, as criteria are interdependent in certain cases. However, the introduced impact assessment and mitigation selection models utilize predominantly independent criteria. For example, researchers consider the CIA (Confidentiality, Integrity, and Availability) security attributes to have no interrelated dependencies. Maček et al. [23] recommend using AHP (Analytic Hierarchy Process) related to CIA. Like the additive value function, AHP assumes each element in the hierarchy is independent of all the others.

The methodology then correlates incident impact in Eq. (1) with mitigation efficiency in Eq. (2), suggesting that a mitigation measure's suitability, efficiency, and rationality are inherently linked to the calculated severity of the incident that is addressed. The veto function reduces the mitigation efficiency, resulting in the effective mitigation score:

$$S^E(M_k) = s^E(M_k) + v^E(M_k)\big(10 - s^E(M_k)\big) \quad (3)$$

If there is no veto, the mitigation efficiency does not deteriorate. In the case of a strict veto, the effective mitigation score reaches the most critical value of 10. For weak veto degrees, veto impacts the mitigation efficiency linearly.

We determine the robustness of selected mitigation actions and strategies in the case study, which is summarized in Section 4.2, by observing the change in the weight vector required for the mitigation M_k to become equally or less effective than any of the initially inferior mitigations M_l. An optimization program is applied for this purpose:

$$\Delta_w = \min \frac{\left[\sum_{j=1}^{n} |w_j - \tilde{w}_j|^P\right]^{1/P}}{\Delta_w^{\max}} \text{ subject to} \quad (4)$$

$$s^E(M_k) = \sum_{j=1}^{n} w_j s_j^E(M_k) \geq s^E(M_l) \quad (5)$$

$$\sum_{j=1}^{n} w_j = 1, w_j^{\min} \leq w_j \leq w_j^{\max}, \forall j = 1, \dots, n \quad (6)$$

CMSS allows for a direct mapping of scores to facilitate the qualitative approach. We can also use the DEX (Decision EXpert) method [6] in our case study.

The group decision-making process implements the Delphi technique. Quantitative and qualitative measures of central tendency are calculated in each round. For impact scores, they are presented as follows:

$$\forall\left(CSE_i, A_j\right): \left(s_{\min}^I, s_{\text{avg}}^I, s_{\max}^I\right) \text{ for } \{DM_1, \dots, DM_m\} \quad (7)$$

$$\forall\left(CSE_i, A_j\right): \left(C_{\min}^I, C_{\text{median}}^I, C_{\max}^I\right) \text{ for } \{DM_1, \dots, DM_m\} \quad (8)$$

The minimum, maximum, and average quantitative scores and the best, worst, and median categories are presented in Eq. (7) and Eq. (8), respectively, based on the opinions of all decision-makers for each combination of a cybersecurity event and an asset compromised by this event. Equations are similar for mitigation scores calculated for individual mitigation actions and complex remediation strategies:

$$\forall\left(CSE_i, A_j, M_k\right): \left(S_{\min}^E, S_{\text{avg}}^E, S_{\max}^E\right) \text{ for } \{DM_1, \dots, DM_m\}, M_k \in \{MA_k, RS_k\} \quad (9)$$

$$\forall\left(CSE_i, A_j, M_k\right): \left(C_{\min}^E, C_{\text{median}}^E, C_{\max}^E\right) \text{ for } \{DM_1, \dots, DM_m\}, M_k \in \{MA_k, RS_k\} \quad (10)$$

4.2 Application of the Process

We apply the group decision-making process in the EPES (Electrical Power and Energy Systems) domain. Decision-makers from diverse organizational levels select mitigation countermeasures to strengthen the resilience of an IT/OT integrated energy distribution network supervised with a SCADA (Supervisory Control and Data Acquisition) system. This case is relevant because cyber-attacks targeted at SCADA and substation tiers of the grid have been documented in the past [3, 5].

Our additional case study focuses on the time-dependent increasing impacts of a cyber-attack on a combined cycle powerplant. Because of their complexity, we will thoroughly present both case studies in a follow-up paper. In this section, we provide a brief recap. Figure 5 depicts the attack vector on the SCADA-supervised IT/OT integrated network. Figure 6 gives a schematic representation of a possible MS_2 mitigation strategy. This ICS (Industrial Control Systems) network strengthening strategy comprises a sequence of relevant elementary mitigation actions from the MITRE ATT&CK framework. Finally,

Table 3 shows the Delphi statistics convergence for the MS_2 strategy in two consecutive rounds of the group decision process. In the table, IS denotes the information system aspect, BP represents the business process, MS is the overall quantitative mitigation score, and MC denotes the qualitative mitigation category.

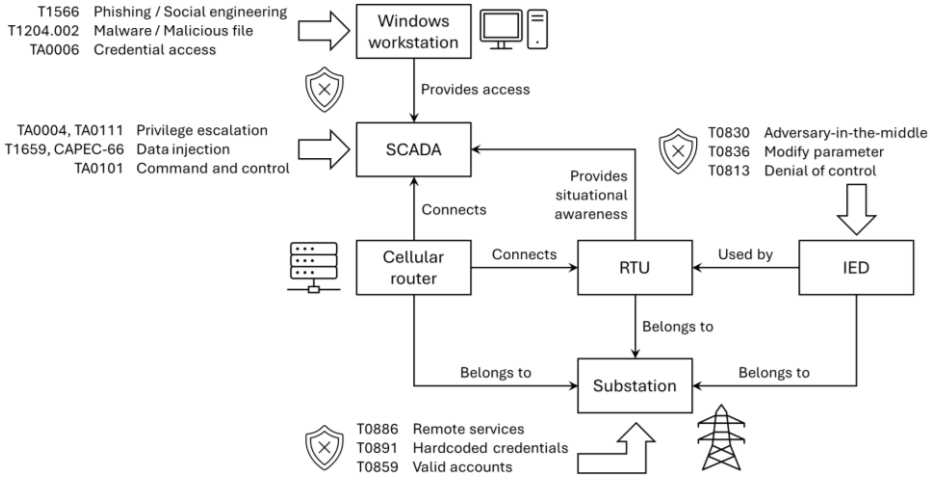

Figure 5: Attack vector on the SCADA-supervised energy distribution network.

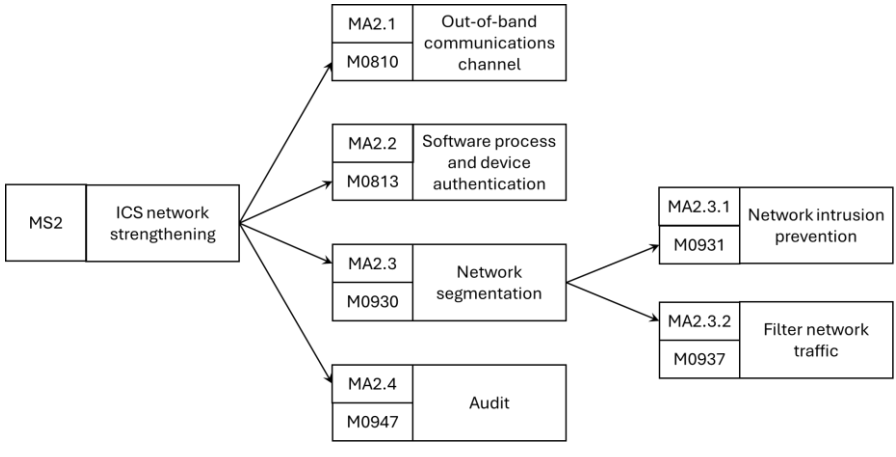

Figure 6: ICS network strengthening mitigation strategy.

Table 3: Convergence of Delphi statistics for mitigation scores.

	MS_2 1st round			MS_2 2nd round		
	Min	Mean	Max	Min	Mean	Max
IS	3.87	5.51	7.32	3.26	4.69	6.33
BP	1.34	4.00	6.26	1.34	3.17	5.08
MS	2.60	4.76	6.79	2.30	3.93	5.70
MC	Low	Medium	Medium	Low	Low	Medium

5 Conclusion

Cybersecurity is an important driver of society because it enhances the safety of digital technologies that are becoming essential in national critical infrastructures, corporate environments, and everyday life. It should be treated as a public good. This means that effective cybersecurity requires human interaction, collaboration, and decision-making. Indeed, this is the main contribution and focus of the presented research. It proposes, analyzes, and applies human-centric decision-making processes to advance the state-of-the-art of cybersecurity.

The presented methodologies help share knowledge, raise situational awareness, and facilitate collaboration and group decision-making. Group decision processes allow stakeholders to incorporate all relevant roles, asset owners, and organizational levels into proactive and reactive cybersecurity strategies. They strengthen CTI exchange, foster strategic and operational intelligence in organizations, and enhance coordination between SOCs and national CERTs. This provides a solid basis for regulatory compliance and lays a foundation for comprehensive cybersecurity programs and policymaking.

The research presented in this paper applied the group decision process for cyber-attack mitigation to the SCADA-supervised network. The case study in Section 4.2 shows that the methodology can perform efficiently. It incorporates the complementary judgments of decision-makers with various degrees of expertise and from different organizational levels. It gives an insight into CTI regarding TTPs and common mitigations. The group decision process allows decision-makers to assess accurately the real costs and effects of all mitigation actions to calibrate financial and operational impacts.

The research has some limitations. Firstly, the methodology performs calculations on data from multiple sources, including vulnerability databases, IoC (Indicators of Compromise) databases, and risk analysis data. This is computationally intensive. Secondly, selecting the best strategy in complex environments increases the intensity, which requires a high degree of automation and integration to accelerate the process. Furthermore, there are challenges associated with resources and human factors. At this point, when collecting vulnerabilities, it is necessary to perform triage to verify a component is present since CPE (Common Platform Enumeration) information is not guaranteed. Human interaction is therefore needed in this process. To avoid it, integration of external services requires significant effort to automate the methodology.

Within the scope of future research, we will enhance these aspects by developing machine learning models to automatically derive initial scores of impact assessment and mitigation selection criteria based on historical data. Decision-makers will then be able to elaborate on these objective scores according to their subjective preferences and knowledge. The framework's evolution will improve automation by using a reinforcement learning model to automatically estimate risks, predict vulnerabilities, eliminate the need for CVE triage, and learn the best strategies. This will facilitate the group decision-making process by providing correct information.

References

[1] Admas, W.S.; Munaye, Y.Y.; Diro, A.A. Cyber security: State of the art, challenges and future directions. Cyber Security and Applications, 2:100031, 2024.
[2] Ainslie, S.; Thompson, D.; Maynard, S.; Ahmad, A. Cyber-threat intelligence for security decision-making. Computers & Security, 132:103352, 2023.

[3] Akbarzadeh, A.; Erdodi, L.; Houmb, S.H.; Soltvedt, T.G. Two-stage advanced persistent threat (APT) attack on an IEC 61850 power grid substation. International Journal of Information Security, 23:2739-2758, 2024.
[4] Alaeifar, P.; Pal, S.; Jadidi, Z.; Hussain, M.; Foo, E. Current approaches and future directions for Cyber Threat Intelligence sharing: A survey. Journal of Information Security and Applications, 83:103786, 2024.
[5] Alanazi, M.; Mahmood, A.; Chowdhury, M.J.M. SCADA vulnerabilities and attacks: A review of state-of-the-art and open issues. Computers & Security, 125(103028), 2023.
[6] Bohanec, M. DEX (Decision EXpert): A qualitative hierarchical multi-criteria method. In Multiple Criteria Decision Making, Studies in Systems, Decision and Control, vol. 407, pages 39-78, Springer, Singapore, 2022.
[7] Bregar, A.: Application of a hybrid Delphi and aggregation-disaggregation procedure for group decision-making. EURO Journal on Decision Processes, 7(1-2):3-32, 2019.
[8] Bregar, A.; Husseis, A.; Flores, J.L. Group decision-making process for cyberattack mitigation. In Proceedings of the International Conference on GDN and ICDSST, vol. 2, pages 268-281, University of Porto, Porto, 2024.
[9] Center for Internet Security. CIS Controls, https://www.cisecurity.org/controls, accessed January 22nd, 2025.
[10] CyberSEAS – Cyber Securing Energy dAta Services, https://cyberseas.eu/, accessed September 27th, 2025.
[11] ENTSO-E. Network Code for cybersecurity aspects of cross-border electricity flows, https://www.entsoe.eu/network_codes/nccs/, 2022.
[12] European Parliament. Directive (EU) 2022/2555 on measures for a high common level of cybersecurity across the Union, https://eur-lex.europa.eu/eli/dir/2022/2555, 2022.
[13] European Parliament. Directive (EU) 2022/2557 on the resilience of critical entities, https://eur-lex.europa.eu/eli/dir/2022/2557/oj, 2022.
[14] Franke, U.; Andreasson, A.; Artman, H.; Brynielsson, J.; Varga, S.; Vilhelm, N. Cyber situational awareness issues and challenges. In Cybersecurity and Cognitive Science, chapter 10, pages 235-265, Academic Press, Cambridge, 2022.
[15] Gartner. Top Cybersecurity Trends for 2024, https://www.gartner.com/, accessed September 27th, 2025.
[16] Gurabi, M.A.; Nitz, L.; Bregar, A.; Popanda, J.; Siemers, C.; Matzutt, R.; Mandal, A. Requirements for playbook-assisted incident response, reporting, and automation. Digital Threats: Research and Practice, 5(3):34, 2024.
[17] Gutzwiller, R.; Dykstra, J.; Payne, B. Gaps and opportunities in situational awareness for cybersecurity. Digital Threats: Research and Practice, 1(3):18, 2020.
[18] Horneman, A. Situational Awareness for Cybersecurity: An Introduction. Carnegie Mellon University, Software Engineering Institute, 2019.
[19] Int. Organization for Standardization. ISO/IEC 27002:2022: Information security, cybersecurity, and privacy protection, https://www.iso.org/standard/75652.html, 2022.
[20] Jalali, M.S.; Siegel, M.; Madnick, S. Decision-making and biases in cybersecurity capability development. Journal of Strategic Information Systems, 28(1):66–82, 2019.
[21] Johnson, C.; Badger, L.; Waltermire, D.; Snyder, J.; Skorupka, C. Guide to Cyber Threat Information Sharing. NIST SP 800-150. NIST, Gaithersburg, 2016.
[22] Keeney, R.L.; Raiffa, H. Decisions with Multiple Objectives: Preferences and Value Tradeoffs. Cambridge University Press, New York, 1993.
[23] Maček, D.; Magdalenić, I.; Ređep, N.B. A Systematic literature review on the application of MCDM methods for information security risk assessment. International Journal of Safety and Security Engineering, 10(2):161-174, 2020.
[24] MITRE. MITRE ATT&CK, https://attack.mitre.org/, accessed January 22nd, 2025.

[25] National Institute of Standards and Technology (NIST). Vulnerability Metrics, https://nvd.nist.gov/vuln-metrics/cvss, accessed April 21st, 2025.
[26] Nelson, A.; Rekhi, S.; Souppaya, M.; Scarfone, K. Incident Response Considerations for Cybersecurity Risk Management. NIST SP 800-61. NIST, Gaithersburg, 2024.
[27] OASIS. OASIS Open Standards, https://www.oasis-open.org/standards/, accessed September 26th, 2025.
[28] Parkin, S.; Kuhn, K.; Shaikh, S.A. Executive decision-makers: a scenario-based approach to assessing cyber-risk perception. Journal of Cybersecurity, 9(1):18, 2023.
[29] Republic of Slovenia. National Cyber Incident Response Plan, https://www.gov.si/assets/vladne-sluzbe/URSIV/Datoteke/Dokumenti/2022-03-NOKI.pdf, 2021.
[30] Republic of Slovenia. Information Security Act (ZInfV-1), Official Journal of the Republic of Slovenia Act, 40/25, https://pisrs.si/pregledPredpisa?id=ZAKO8934, 2025.
[31] Ross, R.S. Managing Information Security Risk: Organization, Mission, and Information System View. NIST SP 800-39. NIST, Gaithersburg, 2011.
[32] Turban, E.; Aronson, J.E.; Liang, T.-P. Decision Support Systems and Intelligent Systems, 7th edition. Prentice-Hall, Upper Saddle River, 2004.
[33] van der Kleij, R.; Schraagen, J.M.; Cadet, B.; Young, H. Developing decision support for cybersecurity threat and incident managers. Computers & Security, 113:102535, 2022.
[34] Wagner, T.D.; Mahbub, K.; Palomar, E.; Abdallah, A.E. Cyber threat intelligence sharing: Survey and research directions. Computers & Security, 87:101589, 2019.
[35] World Economic Forum. We must treat cybersecurity as a public good, https://www.weforum.org/, accessed September 27th, 2025.
[36] Zhang, E.; Wang, G.; Ma, R.; Li, J. An optimal group decision-making approach for cybersecurity using improved selection-drift dynamics. Dynamic Games & Applications, 13:980-1004, 2023.

Multi-criteria model for evaluation of sustainable transportation initiatives

Martin Žnidaršič
Jožef Stefan Institute
Jamova cesta 39, 1000 Ljubljana, Slovenia
martin.znidarsic@ijs.si

Abstract. *This paper provides an overview of a multi-criteria evaluation that was prepared for evaluation of sustainable transportation initiatives for the Municipality of Kranj. The paper outlines the methodology, data sources, evaluation results, limitations and future considerations of the undertaken evaluation approach. The evaluation problem is formalized using a multi-criteria decision analysis approach that involves defining the criteria to measure success, defining the desirability of various values of the criteria and defining their importance. A preliminary evaluation was conducted for two time periods: December 2024 and April 2025, allowing for a demonstration of a comparative assessment. The evaluation indicates that the main differences could be driven by seasonal factors, such as bicycle use and air pollution. We also discuss and stress the importance of data quality, timeliness, and considerations that need to be made with respect to the interpretation of results.*
Keywords. Sustainable Transportation, Multi-criteria Modelling, Evaluation

1 Introduction

Climate changes and their negative aspects are spurring actions world wide to alleviate them. European Union (EU) is particularly active, with the European Green Deal policy initiatives and various related support programmes and missions.

The NetZeroCities is a project supporting the mission *100 Climate-Neutral and Smart Cities by 2030*. In its scope, the projects UP-SCALE (*UP-SCALE Urban Pioneers - Systemic Change Amid Liveable Environments*) and KReATIVE (*Kranj's Resilient Ecosystem for Active Transformation, Innovation, and Visionary Engagement*) tackle various sustainable development topics, among which sustainable transport in the Municipality of Kranj, in particular enhancements of data collection and integration for the purposes of improved urban mobility applications and services for its citizens. In the scope of this work, evaluation was recognized as a relevant aspect for effective monitoring and support of the actions. The evaluation approach was to be as transparent as possible and should allow for continuous, both short term and long term use.

In this paper we present the methodology, data sources, preliminary evaluation results, as well as limitations and future work considerations related to the undertaken evaluation approach. In the formalization of evaluation we followed elements of the Multi-attribute utility theory [4] and in selection of criteria we sought inspiration in Sustainable Urban

Mobility Indicators (SUMI) [3]. An important factor in criteria selection, however, was relevance for the intended user and actual data availability.

The paper is structured as follows: Section 2 introduces the modelling methodology, Section 3 presents the selected criteria and the related value functions, followed by an example of evaluation in Section 4 and conclusion in the last section.

2 Methodology

Evaluations of complex entities or phenomena usually involve many criteria. This is the case also with the evaluation at hand. Namely, it would be difficult to evaluate all the activities through one criterion. As usually in such cases, we will employ multi-criteria evaluation.

In order to perform multi-criteria evaluation, we need to define the criteria or attributes that we take into account. We denote them as: C_1, C_2, ..., C_n. An example of a criterion might be the number of passengers served in the public transport system. Another example might be the number of traffic accidents. And there are of course many others we might want to consider. In the criteria definition phase, it is important to identify all the criteria that are relevant for our evaluation and are at the same time operational in the sense that we can measure or assess them. The criteria can be composed of sub-criteria and sometimes it is beneficial to explicitly model such a relationship. In this sense we distinguish the basic and the aggregated criteria.

The criteria also need definitions of how they can be measured or assessed. For example, the number of passengers might be a monthly number, or an average daily number, or something else. It might be measured relatively precisely with sensors on the buses, or it might be a result of estimates by the bus drivers. Of course, we should strive for as accurate measurements as possible. While many criteria in our evaluation are numeric, some are binary and have a value of true or false, or 0 and 1. These are usually criteria that indicate goal achievements. The measurements of our criteria in a specific situation (period) will be denoted as: x_1, x_2, ..., etc., where x_n represents the measurement of criterion C_n.

Measurements of basic criteria are usually represented in different units, which hinders their comparison and collective use, such as for any weighting and aggregation purposes. In numeric multi-criteria evaluation we usually scale the measurements to a fixed interval, most commonly to the interval [0,1] with a so-called utility function or value function. Such a function transforms any input value (criterion measurement) into 0 for the worst case, 1 for the best input values and in-between for the rest. This way we do not only scale the input values, but can express our preferences at the same time. The preferences can also be piece-wise linear or non-linear in general. We denote the transformed measurements as $v(x_1)$, $v(x_2)$, ..., etc. The binary criteria can be handled in a similar way, with a value of 0 or 1 commonly prescribed to false and true values respectively.

In multi-criteria evaluation, the scaled values of selected criteria are usually combined into an aggregated overall value. There are various approaches to achieve this. In our case, we use a simple additive aggregation model, which is a common choice. The overall value in such a model is defined as:

$$V(v, w) = \sum_{i=1}^{n} w_i v_i(x_i) \tag{1}$$

Where w_i is a weight associated with the criterion C_i and $v_i(x_i)$ is the value function of criterion measurement x_i. For the aggregated criteria we will not use further transformations. Such weights, which all together sum to 1 are to be defined for all the criteria. The weights are used to express the relative importance of criteria.

This relatively simple aggregation approach allows for some transparency of the models and ease of interpretation of the results. There exist, however, even more transparent methodologies that are qualitative in nature, such as DEX [2], but we did not opt for their use for two reasons. Firstly, most of our model's input values are numerical and do not offer a straightforward transformation to qualitative values. Secondly, because we need the model to be relatively sensitive in order to reflect even slight changes of measurements.

For the purpose of our evaluation we observe and analyze both individual criteria and their overall value and their changes during at the start and at the end of a specific evaluation period. However, the methodology as set-up is made to be used also for continuous long term evaluation.

3 Criteria and value functions

In this section we outline the criteria considered in the evaluation and how the criteria values are to be assessed. The criteria to consider were selected based on the project's documentation, SUMI and the available data sources.

3.1 Bus occupancy

The *bus occupancy* criterion stands for the average occupancy of the sensor equipped city buses in the last month. With occupancy, we mean occupancy of seats and standing areas that are dedicated to passengers. Bus occupancy is measured with sensors mounted in a selection of buses. The equipment enables detection of bus entrances and exits, but not mapping of one to another (which entrance and exit were made by the same user), so we can estimate the number of people on the bus at any given moment, but not for example the length of individual rides.

The number of buses considered remained constant throughout the project. However, the routes on which the sensor-equipped buses are used in practice cannot be fixed, as they change dynamically according to the needs and requirements related to the situation at hand (state of the rest of the fleet, etc.).

Bus occupancy data is provided for one minute intervals for all the time when the bus is operational (turned on). Sensory readings are to some extent false, as we noticed that in some cases there are more exits than detected passengers. This could be attributed to people crowding in the exit area and can happen also in the entrance area for the case of entrances. Our estimation of the number of passengers thus uses some corrective measures, such as rolling minimum, mean of cumulative counts, and definition of points in time in which the entrances and exits should not increase, such as during bus movement.

The value function for the bus occupancy criterion is shown in Figure 1. No occupancy is used for the zero value and $3/4$ occupancy or higher as the one which yields the maximum value of 1. According to this value function, constant maximum occupancy or over-occupancy would also hold the value of 1, which is a point to note and potentially revisit.

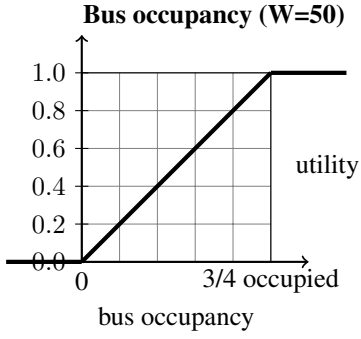

Figure 1: Value function of the *bus occupancy* criterion.

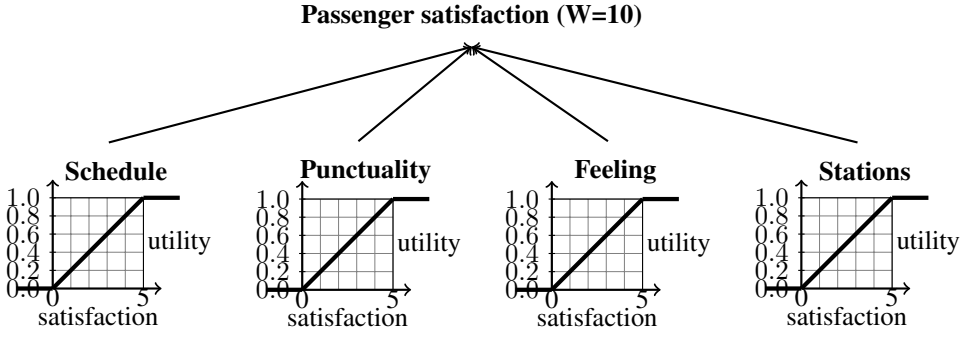

Figure 2: The value functions of criteria related to passenger satisfaction.

3.2 Passenger satisfaction

Satisfaction of the passengers with the public transport system is very important, but it is also a phenomenon that is difficult to assess regularly and in an unobtrusive manner. In our case, the assessment is made through a questionnaire that is employed in two time periods. This is one of the rare criteria that we cannot assess continuously.

The questions about passenger satisfaction in the questionnaire are formulated with an explanation: *Please indicate on a scale from 1 to 5, how would you rate the state of the city's bus service?(1 - very poor, 5 - very good).* followed by the elements to be evaluated: Schedule, Punctuality of buses, Feeling during the ride, Frequency of bus-stops, Cleanliness of bus-stops, Friendliness of bus drivers, Cleanliness of the buses, Payment system, Locations of bus stops, Price, Equipment of bus stops with covered bicycle stands.

For the purposes of our evaluation model, we are considering four of these questions. The data collected in these questions will be averaged and the outcomes transformed with the corresponding value functions and aggregated with equal weights into the overall passenger satisfaction as shown in Figure 2.

3.3 GHG emissions

Emissions of GHG are a common phenomenon to consider in sustainability assessments. We focused on the CO_2 emissions that are directly caused by the bus fleet during opera-

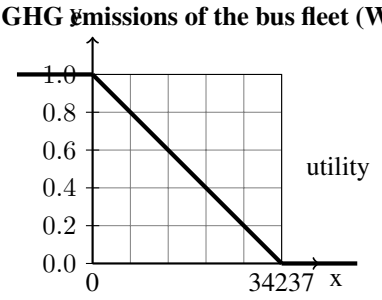

Figure 3: The GHG emissions of the bus fleet value function: the lowest value *worst permitted* corresponds to the entire fleet operating only the vehicles according to the lowest permitted standard.

tion. There were 11 buses used in Kranj for public transport in the time of the project, 3 diesel hybrid ones and 11 electric ones. According to available data, we used 885 g/km as the estimation for the emissions of the diesel hybrid buses and 0 g/km for the electric ones. Our assessment is based on the kilometers actually driven by each bus type in the assessment period. The first threshold of the value function (see Figure 3) is set to 0, which is achievable according to our definition of emissions with a fully electrified bus fleet. For the second threshold, the worst case, we took the maximum distance driven by a bus in a month in the observed period from November 2024 to April 2025 (38686 km in March 2025), which was multiplied with the emissions of diesel hybrids (885 g/km) and the total amount (11) of the buses used. This gives us the threshold value of 34237.

3.4 Local pollution

We have considered the measurements of PM10 and PM2.5 particles, which are publicly available for Kranj[1]. Concentrations of finer PM2.5 particles are more indicative of pollution by traffic[1], so we focus on these. The PM2.5 limit values according to EU standards[2] are at yearly average of $20\mu g/m^3$, which affects our value function as shown in Figure 4.

3.5 Traffic jam delays - punctuality

We measure bus punctuality as the difference between the planned (according to schedule) and real time presence of the bus at the station, which is assessed according to the bus GPS position and time of this measurement. Presence within 50 meters of the station is considered "on station." Any deviation from this is a discrepancy - typically a delay, although early departures are also included. The inputs that we are considering in the value function (see Figure 5) are 30 day averages over all the bus stop events. As a measure of average we are using the median in this case, as it is robust to rare outlier events (e.g., extreme one-time delays due to rare events), which might affect the mean. The fist threshold is at 0, while the second one is at 10 minutes.

[1] https://www.arso.gov.si/zrak/kakovostzraka/podatki/dnevne_koncentracije.html

[2] https://environment.ec.europa.eu/topics/air/air-quality/eu-air-quality-standards_en

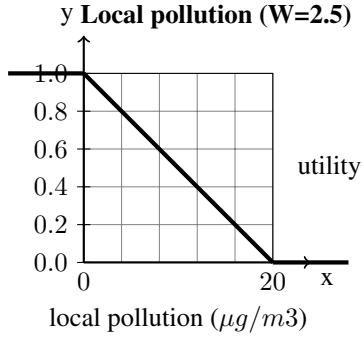

Figure 4: Value function of the *local pollution* criterion.

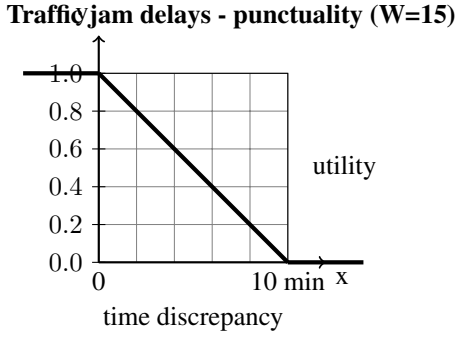

Figure 5: The value function of the bus punctuality criterion.

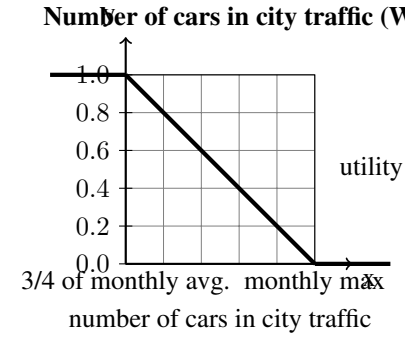

Figure 6: Value function for the number of cars in city traffic.

3.6 Number of cars in city traffic

The number of cars in city traffic is estimated based on automatic measurements of three sections that are close or cross the central part of Kranj. These are: 244 KR Primskovo 2, 718 KR Zlato polje and 913 Kokrica. The measurements of interest are daily average counts of cars based on the readings in the last 30 days. As the measurements are provided separately for two directions of traffic in each section, we use the average of the two directions. The low number threshold of the value function is set at $3/4$ of the average daily measurement in the reference period and the high threshold is set at a maximum average daily measurement in the reference period as shown in Figure 6. As we had data for the year 2024 and the first half of 2025, we considered the year 2024 to be the reference period, which resulted in the low threshold to be set at 17346 and the high one at 30501.

3.7 Use of bicycle sharing

Use of bicycles in general would be a useful criterion to consider, but as cycling counting capabilities in Kranj are currently very limited, we have opted for the use of the information on the *KRsKOLESOM* bicycle sharing system.

Specifically, we assess this criterion by the number of KRsKOLESOM rentals in the last 30 days. In construction of the corresponding value function we consider the amounts of the lowest and highest 30-day rental occurrences in a reference year. The low point of the linear part of the function corresponds to the minimal number of rentals in a 30 day period of the reference year (2023), which is 2383, and the high point to twice the maximal number of rentals in a 30 day period correspondingly. Such a maximum number of rentals in our reference year is 10705. A sketch of the value function is depicted in Figure 7.

3.8 Traffic safety

As a proxy for the traffic safety criterion we use the number of reported traffic accidents, as recorded in official police statistics[3], as an indicator of traffic safety. The utility function defines acceptable performance boundaries based on the lowest and highest monthly

[3] `https://www.policija.si/o-slovenski-policiji/statistika/prometna-varnost`

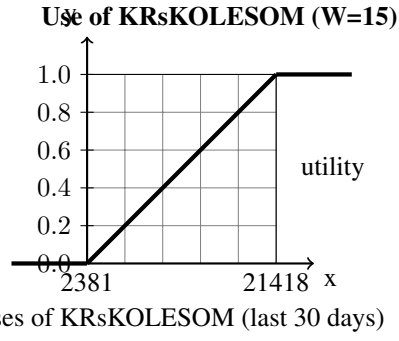

Figure 7: Value function for the use of sustainable mobility KRsKOLESOM.

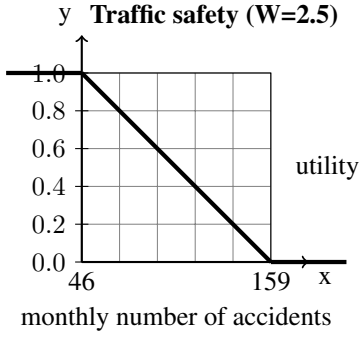

Figure 8: Value function for traffic safety.

accident counts observed over the past decade (2015-2024): 46 accidents (April 2020) and 159 accidents (October 2019). The lower threshold, recorded during a period of COVID-19 related movement restrictions, represents a challenging but aspirational target rather than a typical good result.

4 Evaluation example

In this section we present the initial evaluation results evaluation in two time periods (December 2024, and April 2025), based on the data gathered for the chosen criteria and using the proposed value functions.

For evaluation of the bus occupancy, we used the data from the buses that were active and submitting data in the evaluation period. In the data readings available, there were no strong trends detected. The occupancy remained stable throughout the observed period. In the December period the average occupancy was 4.9% and in the April period it was 4.6%.

Passenger satisfaction was assessed using a questionnaire, which was used in two surveying sessions, one in October 2024 and one in April 2025. The results of the survey in October are used as a proxy for our first evaluation point (which is in fact December). The survey was conducted among citizens of Municipality of Kranj with a representative sample in terms of age, gender and local community. The sample contains 28 respon-

dents. The average scores from the data collected in the survey are as follows: schedule = 3.03/3.25, punctuality = 3.11/2.95, feeling = 4.19/3.90 and stations = 3.85/3.79 for the first/second period.

The estimated total CO_2 emissions of the bus fleet in December are 5369 kg with a daily average of about 192 kg and in April reach 5244 kg and a 187 kg daily average. The fleet of buses that we tracked did not change during the project. The changes in the emission estimates are therefore only due to more or less intensive use of one or the other type of a bus. It is expected that the emissions will get lower, as additional electrical buses are introduced and as hybrid-diesel ones are phased out.

The average daily concentration of PM2.5 particles depends a lot on the (heating) season. In December period, the average daily concentration of PM2.5 particles was 24.260, which exceeds the threshold set for zero value. In April it was 9.730.

The median discrepancy of bus arrivals with the schedule was in December at 84 seconds, with the mean at 121.65, minimum at 0 and maximum at 1943 seconds. In April it was 96, with the mean at 131.13, minimum at 0 and maximum at 858.

Daily average number of cars in the selected road sections differs mainly with respect to distribution of week days and holidays. In our December evaluation period the detected average daily number of cars was 22940.35 in one direction and 23037.94 in the other (an average of these was used as evaluation input). In April, the detected average daily number of cars was 24214.80 in one direction and 24720.23 in the other.

The bicycle sharing system was used 4566 times in December and 9153 times in April. The lower number of uses in December is expected as it is one of the winter months when the use of bicycle sharing tends to be lower.

In December 2024, there were 109 traffic accidents registered in Kranj. There is no official statistics available for the month of April 2025 yet, so we used the latest available value (109 traffic accidents as registered for December) as a proxy and placeholder. The number of traffic accidents remains relatively stable over the years, with notable decrease in the COVID-19 affected years of 2020 and 2021.

The overall evaluation results are shown in Figure 9 and Figure 10. We can see that the latter score is better than the former. However, one must be careful in the interpretation of these results and should exploit the transparency of the evaluation method for critical assessment. In the case of this comparative evaluation, it is clear that the main effects are seasonal, as the biggest difference is in bicycle sharing use and PM2.5 pollution, which are both to a large extent affected by the season.

5 Conclusion

This paper details a methodology and assessment models for evaluating the impacts of sustainable transportation initiatives. While the evaluations offer immediate insights, the core contribution lies in the models themselves, their criteria, value functions, and data sources that were designed collaboratively by the experts from the Municipality of Kranj and the Jožef Stefan Institute. The models are intentionally transparent and simple to facilitate wide stakeholder understanding and continuous adaptation.

Successful evaluation relies heavily on data availability, and our project and the work presented in this paper highlighted the importance of data granularity, timeliness, and abundance. This informed criteria selection and necessitated using proxies in some cases. The report serves as a basis for discussions on data requirements for future assessments. For example, our evaluations and additional supporting analyses showed that the use of

december

	Min	Max	Input	Utility	Weight	Value
Bus Occupancy	0.00	75.00	4.317	0.058	0.500	0.029
Bicycle Sharing Use	2381.00	21418.00	4566.000	0.115	0.150	0.017
Traffic Safety	46.00	159.00	109.000	0.558	0.025	0.014
Schedule	1.00	5.00	3.250	0.563	0.100	0.056
Punctuality	1.00	5.00	3.107	0.527	0.100	0.053
Feeling	1.00	5.00	4.185	0.796	0.100	0.080
Stations	1.00	5.00	3.851	0.713	0.100	0.071
Local PM Pollution	0.00	20.00	24.260	0.000	0.025	-
GHG Emissions	0.00	34237.00	5608.000	0.836	0.025	0.021
Delays	0.00	600.00	84.000	0.860	0.150	0.129
Number of Cars	17346.00	30501.00	22989.000	0.571	0.025	0.014
Total	-	-	-	-	-	**0.289**

Figure 9: Evaluation results for the December period.

april

	Min	Max	Input	Utility	Weight	Value
Bus Occupancy	0.00	75.00	4.392	0.059	0.500	0.029
Bicycle Sharing Use	2381.00	21418.00	9153.000	0.356	0.150	0.053
Traffic Safety	46.00	159.00	109.000	0.558	0.025	0.014
Schedule	1.00	5.00	3.028	0.507	0.100	0.051
Punctuality	1.00	5.00	2.951	0.488	0.100	0.049
Feeling	1.00	5.00	3.900	0.725	0.100	0.072
Stations	1.00	5.00	3.786	0.697	0.100	0.070
Local PM Pollution	0.00	20.00	9.730	0.513	0.025	0.013
GHG Emissions	0.00	34237.00	5682.000	0.834	0.025	0.021
Delays	0.00	600.00	96.000	0.840	0.150	0.126
Number of Cars	17345.00	30501.00	24468.000	0.459	0.025	0.011
Total	-	-	-	-	-	**0.328**

Figure 10: Evaluation results for the April period.

PM2.5 particle concentrations are not the best indicator of traffic related emissions in Kranj, as winter heating seems to be a much more relevant source of this kind of pollution. Municipality is already considering to put in place more targeted and locally detailed systems for traffic related pollution assessments, which will provide more relevant and more abundant data.

The evaluations for December 2024 and April 2025 periods revealed a higher overall score in the latter period. However, interpreting this simply as a positive impact of some initiatives is misleading. Assessments of individual criteria are more meaningful, facilitated by the model's transparency. Acknowledging the influence of multiple factors, such as seasonality, continued use of the developed approach will allow for mitigating such influences, assessing medium to long term effects, and objectively evaluating the impact of decisions.

6 Acknowledgments

This work was partially financed by financial support from the Slovenian Research Agency for research core funding for the programme Knowledge Technologies (No. P2-0103) and by the projects *UP-SCALE* and *KReATIVE*, funded through NetZeroCities under the European Union's Grant Agreement No. HORIZON-RIA-SGA-NZC 101121530.

References

[1] B Artinano et al. "Influence of traffic on the PM10 and PM2. 5 urban aerosol fractions in Madrid (Spain)". In: *Science of the Total Environment* 334 (2004), pp. 111–123.

[2] Marko Bohanec et al. "DEX Methodology: Three Decades of Qualitative Multi-Attribute Modeling". In: *Informatica* 37 (Mar. 2013), pp. 49–54.

[3] European Commission. *Technical support related to sustainable urban mobility indicators (SUMI): Harmonisation guideline*. Final (web) version, 28 August 2020. 2020. URL: `https://transport.ec.europa.eu/system/files/2020-09/sumi_wp1_harmonisation_guidelines.pdf`.

[4] Ralph L Keeney and Howard Raiffa. *Decisions with multiple objectives: preferences and value trade-offs*. Cambridge university press, 1993.

Impact of Filtering Policy Changes on Wikipedia Pageview Metrics

Srdjan Skrbic and Zoran Levnajić
Faculty of Information Studies in Novo mesto, Slovenia
srdjan.skrbic@fis.unm.si zoran.levnajic@fis.unm.si

Abstract. *D*aily number of visits to any Wikipedia article can be obtained very simply. This research-friendly policy enabled the scientific community to study the nature and dynamics of global collective attention. However, Wikimedia Foundation has recently made two major changes in the way article visits (pageviews or viewcounts) are calculated and reported. The first change occurred in December 2019 and was related to bot traffic filtering. The second change took place in May 2020 in order to advance the detection and categorization of automated traffic. These changes improve the quality of pageview time-series by making them more stationary and reducing anomalies. Yet they lead to discontinuities that impact downstream tasks. The goal of this paper is to elucidate these changes and discuss their implications for researchers.

1 Introduction

Wikipedia is the most comprehensive free online encyclopedia. As such, it attracts researchers in both data science and social sciences. It is a major source for understanding public interest, information-seeking behavior, and digital attention. In particular, researchers frequently use pageview (viewcount) statistics from the Wikimedia REST API or Pageviews API to analyze trends over time. This data is free and simple to obtain, which resulted in several interesting papers over the last few years on collective attention and its dynamics [1, 2, 3, 4, 5, 6].

For this reason, it is critical that research community is aware of any changes in how viewcount data is measured and stored by Wikipedia. In this short paper, we report two important and substantial recent changes in this. Namely, it appears that the way of how the number of daily visits to Wikipedia articles (pages) is calculated has changed twice: the first change occurred on the 1st December 2019, and the second on the 1st May 2020.

As best we can tell, these changes are not random but correspond to modifications in Wikimedia data collection and bot filtering policies. Yet these changes introduce unwanted inconsistency of viewcount statistics for data taken before and after the changes. In the rest of the paper we discuss this issue and provide a few recommendations for researchers interested in collective digital attention.

2 Overview of Wikipedia Pageview Data

Pageview data on Wikipedia are generated from web requests to article pages. Initially, the system distinguished only between "user" and "spider" traffic—where spiders were self-identified bots (e.g., Googlebot). This categorization was naive, as many bots do not declare themselves, leading to inflated user traffic metrics. Consequently, analyses using data prior to 2020 may contain artifacts caused by automated agents.

In December 2019, Wikimedia introduced enhanced detection mechanisms for automated traffic that did not self-identify. As described in the Wikimedia Analytics documentation:

> "We identified some automated traffic that was misclassified as user traffic and built filters to remove them. This change affected pageview data as of late 2019." [8]

Although not a full overhaul, this step marked the beginning of improved bot detection and had a measurable effect on data. The total view counts of many popular pages decreased, and variability (standard deviation) in daily views was slightly reduced.

A more profound shift occurred in May 2020, when the Wikimedia team formalized a third classification type: *automated* traffic. This category included non-self-identifying bots previously lumped with user traffic.

An announcement on the Wikimedia Analytics mailing list dated 8 May 2020 explained:

> "A new 'automated' traffic class was introduced, better distinguishing automated traffic from genuine user traffic. This resulted in a 8–10% drop in user-classified pageviews on English Wikipedia." [7]

This change sharply reduced noise in pageview data, especially for pages targeted by scrapers or botnets. After this point, time-series analysis of Wikipedia data became significantly more stable, with a notable drop in view count variance. This makes it much easier to identify and measure the daily number of visits that come from human users. This number is as of May 2020 very reliable.

2.1 Two ways of obtaining pageview data

There are two methodologies for acquiring Wikipedia page view statistics. The first method utilizes the Wikimedia API, which facilitates the extraction of time series data for a designated Wikipedia page over a specified temporal range. A notable advantage of this technique is its capacity to distinguish between human and automated (bot) traffic, thereby enhancing data fidelity. However, empirical measurements indicate that the average response time per page query is approximately 550 milliseconds. Given that the English-language Wikipedia comprises over seven million articles, comprehensive data retrieval for the entire corpus would necessitate an estimated 45 days of continuous querying. Furthermore, Wikimedia's data usage policies discourage parallel querying to ensure responsible and respectful access patterns.

An alternative approach involves the acquisition of daily aggregated page view counts via Wikimedia's publicly available data dumps. These data files, archived and compressed by day, present cumulative view counts per page, inclusive of both human and automated traffic. This method permits significantly expedited and large-scale data collection but lacks the ability to disaggregate views according to traffic source.

3 Our research into collective attention

In this section we describe the research through which we became aware that these changes took place.

3.1 Data collection

We first analyzed the viewcounts of all English Wikipedia pages for the entire year of 2023. The dataset was obtained by directly downloading the corresponding files, rather than using the Wikipedia API. Subsequently, we filtered out all pages that did not reach at least 10,000 total views during 2023. This filtering step served as a rough attempt to identify a smaller subset of relevant or important pages that could be used for further analysis. After applying this threshold, we obtained a list of 38,281 pages.

For the purposes of deeper analysis, we randomly selected 5,000 pages from this filtered set. For these 5,000 pages, we retrieved all available historical viewcounts from 2-Jul-2015 through 1-Jan-2025 using the Wikipedia API. Importantly, when querying the API, we restricted the data to include only views attributed to human users ("user" category), excluding automated traffic.

The resulting dataset was organized into a tabular structure where each row corresponds to a single day and each column represents a page. Thus, each row can be interpreted as a vector containing the viewcounts of all 5,000 selected pages for that day. In this way, the full dataset consists of 3,471 such vectors.

We then measured distances between these vectors to capture temporal similarities and changes in page view distributions. Euclidean distance was employed as the primary metric, although alternative distance measures could also be considered given the high dimensionality of the vectors. We then calculated distances between the first available day (2-Jul-2015) and all subsequent days.

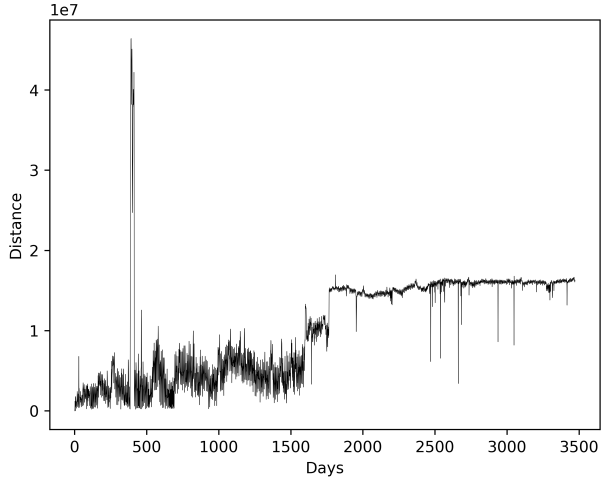

Figure 1: Distances between the first available day and subsequent days.

The analysis revealed how the distribution of Wikipedia views diverged over time from the initial baseline. The results of this analysis are shown in Figure 1, where each

point represents the distance between the first day and a subsequent day. Figure 2 shows the same plot, zoomed in to the days from 1500 to 1900.

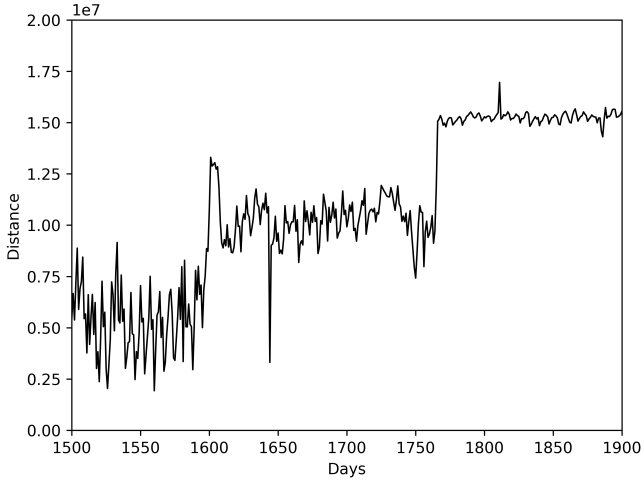

Figure 2: Distances between the first and subsequent days zoomed from day 1500 to day 1900.

3.2 Analysis and interpretation

It is immediately noticeable that a significant change occurred around the 1600th day and later around the 1765th day. The 1600th day corresponds to 18th November 2019, meaning that the first shift in the trend took place in late November 2019. The second shift occurred around 1st May 2020.

When natural intelligence encounters such results, the first thought is often an error in data acquisition or preprocessing—a flawed dataset. At the same time, one might secretly hope that it represents a genuine discovery. Specifically, distances between vectors can be interpreted as a characterization of the dynamics of human attention. In this context, two very sharp changes in trend were detected. One of them coincides approximately with the onset of the COVID-19 crisis, suggesting the possibility of a meaningful discovery. However, the second shift in May 2020 is less straightforward. It might correspond to the end of widespread lockdowns.

Nevertheless, as was suggested in the earlier part of this work, these findings unfortunately do not reflect such discoveries. Instead, we clearly identified that the observed changes are due to modifications in the way Wikipedia itself calculates and reports view-counts.

4 Conclusion and Implications for Researchers

These filtering changes create structural breaks in the Wikipedia pageview time series. Any longitudinal analysis spanning the 2015–2024 period must account for:

- Overestimation of pageviews before Dec 2019.
- Transition uncertainty from Dec 2019 to Apr 2020.
- Cleaner, lower-variance data after May 2020.

We recommend splitting time series into three segments for normalization:

1. Pre-Dec 2019: High variance, includes many bots.
2. Dec 2019–Apr 2020: Transitional filtering, partial bot removal.
3. Post-May 2020: Reliable user-only traffic.

The accuracy and interpretability of Wikipedia pageview data depend on understanding these back-end changes. The introduction of refined bot filtering mechanisms in late 2019 and early 2020 significantly affected statistical properties of the data. Researchers using these datasets must adjust for these discontinuities to ensure robust conclusions. In any case, using viewcount data starting from May 2020 should be – as best we can tell – very reliable and in no need of adjustments.

References

[1] Ruth García-Gavilanes, Anders Mollgaard, Milena Tsvetkova, and Taha Yasseri. Dynamics and biases of online attention: the case of aircraft incidents. *Royal Society Open Science*, 3(10):160460, 2016.

[2] Patrick Gildersleve, Renaud Lambiotte, and Taha Yasseri. Between news and history: identifying networked topics of collective attention on wikipedia. *Journal of Computational Social Science*, 6(2):845–875, Oct 2023.

[3] Milan Jović, Lovro Šubelj, Tea Golob, Matej Makarovič, Taha Yasseri, Danijela Boberić Krstićev, Srdjan Skrbic, and Zoran Levnajić. Terrorist attacks sharpen the binary perception of "us" vs. "them". *Scientific Reports*, 13(1):12451, Aug 2023.

[4] Ryota Kobayashi and et al. Modeling collective anticipation and response on wikipedia. In *Proceedings of the International Conference on Web and Social Media (ICWSM)*, 2021.

[5] Florian Meier. Using wikipedia pageview data to investigate public interest in climate change at a global scale. In *Proceedings of the 16th ACM Web Science Conference*, WEBSCI '24, page 365–375, New York, NY, USA, 2024. Association for Computing Machinery.

[6] Thorsten Ruprechter, Keith Burghardt, and Denis Helic. The wealth and regional gaps in event attention and coverage on wikipedia. *PLOS ONE*, 18(3):e0282425, 2023.

[7] Wikimedia Analytics Mailing List. Announcement of automated traffic filtering. `https://lists.wikimedia.org/hyperkitty/list/analytics@lists.wikimedia.org/message/LFN6HCGIFPRLZN3SJQ6IMMESVSHYKJWU/`, May 2020. Accessed: 2025-09-28.

[8] Wikimedia Foundation. Botdetection – data lake. https://wikitech.wikimedia.org/wiki/Analytics/Data_Lake/Traffic/BotDetection. Accessed: 2025-09-28.

Can Artificial Intelligence Invent in Slovenia?

International Conference on Information Technologies and Information Society (ITIS)

Ana Hafner
Rudolfovo – Science and Technology centre Novo mesto and Faculty of Information Studies Novo mesto
Podbreznik 15, 8000 Novo mesto, Slovenia
Ljubljanska cesta 31A, 8000 Novo mesto, Slovenia
ana.hafner@rudolfovo.eu

Abstract: *The rapid development of artificial intelligence in recent years has sparked global debate on whether non-human systems can be recognised as inventors under patent law. This paper examines the Slovenian legal framework to determine if AI-generated inventions can be protected within the existing system of intellectual property rights. It analyses the Slovenian Industrial Property Act and the Copyright Act, and further European Union patent legislation - European Patent Convention. Findings of this paper show that artificial intelligence can invent but cannot act as an inventor, although the Slovenian Industrial Property Act does not explicitly define an inventor as a natural person and may therefore leave the door open for non-human entities. This study contributes to broader debates on how small jurisdictions such as Slovenia face the challenges posed by artificial intelligence-driven innovation.*

Key Words*: artificial intelligence, intellectual property, industrial property, patents, inventors*

1 Introduction

Since the emergence of ChatGPT, there has been almost no scientific conference, whether in economics, management, informatics, education, or any other field, where a single topic has not overshadowed all others: artificial intelligence (AI). With curiosity, fear, and uncertainty, we wonder how profoundly it will reshape our disciplines, redefine established practices, and challenge long-standing assumptions. For some, AI represents a powerful enabler of innovation, productivity, and creativity; for others, it raises concerns about ethics, trust, and the displacement of human expertise. The intensity of these discussions reflects societal unease with the speed and unpredictability of change, especially since the most influential artificial intelligence tools are in the hands of a handful of the richest individuals.

There are at least four intersections between AI and intellectual property (IP) as presented in Table 1.

Table 1: Intersection of AI and IP

AI's role	IP issue	Key questions

AI as a creator	Authorship, inventorship	Can AI hold or transfer IP rights?
AI as a tool	IP management automation, IP search, detection and litigation	How does AI affect patent quality and strategy? How reliable are AI-based infringement systems?
AI as an object	Protecting AI models and data	What IP regime best suits algorithms and datasets?
AI and ethical dilemmas	Access to AI, openness, AI as a source of plagiarism and AI as a plagiator	How should IP adapt to promote fair AI innovation? To what extent may students or researchers use AI as a writing assistant? Can AI "plagiarise" in its learning process?

In this paper, we will focus only on the first issue, AI as a creator. AI can now write novels, create music, and produce artistic pictures. But can AI invent? This question reaches beyond creativity in the artistic sense and touches the very core of scientific and technological progress. An invention that can be protected as a patent implies novelty, inventive step, and industrial applicability. While AI can generate countless variations of existing knowledge, the debate remains whether it can truly identify problems, conceptualise solutions, and contribute to society in a way that meets the standards of a patentable invention. This paper does not have an ambition to answer this question literally, but rather in the sense: if AI could invent, can it file a patent in Slovenia? So the basic research problem is, how do intellectual property frameworks, particularly the Slovenian law, respond to these challenges? Issues that existing laws were not designed to address are AI's capacity to autonomously create, modify, and utilise IP, which raise complex questions regarding authorship, inventorship, ownership, and enforcement [1].

2 Method

The method employed in this paper is a review and analysis of current Slovenian legislation on copyright and industrial property, with a particular focus on how these legal frameworks address the challenges associated with AI. The study examines relevant provisions of the Copyright and Related Rights Act, also known as the Copyright Act (ZASP), and the Industrial Property Act (ZIL-1), as well as their interpretation in light of European and international standards. Special attention is given to questions of authorship and inventorship when works or inventions are generated with the assistance of, or autonomously by, AI systems.

3 Results

Across the EU, member states exhibit varying degrees of adaptation to AI-related IP challenges. Countries such as Germany, France, and the Netherlands are actively discussing potential legal reforms, while others, including Poland, Greece, and Romania, tend to rely more on existing frameworks and await further guidance from the EU [1]. And to which group of countries does Slovenia belong? From a review of online news, it

appears that the Slovenian Intellectual Property Office (SIPO) has acknowledged the complexities introduced by AI in the realm of intellectual property (IP), e.g. [2] [3], but direct views from the SIPO are missing, and AI is also not included in the recent National Intellectual Property Strategy 2030 [4]. We can conclude that Slovenia might be one of the countries which are waiting for further guidance [1]. Discussion on AI and IP is, however, very active in the nongovernmental sector, specifically within the Open Data and Intellectual Property Institute [5], which has organised lectures and other events related to this topic [6] [7] and aims to raise broader awareness of the topic, e.g. [8].

In Slovenia, the Copyright Act (*Zakon o avtorski in sorodnih pravicah*, ZASP) [9] primarily protects works created by human authors. Article 10 explicitly states: "An author is a *natural person* who created a work of authorship."

Current legislation does not explicitly address AI-generated content. A recent paper by Bogataj Jančič and Purkart [10] addresses text and data mining in the Slovenian legal system, specifically the implementation of the text and data mining (TDM) exceptions outlined in Articles 57a and 57b of the Copyright Act. According to the authors, these Articles provide both progressive and problematic elements of the European TDM exceptions. The TDM exceptions permit the digitisation of analogue works for TDM, remote access to content, and – specifically under the TDM exception for scientific research – the sharing of results for TDM purposes. This represents a notably progressive implementation that could serve as a model elsewhere. Rights holders are also required to ensure that beneficiaries of both exceptions can effectively conduct TDM, and they must respond within 72 hours; otherwise, they may face sanctions. Consequently, the Slovenian legal framework provides a favourable basis for developing generative AI models. The main shortcoming of the Slovenian implementation is that it does not explicitly recognise access to freely available online content as lawful access, a point clearly affirmed in Recital 14 of the DSM Directive. As a result, AI developers in Slovenia may be placed at a significant disadvantage, though it is reasonable to expect that legislators will eventually address this issue [10]. Nevertheless, researchers working on open-access large language models (LLMs) for Slovenian or other languages retain a solid legal foundation for collecting texts, building datasets, sharing them with others, and developing LLMs under the Slovenian TDM exception [10].

Besides the Copyright Act, Slovenia also has the Industrial Property Act (*Zakon o industrijski lastnini* - ZIL-1) [11], and in this case, it is less clear than in the case of copyrights, who can be considered an inventor in the case of a patent (or designer in the case of industrial design). The Industrial Property Act does not explicitly state that the inventor or designer must be a natural person; however, in a few places within this Act, we can conclude this indirectly, specifically in Articles 106 and 115.
Article 106 states: "The following data shall be entered in the patent register: /…/ data on the inventor (*surname, first name* and address) …"
Article 115 states: "(1) The inventor, his *heir or other legal successor* may, by filing a lawsuit with the competent court …"; "(2) The designer of the product design, his *heir or other legal successor* may, by filing a lawsuit with the competent court …" and "(3) The action referred to in the first or second paragraph of this Article may also be filed by *a person* who is entitled to the rights under a patent or design, if the patent is granted or the design is registered in the name of the inventor or designer or another *person* …"

We can assume that only a natural person must have a name and surname and can have a heir, but this is actually not necessary at all. AI could also have a first name and surname or any other identifier that allows it to interact socially or legally with humans. Unlike natural persons, an AI's identity doesn't arise from birth, culture, or family lineage; it is assigned, maintained, and potentially modified by its creators or users. Similarly, the concept of a "heir" could be reinterpreted: an AI might "inherit" digital assets, data rights, or decision-making authority from another AI or from a human owner, but this inheritance would not involve biological connections.

Therefore, it is not easy to find in Slovenian law why AI cannot be considered an inventor because Slovenian law does not explicitly define an inventor as a natural person, which theoretically leaves room for broader interpretations. However, in practice, patent offices and courts have consistently assumed that inventors are human. This creates a practical barrier: even if AI generates an invention independently, there is no clear mechanism for granting it patent rights, managing its obligations, or enforcing its interests. As a result, the law implicitly excludes AI from being recognised as an inventor, not because of a formal prohibition, but because existing legal structures are built around human or organisational agents. This gap highlights a tension between the formal text of the law, which could be interpreted to allow non-human inventors, and its application, which remains anchored in human-centric assumptions. To put it simply: there is no formal prohibition for AI inventors, but no mechanism exists to grant rights to AI since inventorship is implicitly human-centred.

At the EU level, the legal situation is similar. In European patent law, human inventorship is implicitly present [12], although, as implemented through the European Patent Convention (EPC) [13], it does not explicitly define an inventor as a natural person. But in practice, the European Patent Office (EPO) has consistently treated inventors as humans with the basic assumption that each invention has an individualised human inventor [12]. This is because the patent system assumes that inventors can hold rights, be credited for their work, and be accountable for legal obligations. All these are roles that AI cannot currently fulfil.

Notably, in recent years, there have been high-profile cases, such as the DABUS AI case at the European Patent Office [12] [14], where applications listing AI as the inventor were rejected by the EPO and national offices across the EU. The rationale given was that only a natural person can be an inventor under the current interpretation of patent law [12], [15], even if the statutory text does not explicitly state this. This illustrates a broader tension: while EU law may not formally exclude AI from inventorship, its application is firmly human-centred. The debate is ongoing, with some scholars and policymakers suggesting reforms.

4 Discussion and conclusion

This paper examines the legal status of AI as an inventor under Slovenian law and also compares it to EU patent law. While neither Slovenian law nor the EPC explicitly defines an inventor as a natural person, practical application has consistently treated inventors as humans, reflecting the legal system's reliance on human-centric concepts of rights. This results in the gap between the formal neutrality of the law and its human-centred

enforcement: AI can invent but cannot be an inventor.

Case studies, including the DABUS AI cases [14], illustrate how current practice excludes AI from formal recognition, despite its capacity for autonomous invention. In this case, different patent offices were asked to decide whether a patent could be granted for an invention listing the AI system DABUS as the inventor. All applications were ultimately rejected, albeit for different reasons. The European Patent Office based its decision on formal procedural rules, the UK Intellectual Property Office examined substantive legal considerations, and the US Patent and Trademark Office relied on the interpretation of statutory language [16].

Since inventorship is tied to concepts such as intention, responsibility, and rights, which are attributes that legal systems assign only to natural persons, it is probably entirely appropriate that AI cannot be recognised as an inventor. However, the perceived tension stems from the increasing ability of AI systems to generate inventive outputs. As Albayrak [14] noted, the question of AI ownership of inventions under patent law remains unresolved, with no clear legal consensus to date. Determining whether the AI itself or its human operator should be recognised as the inventor is a challenge that the legal system has yet to address. Patent law must evolve in step with advancements in AI, ensuring that AI-generated innovations are properly encouraged and protected. For Slovenia, it is not necessary to wait for EU-level reforms; national policymakers could already take proactive steps to adapt the legal framework, clarify ownership issues, and provide guidance for inventors and businesses.

5 Acknowledgements

Author would like to thank Public Agency for Scientific Research and Innovation of the Republic of Slovenia (ARIS) and Ministry of Higher Education, Science and Innovation (MVZI) for support of the “CRP – Podpora IL” project (Analysis of the supportive environment for effective management and exploitation of intellectual property in Slovenia); website: https://www.rudolfovo.eu/en/nacionalni-projekti/crp-il%3A-.

6 References

[1] V. Marchenko, A. Dombrovska, and V. Prodaivoda, ‘COMPARATIVE ANALYSIS OF REGULATORY ACTS OF THE EU COUNTRIES ON THE PROTECTION OF INTELLECTUAL PROPERTY IN THE CONDITIONS OF THE USE OF ARTIFICIAL INTELLIGENCE’, *Public Administration and Law Review*, no. 3(19), pp. 44–66, Sept. 2024, doi: 10.36690/2674-5216-2024-3-44-66.

[2] ‘Študija o vplivu umetne inteligence na kršenje ter uveljavljanje avtorske pravice in modelov | GOV.SI’, Portal GOV.SI. Accessed: Sept. 28, 2025. [Online]. Available: https://www.gov.si/novice/2022-04-05-studija-o-vplivu-umetne-inteligence-na-krsenje-ter-uveljavljanje-avtorske-pravice-in-modelov/

[3] K. Žvokelj, STATEMENT OF THE REPUBLIC OF SLOVENIA DELIVERED BY THE DIRECTOR OF THE SLOVENIAN INTELLECTUAL PROPERTY OFFICE ‘a_63_stmt_slovenia.pdf’. Accessed: Sept. 28, 2025. [Online]. Available: https://www.wipo.int/edocs/mdocs/govbody/en/a_63/a_63_stmt_slovenia.pdf?utm_source=chatgpt.com

[4] *National intellectual property strategy 2030*, Electronic ed. Ljubljana: Slovenian

Intellectual Property Office on behalf of the Government of the Republic of Slovenia, 2024.
[5] ‘ODIPI’. Accessed: Sept. 29, 2025. [Online]. Available: https://www.odipi.si/en/
[6] ‘ODIPI | ODIPI at AI 4 Science’. Accessed: Sept. 29, 2025. [Online]. Available: https://www.odipi.si/en/odipi-at-ai-4-science/
[7] ‘ODIPI | Copyright Law Foundations in Building LLM in Slovenia: Exceptions in the Copyright Law, Right Clearance via Licenses and Obstacles in Practice’. Accessed: Sept. 29, 2025. [Online]. Available: https://www.odipi.si/en/copyright-obstacles-in-building-llm-in-slovenia-exceptions-in-the-copyright-law-right-clearance-via-licenses-and-obstacles-in-practice/
[8] ‘ODIPI | “Text and Data Mining in the Slovenian Legal System” – article by Dr. Maja Bogataj Jančič and Ema Purkart’. Accessed: Sept. 29, 2025. [Online]. Available: https://www.odipi.si/en/text-and-data-mining-in-the-slovenian-legal-system-article-by-dr-maja-bogataj-jancic-and-ema-purkart/
[9] ‘Zakon o avtorski in sorodnih pravicah (ZASP) (PISRS)’. Accessed: Sept. 29, 2025. [Online]. Available: https://pisrs.si/pregledPredpisa?id=ZAKO403
[10] M. B. Jančič and E. Purkart, ‘Text and Data Mining in the Slovenian Legal System’, Stockholm IP Law Review, vol. 7, no. 2, 2024.
[11] ‘Zakon o industrijski lastnini (ZIL-1) (PISRS)’. Accessed: Sept. 29, 2025. [Online]. Available: https://pisrs.si/pregledPredpisa?id=ZAKO1668
[12] E. Stanková, ‘HUMAN INVENTORSHIP IN EUROPEAN PATENT LAW’, *C.L.J.*, vol. 80, no. 2, pp. 338–365, July 2021, doi: 10.1017/S0008197321000507.
[13] ‘European Patent Convention | epo.org’. Accessed: Sept. 30, 2025. [Online]. Available: https://www.epo.org/en/legal/epc
[14] F. D. I. Albayrak, ‘Artificial Intelligence and Patent Law: Patent Applications for DABUS’, in *Artificial Intelligence*, CRC Press, 2024.
[15] I. Buzu, ‘The Inventorship Paradox within Generative AI’, *Intellectus*, vol. 2024, p. 34, 2024.
[16] A. Engel, ‘Can a Patent Be Granted for an AI-Generated Invention?’, *GRUR Int*, vol. 69, no. 11, pp. 1123–1129, Nov. 2020, doi: 10.1093/grurint/ikaa117.

The Use of Artificial Intelligence in Education

Katarina Rojko
Faculty of Information Studies in Novo mesto
Ljubljanska cesta 31A, 8000 Novo mesto, Slovenia
katarina.rojko@fis.unm.si

Abstract: *The use of artificial intelligence (AI) in education is a topic of growing research interest as its application in schools and universities is growing exponentially. This article addresses five research questions: current trends and statistics on the use of AI in education, the advantages and disadvantages of its use, the different perspectives of students and teachers, AI-based tools in education, and how AI could change education in the coming years. The findings highlight a significant growth in the use of AI, from adaptive learning platforms and automated assessment tools to personalised teaching systems. While AI improves efficiency, personalisation, and accessibility, it also raises concerns about data privacy, equity, and over-reliance on technology. A comparative theoretical analysis shows that students often emphasise the convenience and benefits of AI for inclusion, while teachers are more cautious and focus on pedagogical effectiveness and ethical issues. An overview of the main categories of AI tools used in education includes tools for people with disabilities, intelligent tutors and learning agents, chatbots, personalised learning systems, and visualisations and virtual reality. In the future, AI is expected to influence policy debates on curriculum design, teacher training, data management, and digital inclusion. By exploring the opportunities and challenges, while also connections, contradictions, and gaps in the literature, this study advises educators and students to use AI cautiously, informedly, and responsibly, ensuring that artificial intelligence supports rather than replaces human-centred teaching and learning.*

Key Words: *Artificial Intelligence in Education, Transformation of Education, Teaching and Learning*

1 Introduction

Artificial intelligence (AI) is rapidly transforming the educational landscape, offering innovative solutions to enhance teaching and learning processes. The integration of AI in education is multifaceted, impacting various aspects of learning and teaching efficiency. Nonetheless, in general, opinions prevail that AI can facilitate teachers' administrative tasks and help implement innovative teaching methods to make learning more interactive and interesting. But first, educational institutions must understand how to make optimal use of AI and prepare educators and students to exploit its capabilities effectively [1].

Despite its benefits, integrating AI into education presents several challenges, including ethical concerns about privacy, bias, and accountability. Ensuring responsible AI governance and addressing these ethical issues are critical to leveraging AI's full potential in education [2] [3]. Furthermore, resistance from educators and the need for substantial financial support to build adequate infrastructure are significant barriers to AI adoption

[4] [5]. On the other hand, “non-resistance” from students is worrying, so it is essential to educate students about the pitfalls and responsibilities of using AI for learning.

For this reason, we decided to set the following research questions:
- What are the key statistics and trends on the use of AI in education?
- What are the advantages and disadvantages of using AI in education?
- How do students and teachers differ in their views on the use of AI in education?
- What AI tools are available for education?
- How could AI change education in the coming years?

2 Methods

This theoretical study adopted a qualitative research design grounded in a systematic review of the recent scholarly literature on AI in education. The methodological approach consisted of three sequential stages: identification of relevant sources, application of inclusion and exclusion criteria, and qualitative analysis through thematic synthesis.
In the identification phase, peer-reviewed journal articles and conference proceedings published between 2023 (the year 2023 marks a critical turning point in AI development and adoption, particularly following the emergence and widespread dissemination of advanced generative AI systems) and 2025 were retrieved from academic databases, including Scopus and Google Scholar. Furthermore, the Consensus AI tool was used to supplement the literature database.
The selection phase entailed screening based on predefined inclusion criteria: (a) explicit focus on AI applications within educational contexts, and (b) contribution of empirical data, theoretical insights, or policy analysis. Exclusion criteria included sources that addressed purely technical aspects of AI development without a substantive connection to educational practice or theory.
For the analysis phase, a qualitative content analysis was employed. Publications were systematically reviewed according to categories derived from the guiding research questions: trends and statistical insights; advantages and disadvantages of AI adoption; students' versus teachers' perspectives; available AI tools; and projected future transformations. Comparative analysis across studies enabled the identification of convergences and emergent themes. Then, AI was used to synthesise key findings from each relevant source. AI was also used to stylistically and linguistically improve the paper's text, for translations, to edit the literature list, and to harmonise the text with the instructions.

3 Results

3.1 Statistics and trends of the use of AI in education

AI is increasingly being integrated into educational systems, offering a range of benefits and applications. There is a rapid global uptake of generative AI adoption, especially in higher education [6, 7] and STEM fields [8, 9]. In general, as regards the regional and demographic disparities, there is evidence that male, urban, and IT students are more likely to adopt AI [10].

On the teachers' side, empirical studies reveal that the leading countries in AI in education are China, the US, Russia, and the UK, based on research output and implementation, while 90% of AI in education research was published since 2019 [11].

There was also a significant positive effect of AI on academic achievement, measured at 0.857–0.924 [12, 13], while AI can also reduce administrative burden. E.g. AI systems can automate the grading process, reducing the time required for assessment by 60-70% and increasing accuracy by 25-30% [14].

As regards the usage frequency among students, a recent study [15] reveals that:
- 8% to 31% of students use AI daily in their studies.
- 6% to 15% of students use AI 3-4 times per week.
- 17% to 20% of students use AI 1-2 times per week.
- 14% to 40% of students use AI 1-2 times per month.
- 20% to 29% of students reported never using AI in their studies.

On the student side, as much as 64–70% of students find AI-driven personalised learning helpful [8, 9]. Moreover, it was proven that AI-driven personalised learning can improve learning efficiency by 35% and knowledge mastery by 40% over traditional methods [16]. An important benefit of AI is also observed in improved engagement, since AI tools have been shown to increase student engagement from 45% to 78%, and higher completion rates, as assignment completion rates increased from 67% to 85% with the use of AI [17]. Lastly, it was also proven that personalised AI-driven learning systems can improve overall student performance by 15-20% [14].

3.2 Advantages and Disadvantages of Using AI in Education

Besides the many advantages of AI in education, as presented in the statistics above, this section also warns about its disadvantages. Both advantages and disadvantages are noted in relation to students, but also to teachers. Nonetheless, as teachers are generally more cautious, students are generally less critical and use AI less carefully.

Focused on the advantages, we already mentioned benefits for students: personalised learning, which is enabled by AI by adapting educational content to individual student needs, enhancing engagement and comprehension [18, 19], and intelligent tutoring systems that provide tailored feedback and support, which improve academic outcomes [20, 21].

The second benefit for students is enhanced student engagement, as AI tools can increase motivation and engagement through interactive and adaptive learning environments [21, 22] and, through gamification and interactive content, foster a more engaging learning experience [23, 24].

The third benefit for students is accessibility and inclusivity, as AI supports inclusive education by providing resources and tools that cater to diverse learning needs and styles [22, 24] and by offering round-the-clock interaction and support, making education more accessible [23].

On the side of teachers, studies reveal two main benefits. The first is improved performance and feedback: AI systems can provide immediate feedback and assessments [25, 22] and streamline the evaluation process, ensuring timely and consistent feedback [26, 27]. The second main advantage for teachers results from the automation of administrative tasks, which frees up time for educators to focus on teaching [18, 19] and enhances task management and data processing, making educational processes more efficient [28, 23].

As regards the disadvantages, they mostly affect both students and teachers, although in most studies the focus is on students. There are two main challenges where students are less educated and critically evolved:
- bias and accuracy issues, as AI systems can exhibit biases based on the data they are trained on, leading to unfair or inaccurate outcomes, and insufficient labelled data can result in errors and unreliable performance [29].
- technological dependence in the sense of over-reliance on AI, which can lead to a lack of human interaction and reduced interpersonal engagement in the classroom [26, 30], and a risk that students' dependence on technology potentially hinders their critical thinking and problem-solving skills [30].

There are also significant ethical, privacy and security concerns [18, 27], and therefore ensuring the ethical use of AI and protecting student data are critical challenges that need to be addressed [27]. Moreover, there are cost and implementation challenges, including potentially high upfront costs and the need for substantial investment in AI infrastructure, which can be barriers to implementation [27], as well as maintenance and updating AI systems [28].

It is important to note that substantial training and ongoing education are needed in relation to the proper and well-exploited potential of AI tools in education, which can be a significant challenge for teachers [29]. At the same time, the integration of AI may also lead to a reduction in traditional educational jobs [20, 28].

3.3 Teachers' Perspectives on AI in Education

To present the differing perspectives of students and teachers on the use of AI in education, we summarise insights from several studies. First, we present the teachers' perspective, and after the students'.

Teachers' perceptions of AI vary significantly, and it was proven that it is in relation to their experience. Namely, those with more than 10 years of teaching experience had more open attitudes towards AI than those with less than 5 years of experience [31]. Nonetheless, in general, teachers perceive more limitations than benefits in using AI, with concerns about inappropriate use and lack of critical review of AI-generated results [32]. Challenges include the potential to undermine human connection, ethical concerns, and the need for greater effort than traditional methods [33, 34], while benefits identified by teachers include facilitating tasks, access to resources, and the potential for personalised learning and enhanced student engagement [32, 33]. Teachers also see AI as a tool to support teaching by preparing lesson plans, generating teaching materials, and diagnosing learning difficulties [34]. However, there is significant concern about AI's impact on the risks associated with verifying knowledge [35].

As regards students' perspectives on AI in education, they generally view AI as a collaborator in their learning process, helping them generate ideas, work on tasks, and engage socially [36]. For this reason, they use AI primarily to search for information, simplify and correct their work, generate content, and assist with assignments [35, [37]. Students appreciate AI for its ability to provide individualised instruction, automate tasks, and enhance the learning experience [38, 39]. At the same time, concerns include AI's inaccuracy and potential for errors, which decrease trust in AI tools [35]. For this reason, they also recognise the irreplaceable role of human teachers [40].

3.4 AI tools in education

AI is increasingly integrated into educational settings, offering a range of tools that enhance learning, teaching, and administrative processes. In continuation, we list some key AI tools and their applications in education:

- disability tools are designed to assist students with disabilities, providing tailored support to enhance their learning experience [41],
- intelligent tutors and teachable agents, which offer personalised tutoring by adapting to individual student needs, helping to improve learning outcomes [41, 42],
- chatbots like ChatGPT, Gemini, and Microsoft Copilot are used for various educational purposes, including answering student queries, providing feedback, and facilitating interactive learning experiences [43, 44],
- personalised learning systems (PLS), which adapt content and learning paths to meet the unique needs of each student, enhancing engagement and learning efficiency [41, 45] [46, 47], and
- visualisations and virtual reality (VR) tools to create immersive learning experiences, making complex concepts easier to understand [41].

Studies also focus on some specific AI tools, as:

- ChatGPT large language model (LLM), which is most widely used for generating lesson plans, providing differentiated instruction [48] [49], offering automated and personalised feedback to students [50, 51], and is also used for creating customised content, and providing virtual tutoring [46, 47].
- MagicSchool, which assists teachers in creating detailed lesson plans, generating learning content, and providing ideas for homework and quizzes [49].
- AI-enhanced collaborative learning tools, which facilitate group work and personalised learning through feedback and recommender algorithms, improving student engagement and motivation [42].
- assessment tools, which automate grading, provide real-time feedback, and conduct data analysis to identify patterns in student performance, allowing for targeted interventions [52, 53].

These tools provide benefits, as already mentioned: personalisation, greater student engagement, and greater efficiency for teachers in administrative tasks. On the other hand, the challenges with tools are mostly related to the lack of technical expertise and equity. For this reason, educators require additional training to integrate and utilise AI tools effectively [54, 55], while ensuring equitable access to AI technologies for all students is crucial to avoid widening the digital divide [56].

3.5 Transformation of education through AI

AI is expected to continue rapidly transforming education in the coming years. Foreseen key transformations in education through AI include:

- personalised learning, since AI can tailor educational content to individual student needs, provide real-time feedback and adapt to students' learning styles and progress [57, 58],
- enhanced student engagement, as AI-driven tools such as chatbots and gamification elements can increase student motivation and participation by making learning more interactive and enjoyable [59],

- administrative efficiency, namely AI can automate routine administrative tasks, such as admissions processing, course scheduling, and resource allocation, leading to improved operational efficiency and reduced administrative burdens [60, 61], and
- adaptive assessment and feedback by automating grading and providing instant feedback, allowing teachers to focus on qualitative assessments and personalised support [62].

Due to challenges and ethical considerations, educational institutions will have to:
- ensure the protection of student data and maintain transparency in data usage [63, 64],
- develop clear data usage policies and adopt a privacy approach to mitigate these risks [64, 65], and
- provide ethical frameworks and oversight mechanisms, which are necessary to guide the responsible use of AI in education [63, 66].

Last concern, yet not of lesser significance for the future development of AI use in education, is related to the human interaction and teacher roles, as many studies warn that AI should not replace human interaction in education. For this reason, maintaining a balance between AI-driven tools and human oversight is crucial to preserve the human aspect of teaching [65, 29].

4 Discussion

This section briefly summarises findings from the literature review presented in the Results section, answering our research question:

What are the key statistics and trends on the use of AI in education?

AI adoption in education is growing rapidly, especially in higher education and STEM, with higher use among male, urban, and IT students. AI improves academic achievement, reduces grading time, and increases its accuracy. Around 8–31% of students use AI daily, while 20–29% never use it [15]. Personalised AI learning enhances efficiency, knowledge mastery, and engagement by 30-40% [16], while assignment completion and overall performance by 15-20% [17, 14].

What are the advantages and disadvantages of using AI in education?

AI in education offers numerous benefits, including personalised learning, increased efficiency, and enhanced student engagement. However, it also presents challenges, including technological dependence, ethical concerns, and implementation costs. Balancing AI use with human intervention and addressing these challenges is crucial for maximising the benefits of AI in education.

How do students and teachers differ in their views on the use of AI in education?

While both students and teachers recognise the potential benefits of AI in education, their perspectives differ significantly. Teachers are more cautious, focusing on the ethical implications and practical challenges, whereas students are more enthusiastic about the convenience and support AI provides in their learning processes. These differences highlight the need for targeted training and ethical guidelines to ensure the effective and responsible integration of AI in education.

What AI tools are available for education?

The main categories of AI tools include disability tools, intelligent tutors and teachable agents, chatbots, personalised learning systems, visualisations and virtual reality tools, and specific AI tools such as AI-enhanced collaborative learning tools, generative AI

tools, and assessment tools. These educational tools offer significant potential to enhance learning experiences, streamline administrative tasks, and provide personalised support to students. However, careful consideration of ethical implications and equitable access is essential to maximise their benefits and address potential challenges.

How could AI change education in the coming years?
In summary, AI has the potential to revolutionise education by providing personalised learning experiences, enhancing student engagement, improving administrative efficiency, and offering adaptive assessment tools. However, addressing ethical concerns and ensuring equitable access to AI technologies are critical to realising its full potential in transforming education.

5 Connections, contradictions, and gaps in the literature

Here, some of the connections and contradictions among the sources across the five subsections (3.1–3.5) of the Results section are presented. There are 5 main topics where connections among sources are obvious:
- Widespread adoption and positive outcomes.
- Teachers' benefits.
- Ethical, privacy, and security concerns.
- Training and skills gap.
- Human interaction as a persistent value.

As regards the Widespread adoption and positive outcomes, subsections 3.1, 3.2, and 3.5 form a consistent line of argument that AI brings measurable performance and engagement benefits. Namely, multiple sources [6–9, 11–17, 18–21, 57–62] agree that AI adoption in education is rapidly increasing globally and across educational levels. There is also evidence that AI improves academic achievement [12, 13], learning efficiency and knowledge mastery [16], engagement and completion rates [17]. Especially subsection 3.5 then builds on these statistics by presenting the transformation of education through AI—personalised learning [57, 58], engagement [59], and administrative automation [60–62].

Teachers' benefits are brought to front in subsections 3.2, 3.3, and 3.4, namely, the descriptions of AI tools in 3.4 concretely support the theoretical advantages mentioned in 3.2 and 3.3. Precisely, in both 3.2 and 3.3, is said that AI can facilitate teachers' work by automating grading, providing feedback, and supporting lesson planning [18, 19, 25–27, 28, 34, 48–51], while in subsection 3.4 the tools such as MagicSchool and ChatGPT [46–51], exemplify these applications, offering practical evidence for the benefits mentioned earlier.

Ethical, privacy, and security concerns appear repeatedly throughout the paper; In subsection 3.2, those concerns are mentioned [18, 27] as key disadvantages, in subsection 3.3, teachers emphasise that these are the main barriers to adoption [33–35], and in 3.5, they are combined into recommendations for institutional policy—data protection, transparency, and ethical frameworks [63–66].

Training and skills gap are highlighted in subsections 3.2, 3.4, and 3.5, e.g. the need for teacher training appears in subsection 3.2 as a “significant challenge for teachers” [29], in subsection 3.4 as a prerequisite for effective AI integration [54, 55], while subsection 3.5 implicitly under the call for institutional preparedness [29, 63–66].

Lastly, Human interaction as a persistent value theme ties together ethical and pedagogical dimensions, showing that AI should augment, not replace, human educators. Several studies emphasize maintaining human roles in education: subsection 3.2 warns of “technological dependence” and reduced interpersonal engagement [26, 30], subsection 3.3 cites teachers’ and students’ fear that AI might “undermine human connection” [33–35], while subsection 3.5 concludes that maintaining a balance between AI and human oversight is essential [29, 65].

There are also 5 main topics where contradictions among sources are obvious:
- Teachers’ openness vs. reported scepticism.
- High effect sizes vs. reported concerns.
- Students as both heavy and cautious users.
- Regional and demographic gaps vs. equity goals.
- Automation benefits vs. job security concerns.

Regarding Teachers’ openness vs. reported scepticism, subsection 3.3 reports that teachers with >10 years of experience are more open to AI [31], yet the same subsection and other sources [32–34] state that teachers generally perceive more limitations than benefits. Thus, the openness of senior teachers clashes with the overall trend of teacher caution and critical attitudes. This could reflect differing methodologies, contexts, or sample demographics in different sources.

High effect sizes vs. reported concerns are seen comparing sources from the subsections 3.1 and 3.2, which show very high quantitative benefits (e.g., +35–40% efficiency [16], +15–20% performance [14], +0.857–0.924 effect size [12, 13], implying strong effectiveness. However, subsections 3.2 and 3.3 emphasise bias, inaccuracy, and lack of trust, among teachers and students [29, 35]. This suggests a gap between performance metrics and perceived trust or usability.

Concerning the contradiction that Students as both heavy and cautious users, subsection 3.1 shows that a large share of students use AI frequently, with up to 31% daily users [15], and 64–70% find AI helpful [8, 9]. Yet subsection 3.3 portrays students as concerned about inaccuracy and valuing human teachers’ irreplaceability [35, 40]. This suggests an emerging critical but dependent user attitude among the students.

As regards the Regional and demographic gaps vs. equity goals, subsection 3.1 identifies regional and demographic disparities (male, urban, IT students are more likely adopters [10], while subsections 3.4 and 3.5 stress equitable access and avoiding the digital divide [56]. Contradiction is thus, while inclusivity is an explicit objective, empirical evidence shows inequitable adoption, indicating a misalignment between aspiration and current reality.

Automation benefits vs. job security concerns are stated in subsection 3.2, which, on one hand, praises automation for improving efficiency [18, 19, 25–28] but, on the other hand, also warns that it may reduce traditional educational jobs [20, 28]. This suggests that efficiency gains come at the potential expense of employment—reflecting a classic automation paradox.

Several gaps in the literature can be identified, both explicitly (where sources note limitations) and implicitly (where important issues are not yet addressed or are treated insufficiently). Here, some of the key literature gaps among the sources across the five subsections (3.1–3.5) of the Results section are presented. There are 5 main topics where connections among sources are obvious:
- Empirical and contextual gaps.
- Pedagogical and methodological gaps.
- Ethical, legal, and policy gaps.
- Human–AI interaction and teacher role gaps.
- Technical infrastructure and evaluation of AI tools' gaps.

Empirical and contextual gaps comprise limited longitudinal and comparative studies, insufficient data on primary and secondary education, and demographic disparities largely unexplored. As of limited longitudinal studies, most statistics in the subsections 3.1 and 3.2 are recent (post-2019), showing no long-term evaluation. Although the reason for the absence of longitudinal evidence lies in the recent uptake of AI in education, this prevents drawing conclusions on the actual educational impact and the sustainability of AI integration. There is also a limited amount of regional comparisons (e.g. among China, the US, Russia, and the UK [11]. There are also insufficient data on primary and secondary education in this paper, focusing on higher education and STEM fields [8, 9]. Moreover, demographic disparities are largely unexplored, e.g. only one source mentions that male, urban, and IT students are more likely adopters of AI [10], but this is not further examined. There is also no data on socioeconomic status, although inclusivity is mentioned in several sources [22, 24, 56].

Pedagogical and methodological gaps include a missing discussion and a lack of frameworks for critical AI literacy. With missing discussion, we warn that many studies recommend AI tools for efficiency [16–19, 25–28], and engagement [17, 59], but teachers' perspectives show confusion or scepticism, probably due to this missing framework. Also, a lack of AI literacy and critical thinking education frameworks is well visible, as while students' lack of critical awareness and overreliance on AI are noted [26, 30], there's no exploration of teaching strategies to develop AI literacy, ethical awareness, or critical thinking. Teachers' training is mentioned only in technical or operational terms [29, 54, 55].

There are clear Ethical, legal, and policy gaps, since ethics and privacy are discussed only superficially, not practically, and there is inadequate focus on AI bias and accuracy. In this sense, subsections 3.2 and 3.5 emphasise the importance of ethics, privacy, and transparency [18, 27, 63–66] but provide no frameworks, case studies, or policy evaluations. Furthermore, bias and accuracy are briefly noted [29] but not explored empirically.

Human–AI interaction and teacher role gaps include ambiguity in the future role of teachers and a limited understanding of student agency. While AI is said to assist teachers,

some sources warn of job reduction [20, 28] and loss of human connection [33–35, 65]. There is also no case presented of teacher–AI collaboration, suggesting that there is an absence or a lack of a conceptual or empirical framework for teacher–AI co-teaching models. Limited understanding of student agency is revealed in the finding that in the sources, students are depicted as either users or recipients of AI, not as active agents co-shaping AI systems.

Technological and evaluation of AI tools' gaps are also visible in the sources. There is a missing discussion on technical infrastructure and a lack of deeper evaluation of specific AI tools for education. Although the present study purposefully avoided sources addressing purely technical aspects of AI development without a substantive connection to educational practice or theory, there is still a place for a broader discussion on cost and implementation challenges, which are only mentioned [27, 28] but not quantified or compared across contexts. Furthermore, a more profound evaluation of specific tools, including comparative evaluation of effectiveness, would help to assess the pedagogical effectiveness of different AI tools, namely, there is a significant variation in design, quality, and user experience.

The study of sources thus reveals that there are already various studies documenting early adoption, engagement, and quantitative performance gains, but on the other hand, there is a lack of exploration of the social, pedagogical, ethical, and systemic dimensions of AI's role in education.

6 Conclusion

The future of AI in education looks promising, with ongoing advancements in machine learning, natural language processing, and predictive analytics. However, various dilemmas are justified, while there are also some contradictions and gaps in the literature. Moreover, it is difficult to ignore the fact that the use of AI actually incurs significant costs, and for this reason, it might be only a matter of time before even basic use will no longer be "for free". There is a danger that the current period, in which AI is rapidly becoming embedded in education and everyday life, is merely a transitional one. The use of AI is becoming self-evident and almost necessary, while AI services are already available under the so-called freemium model (limited and/or basic functionalities available for free, the rest for a fee). What if all AI services become exclusively fee-based, when the population is already so dependent on AI that there's no way out, and we're forced to pay for everything? Such development would cause even greater inequalities among individual students, teachers, and educational institutions—the poor would become even more inadequate, and the rich would become even richer.

7 Acknowledgements

This research was partially supported by the RRP pilot project "Applied Computer Skills" financed by the Slovenian Ministry of Higher Education, Science and Innovation, and the European Union—NextGenerationEU, and by ARIS Program P5-0445.

8 References

[1] Junaidi, J. Education reform in the application of artificial intelligence in supporting learning media: AI and education for Sustainable Development Goals (SDGS). AI, Policy, and the Future of Human-Centered Education, 2025.

[2] Torrisi-Steele, G. AI in Higher Education: Opportunities and Challenges. Foundations and Frameworks for AI in Education, 2025.
[3] Burkett, J.R. Artificial intelligence and school leadership: Using ethical leadership as a framework for monitoring the use of AI in schools. Responsible AI Integration in Education, 2025.
[4] Rosmayanti, V. Artificial intelligence in social sciences: Transforming education and research. AI Use in Social Sciences, 2025.
[5] Vasou, M.; Kyprianou, G.; Amanatiadis, A.; Chatzichristofis, S.A. Transforming Education with AI and Robotics: Potential, Perceptions, and Infrastructure Needs. Lecture Notes in Business Information Processing, 2025.
[6] Sergeeva, O.V.; Zheltukhina, M.R.; Shoustikova, T.; Tukhvatullina, L.R.; Dobrokhotov, D.A.; Kondrashev, S.V. Understanding higher education students' adoption of generative AI technologies: An empirical investigation using UTAUT2. Contemporary Educational Technology, 2(1):15-35, 2025.
[7] Kalashi, K.; Montazer, G.A.; Saed, S.; Olyaee, S. Adoption of Chatbots in Higher Education: Global Insights and Recommendations for Iran. In Proceedings of the 2025 12th International and 18th National Conference on e-Learning and e-Teaching, pages 120-135, Tehran, Iran, 2025.
[8] Slepankova, M.; Kilianova, K.; Kockova, P.; Kostolanyova, K.: Kotyrba, M.; Habiballa, H. Student Perceptions and Preferences in Personalized AI-driven Learning. Acta Informatica Pragensia, 12(2):45-67, 2025.
[9] Kovalchuk, V.; Reva, S.; Volch, I.; Shcherbyna, S.; Mykhailyshyn, H.; ; Lychova, T. Artificial intelligence as an effective tool for personalized learning in modern education. Environment Technology Resources - Proceedings of the 16th International Scientific and Practical Conference, pages 88-101, Kyiv, Ukraine, 2025.
[10] Kour, J.; Bhatia, R.; Joshi, R. Artificial Intelligence in Education: Insights From Gender and Locality-Based Perspectives. Impacts of AI on Students and Teachers in Education 5.0, 2025.
[11] Maphosa, V.; Maphosa, M. Artificial intelligence in higher education: a bibliometric analysis and topic modeling approach. Applied Artificial Intelligence, 37(6):512-534, 2023.
[12] Liu, X.; Guo, B.; He, W.; Hu, X. Effects of Generative Artificial Intelligence on K-12 and Higher Education Students' Learning Outcomes: A Meta-Analysis. Journal of Educational Computing Research, 64(3):412-435, 2025.
[13] Dong, L.; Tang, X.; Wang, X. Examining the effect of artificial intelligence in relation to students' academic achievement: A meta-analysis. Computers and Education: Artificial Intelligence, 6:100123, 2025.
[14] Savytska, I.; Bulgakova, O.; Zbaravska, L.; Ruciņš, Ā.; Aboltins, A.; Mushenyk, I.; Vasileva, V.; ; Kultenko, V. Application of Artificial Intelligence to Automatically Verify Student Calculations in Higher Education Institutions. Environment Technology Resources - Proceedings of the 16th International Scientific and Practical Conference, pages 78-91, Kyiv, Ukraine, 2025.
[15] Nechyporenko, V.; Hordiienko, N.; Pozdniakova, O.; Pozdniakova-Kyrbiatieva, E.; Siliavina, Y. How often do University Students use Artificial Intelligence in Their Studies? WSEAS Transactions on Information Science and Applications, 22(5):112-130, 2025.
[16] Chen, Y.; Liu, Q. Real-time Optimization Research and Implementation of Personalized Learning Path Based on Multimodal Artificial Intelligence. In Proceedings of the 2025 International Conference on Artificial Intelligence and Educational Systems,

pages 56-69, Beijing, China, 2025.
[17] Shtayyat, A.; Gawanmeh, A. Enhancing Educational Diversity and Inclusion Through AI-Enabled Adaptive Moodle Architecture. In Proceedings of the 2025 1st International Conference on Computational Intelligence Approaches and Applications, pages 101-112, Amman, Jordan, 2025.
[18] Aboimova, I.; Kirillova, E.; Otcheskiy, I.; Kulikova, S.; Vaslavskaya, I.; Polozhentseva, I. Use of Artificial Intelligence in the Educational Environment: A Qualitative Study of the Technology's Potential. Relacoes Internacionais no Mundo Atual, 9(2):34-51, 2024.
[19] Manglani, J.; Trivedi, A.; Madhani, J. The impact of artificial intelligence (AI) on education - A review paper. Digital Transformation and Sustainability of Business, 8(1):22-41, 2025.
[20] Dragojević, T.N.; Letić Lungulov, M.M. Artificial Intelligence in the Educational Context: Value and Challenges. Didactica Slovenica - Pedagoska Obzorja, 39(1):12-27, 2024.
[21] Bekdaş, M. The Pros and Cons of ChatGPT in Foreign Language Teaching and Its Impact on Student Motivation. Arab World English Journal, 16(4):110-125, 2025.
[22] Al-Shahrani, H.A.Z.; Albahiri, M.H.; Alhaj, A.A.M. Exploring the Benefits and Challenges of Utilizing Artificial Intelligence in Education from the Perspective of Teaching Cadre at Bisha University. Educational Process: International Journal, 14(3):201-219, 2025.
[23] Kirillova, E.; Klochko, E.; Akhmetshin, E.; Kozachek, A. Legal Aspects of the Use of Artificial Intelligence in the Educational Environment of Universities. In Proceedings of the 2025 Communication Strategies in Digital Society Seminar, ComSDS 2025, pages 45-58, Moscow, Russia, 2025.
[24] Julien, G. The significance of artificial intelligence (AI) and inclusive education. Integrating the Biopsychosocial Model in Education, 2024.
[25] Dai, K.; Liu, Q. Leveraging artificial intelligence (AI) in English as a foreign language (EFL) classes: Challenges and opportunities in the spotlight. Computers in Human Behavior, 146:107-120, 2024.
[26] Babo, L.; Mendonca, J.M.P.; Queiros, R.; Silva, L.; Almeida, P.; Mascarenhas, D.; Mascarenhas, D. Exploring HEIs Students' Perceptions of Artificial Intelligence on their Learning Process. In EEITE 2024 - Proceedings of 2024 5th International Conference in Electronic Engineering, Information Technology and Education, pages 55-70, Lisbon, Portugal, 2024.
[27] Pan, M.; Wang, J.; Wang, J. Application of Artificial Intelligence in Education: Opportunities, Challenges, and Suggestions. In Proceedings of the 2023 13th International Conference on Information Technology in Medicine and Education, pages 120-133, Shanghai, China, 2023.
[28] Al-Tkhayneh, K.M.; Alghazo, E.M.; Tahat, D. The Advantages and Disadvantages of Using Artificial Intelligence in Education. Journal of Educational and Social Research, 13(2):45-58, 2023.
[29] Gouveia, A.J.; Costa, R.; Gomes, S.; Silva, D.; Briga, L. Understanding AI Integration Challenges in Education: A Brief Systematic Literature Review. Communications in Computer and Information Science, 165:77-92, 2025.
[30] Álvarez-Herrero, J.-F. Opinion of university students in education on the use of AI in their academic tasks. European Public and Social Innovation Review, 5(1):33-49, 2024.
[31] Prasetya, Y.Y.; Reba, Y.A.; Muttaqin, M.Z.; Taufiqulloh, T.; Mataputun, Y. Teachers' Perception of Artificial Intelligence Integration in Learning: A Cross-Sectional Online

Questionnaire Survey. In Proceedings of the International Conference on Education and Technology, ICET 2024, pages 99-110, Jakarta, Indonesia, 2024.
[32] Frutos, N.D.D.; Carrasco, L.C.; Maza, M.S.D.L.; Etxabe-Urbieta, J.M. Application of Artificial Intelligence (AI) in Education: Benefits and Limitations of AI as Perceived by Primary, Secondary, and Higher Education Teachers. Revista Electronica Interuniversitaria de Formacion del Profesorado, 27(4):123-139, 2024.
[33] Çela, E.; Vajjhala, N.R.; Fonkam, M.M. Teachers' roles and perspectives on AI integration in schools. Teachers' Roles and Perspectives on AI Integration in Schools, 2024.
[34] Wardat, Y.; Tashtoush, M.A.; AlAli, R.; Saleh, S. Artificial Intelligence in Education: Mathematics Teachers' Perspectives, Practices and Challenges. Iraqi Journal for Computer Science and Mathematics, 8(2):55-70, 2024.
[35] Romaniuk, M.W.; Łukasiewicz-Wieleba, J. Generative Artificial Intelligence in the teaching activities of academic teachers and students. International Journal of Electronics and Telecommunications, 70(1):15-29, 2024.
[36] Long, Z.; Luo, D.; Gao, H.; Xu, S.; Hu, X. What Makes the Ideal AI Collaborator? Exploring Student and Teacher Perspectives on Roles, Support, and Challenges. In Global Chinese Conference on Computers in Education Main Conference Proceedings (English Paper), pages 45-61, Shanghai, China, 2025.
[37] Zhou, Y. The impact of the artificial intelligence (AI) Art Generator in pre-service art teacher training. In Proceedings of the 2024 IEEE Conference on Artificial Intelligence, CAI 2024, pages 78-89, Los Angeles, USA, 2024.
[38] Nawaila, M.B.; Erçağ, E.; Kanbul, S.; Akdağ, S.; Sertttaş-Yırtıcı, Z. Artificial Intelligence in Education: Discussing the Ethics. Sustainable Civil Infrastructures, 2025.
[39] Alam, A.; Mohanty, A. Facial Analytics or Virtual Avatars: Competencies and Design Considerations for Student-Teacher Interaction in AI-Powered Online Education for Effective Classroom Engagement. Communications in Computer and Information Science, 180:33-50, 2023.
[40] Dolenc, K.; Brumen, M. Exploring social and computer science students' perceptions of AI integration in (foreign) language instruction. Computers and Education: Artificial Intelligence, 5:100112, 2024.
[41] Yoo, A. Artificial intelligence in classrooms. A Biologist's Guide to Artificial Intelligence: Building the foundations of Artificial Intelligence and Machine Learning for Achieving Advancements in Life Sciences, 2024.
[42] Kovari, A. A systematic review of AI-powered collaborative learning in higher education: Trends and outcomes from the last decade. Social Sciences and Humanities Open, 7:100379, 2025.
[43] Guven, C.; Altinpulluk, H. Unveiling the untapped potential: Revolutionizing education with cutting-edge artificial intelligence tools. Integration Strategies of Generative AI in Higher Education, 2024.
[44] Vidalis, S.M.; Subramanian, R.; Najafi, F.T. Revolutionizing Engineering Education: The Impact of AI Tools on Student Learning. ASEE Annual Conference and Exposition, Conference Proceedings, 2024.
[45] Kovari, A.; Katona, J. Transformative Applications and Key Challenges of Generative AI. CANDO-EPE 2024 - Proceedings: IEEE 7th International Conference and Workshop Obuda on Electrical and Power Engineering, pages 12-25, Budapest, Hungary, 2024.
[46] Yogi, M.K.; Chowdary, Y.R.; Santhoshi, C.P.R.S. Impact of Generative AI Models on Personalized Learning and Adaptive Systems. Empowering Digital Education with

ChatGPT: From Theoretical to Practical Applications, 2024.
[47] Gupta, S.; Dharamshi, R.R.; Kakde, V. An Impactful and Revolutionized Educational Ecosystem using Generative AI to Assist and Assess the Teaching and Learning benefits, Fostering the Post-Pandemic Requirements. In Proceedings of the 2nd International Conference on Emerging Trends in Information Technology and Engineering, ic-ETITE 2024, pages 44-59, Mumbai, India, 2024.
[48] Guilbault, K.M.; Wang, Y.; McCormick, K.M. Using ChatGPT in the Secondary Gifted Classroom for Personalized Learning and Mentoring. Gifted Child Today, 48(3):210-225, 2025.
[49] Kiryakova, G. Artificial intelligence as a supportive tool for teachers' activities. In Proceedings of the International Conference on Virtual Learning, pages 33-45, Sofia, Bulgaria, 2024.
[50] Alers, H.; Malinowska, A.; Meghoe, G.; Apfel, E. Using ChatGPT-4 to Grade Open Question Exams. Lecture Notes in Networks and Systems, 321:112-125, 2024.
[51] Aşik, G.; Öztüfekçi, A. Exploring artificial intelligence in assessment: Unpacking the whats and hows in education. Generative Artificial Intelligence Applications: Holistic Reflections From The Educational Landscape, 2025.
[52] Vetrivel, S.C.; Vidhyapriya, P.; Arun, V.P. The Role of AI in Transforming Assessment Practices in Education. AI Applications and Strategies in Teacher Education, 2024.
[53] Dubey, P.K.; Crevar, A.R.; Rischard, M.K. Relevance of Artificial Intelligence in Assessment: A Balanced Approach. Foundations and Frameworks for AI in Education, 2025.
[54] Mohammed, M.A.E.; Al-Osail, A.F.; Assiry, B.; Ibrahim, S.A.; Abdellatif, M.S.; Alagiri, M.T.A.A.; Elbagoury, M.A.E. Leveraging artificial intelligence for enhanced electronic course design and student achievement: Unlocking the potential of AI in education. Research Journal in Advanced Humanities, 2025.
[55] Mupaikwa, E. The Application of Artificial Intelligence in Educational Administration. AI Adoption and Diffusion in Education, 2025.
[56] Singh, P. Artificial intelligence and student engagement: Drivers and consequences. Cases on Enhancing P-16 Student Engagement With Digital Technologies, 2024.
[57] Dan, L.; Mohamed, H.B.; Abuhassna, H. A bibliometric analysis of artificial intelligence: Impact on student motivation in education. AI in Education, Governance, and Leadership: Adoption, Impact, and Ethics, 2025.
[58] Wadhwa, R.; Rabby, F.; Bansal, R.; Hundekari, S. Drivers and impact of artificial intelligence on student engagement. AI Algorithms and ChatGPT for Student Engagement in Online Learning, 2024.
[59] Ouariach, S.; Ouariach, F.Z.; Khaldi, M. Artificial Intelligence as a Harbinger of Engagement and Collaboration. Ethics and AI Integration Into Modern Classrooms, 2025.
[60] Kayyali, M. AI in Higher Education: Revolutionizing Curriculum and Administration. AI Adoption and Diffusion in Education, 2025.
[61] Suazo-Galdamés, I.C.; Chaple-Gil, A.M. Impact of Intelligent Systems and AI Automation on Operational Efficiency and User Satisfaction in Higher Education. Ingenierie des Systemes d'Information, 2025.
[62] Sun, J.; Kwong, C.-F.; Buticchi, G. The Potential of AI in Electrical and Electronic Engineering Education: A Review. In Proceedings of the 2024 IEEE 11th International Conference on E-Learning in Industrial Electronics, ICELIE 2024, pages 33-48, Beijing, China, 2024.
[63] Yuri, R.; Caro, C.; Rituay, A.; Llanos, K.; Perez, D.; Sánchez Bardales, E.; Santos,

R. Ethical Challenges Associated with the Use of Artificial Intelligence in University Education. Journal of Academic Ethics, 23(2):101-120, 2025.
[64] Ismail, I.A. Protecting Privacy in AI-Enhanced Education: A Comprehensive Examination of Data Privacy Concerns and Solutions in AI-Based Learning. Impacts of Generative AI on the Future of Research and Education, 2024.
[65] Manoharan, G.; Ashtikar, S.P.; Dudhagara, C.R. Ethics of AI in the Educational Sector - Navigating the Moral Landscape. Communications in Computer and Information Science, 175:55-72, 2025.
[66] Wang, V. Ethics and equity in AI-driven education. AI Integration Into Andragogical Education, 2025.

How should we benchmark community detection algorithms in complex networks?

Robi Pritržnik
Faculty of Information Studies
Ljubljanska cesta 31a, 8000 Novo mesto, Slovenia
robi.pritrznik@fis.unm.si

Abstract. *In this article we discuss how should we benchmark community detection algorithms in complex networks. We compare the community detection algorithms Louvain, Leiden, Label Propagation, Fast Label Propagation, Greedy modularity, Infomap, Walktrap and Girvan-Newman on complex networks of the Zachary karate club, Synthetic Network, a social network from X (Twitter), a neuroscience network, a email network and a patent citation network in the USA. We find that the speed of algorithms depends on the size and structure of networks. It turns out that among the considered algorithms for community detection in large networks, the Leiden algorithm is the most suitable, while on average the Fast Label Propagation algorithm performed the fastest in all cases. It is shown that on LFR benchmark network, algorithms successfully detect the same number of communities, however when we apply the same algorithms on complex networks the results are variable based on specific algorithm.*

Keywords. community detection, networks and graphs, network analysis, complex networks

1 Introduction

With the development of complex systems and their study, mathematical models have been designed, which are derived from graph theory, with which we attempt to model these complex systems and study their properties. With models, we primarily want to gain insight into the structure and organization of a complex system and thereby enable the analysis and interpretation of its operation.

In the following section we provide an theoretical overview of graphs and networks following with network analysis techniques such as community detection. We focus on benchmarking community detection algorithms.

Different approaches towards benchmarking community detecting algorithms have been proposed and most of the approaches rely on benchmarking on synthetic networks with known network structure. This opens a question on how such algorithms act in different networks.

1.1 Graphs and networks

A graph is defined as an ordered pair G = (V, E), where V is the set of vertices and E is the set of pairs of vertices, called edges. The set of vertices in a graph G is thus denoted by V(G), and the set of edges in a graph G is denoted by E(G).

Networks are graphs that have a prescribed meaning or scope. Networks appear in many fields, such as social networks, information and communication (ICT) networks, biological networks, logistics and transport networks [4].

Networks represent an abstraction of different systems and primarily serve as a representation in the form of a mathematical model [4]. Table 1 shows the areas of application of networks by presenting the nodes and edges of an individual network.

Table 1: Examples of real networks and their properties

Domain	**Network**	**Nodes**	**Edges**
social	Facebook	people	friendships
logistics	railway	stations	tracks
ICT	World Wide Web	web pages	hyperlinks
economic	stock exchange	investors	trades
neuroscience	brain activity	neurons	synapses
biological	animal groups	animals	feeding relations
bibliographic	citations	articles	citations

1.2 Community detection

With the emergence of the field of network analysis, many studies have focused on analyzing and extracting topological features of networks. One such feature is the community structure. A community is a subgraph in which the nodes are densely connected [4].

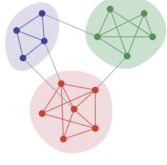

Figure 1: Communities in network [4]

Although a community is not formally defined, the theory and methods for community detection are still developing. Basically, it is a problem in which we want to identify nodes that are strongly connected to each other [6]. With the help of community detection, we want to identify and group nodes into communities that are similar in their properties. In other words, community detection is the process of assigning one or more communities to nodes in a network to which they belong.

Networks contain many nodes and connections and are usually difficult to represent graphically, but if we can identify communities, they are easier to visualize and represent the complex structure of the system.

In the community detection process, we need to algorithmically scan the network to identify communities. We want the algorithms to work quickly and independently of

the network size. Given this, it is important that the time complexity of an individual algorithm is as small as possible. The time complexity of an algorithm is a function of the input parameters of the algorithm. In general, the problem of community detection in networks is an NP-hard problem [7].

1.3 Modularity in networks

One of the fundamental properties that describes a network is its modularity. Modularity represents the ratio between the actual number of edges in the network and the expected number of edges, assuming that edges between vertices were assigned randomly while preserving the vertice degree. The expected number of edges between vertices i and j is equal to $\frac{k_i k_j}{2m}$, where k_i and k_j denote the degrees of vertices i and j, and m is the total number of edges. Modularity is denoted by the letter Q [10].

$$Q = \frac{1}{2} \sum_{ij} \left(A_{ij} - \gamma \frac{k_i k_j}{2m} \right) \delta(c_i, c_j)$$

In the previous equation, m denotes the total number of links, A is the adjacency matrix of the graph, k represents the degree of nodes i and j, and γ is the resolution parameter. δ takes the value 1 if c_i and c_j belong to the same community, and 0 otherwise (Newman, 2010). The values c_i and c_j denote the community labels. The resolution parameter γ allows favoring of communities: if γ is smaller than 1, the method favors smaller communities, otherwise it favors larger ones. This parameter was introduced because modularity without it may fail to detect smaller communities. By including γ, we enable a certain degree of sensitivity in community detection.

Intuitively, if we have a network that is divided into communities that have many connections within the community and few between individual communities, then the modularity of such a network is high, since the network is well divided into individual communities. However, if the network does not have a clear division into communities, the modularity is low. [4]

1.4 Community detection algorithms

In the following, we discuss various methods for detecting communities in graphs and networks. We also describe the algorithms discussed and their basic properties.

Girvan-Newman algorithm detects communities by removing edges from the network. The algorithm works by first calculating betweenness centrality [1]. for all nodes in the network. It then iteratively removes links and calculates betweenness centrality until it reaches mutually unconnected components or isolated communities. [12].

Greedy modularity algorithm uses a greedy optimization method of modularity until it reaches its maximum [9]. It can get stuck in a local optimum, since it relies on a heuristic and focuses only on the best solution at each step, which is not necessarily the best global solution for the entire problem [19].

Louvain algorithm is a well-known and widely used algorithm that also relies on modularity for partition optimization [1]. An advantage of the Louvain algorithm is its good performance on large networks, while its drawback is that the number of detected communities depends on the partitioning of the network. Due to the nondeterministic

[1] Betweenness centrality is defined as the ratio of the number of shortest paths between nodes h and j through node i, to all shortest paths between nodes h and j [3]

nature of the network, different results may be obtained across different iterations of the algorithm [19].

Leiden algorithm is an improvement over the Louvain algorithm and consists of three phases. In the first phase, nodes are explored locally, in the second phase partitions are refined, and in the third phase the network is aggregated into a new network based on the improved partition. [20]

Fast Label Propagation and Label Propagation. Algorithms based on Label propagation approaches are very fast and relatively easy to implement, since they rely on the idea of assigning labels to nodes, where a community is represented by nodes that share the same label. Fast Label Propagation algorithm is an improvement over Label Propagation by optimizing the order of iterations when examining node neighbors [20]. Both algorithms are suitable for large networks, as they are simple to use and applicable to different types of communities.

Random walks are based on the simulation of movements through graphs and define communities as those parts of the walks where the walk takes the most time. A random walk in a network is performed iteratively. The condition for a walk to take a long time in a community is high internal connectivity within the community, which is also one of the characteristics for community detection. The most well-known examples of this approach are the **Infomap algorithm** [16] and the **Walktrap algorithm** [15]. The main difference between the two algorithms is that in the Infomap algorithm, the random walk flow is conditioned by a map equation that minimizes movement within individual communities. The space of possible walks is determined deterministically by greedy search and uses a simulated cooling approach [2]

2 Methods

Analysis of just one network for benchmarking community detection algorithms is not optimal for understanding the dynamics of different algorithms. [11]. Therefore we need to use a different approach for benchmarking. Application of real networks for community detection shows that the performance of specific algorithms is affected by networks characteristics, therefore we also need to use real networks and not just synthetic networks. [13]. It has also been shown that usage of different benchmark approaches on one network provides different algorithm ranking, so we need to extend the analysis on multiple networks. [8]

We use two methods for benchmarking community detection algorithms. In the first method we benchmark community detection algorithms on LFR benchmark network. [11]. Such networks have known communities, therefore we can use them to run different community detection algorithms and efficiently compare the accuracy results by comparing the number of detected communities and modularity. The benefit of this method is that we can provide an exact accuracy of specific algorithm for such networks because of the known communities upfront.

In the second method we iteratively run comparison on different complex networks and compare the results on each algorithm. In this method we usually do not know exactly how many communities specific network has, however we extend the accuracy of community detection benchmarking to a wider spectrum of different networks and provide a more realistic insight into accuracy and dynamics of specific algorithms.

We first analyze the algorithms on LFR benchmark network and later compare selected algorithms for community detection on different networks. The selected networks

are the Zachary's karate club, Synthetic Network, Social Network with news tweets from X (former Twitter) with the keyword "Ukraine", the Citation Network from patents in the USA, and the Email Network of e-mail communication from a research organization in the EU. We compare the algorithms with each other on each of these networks and compare the structural properties of the identified communities using a proposed benchmarking analysis.

We ran the algorithms for each network separately, and we also ran each algorithm in 100 iterations in each network to identify any deviations in a larger sample that occur due to the operation of individual algorithms that use random selections when finding a solution and heuristics with which the algorithms try to reduce the search space for suitable solutions.

The experiment was done locally on Apple MacBook Pro M2 with 32GB of RAM. All results are then analyzed averaged for all runs. All algorithms were implemented using NetworkX and CDlib Python libraries.

The criteria on which we compare the algorithms are: the number of detected communities, the number of nodes in the largest identified community, the average community size, modularity, the average density, the average node degree, the average clustering coefficient, and the execution time of the algorithm.

Table 2: Networks and their properties

Network	**Number of nodes**	**Number of edges**
Neuroscience network	30	100
Zachary's karate club	34	78
Synthetic Network	150	416
X (Twitter)	4874	5418
EU network	265.214	420.045
Citation network	3.774.768	16.518.948

3 Results

3.1 Benchmarking Community Detection Algorithms on LFR benchmark Network

First we compare the algorithms on LFR benchmark graph, created using parameters as followed: n = 150, tau1 = 2.5, tau2 = 1.5, mu = 0.1, min_degree = 10, max_degree = 30, min_community = 20, max_degree = 40. The parameters ensure well separated communities in the generated network and ensuring connections between communities.

As seen in Figure 3 most of the algorithms correctly detect the number of communities, with exemption of Girvan-Newman which detected 4 communities. Also the Fast Label Propagation algorithms has due to its characteristics a bit of variability in the number of detected communities which varies from 5 to 6. The LFR benchmark network has 6 communities. Therefore, we see that on LFR benchmark network, which has specific structure so that communities are known upfront we can detect communities using most of the algorithms.

We can also see in Table 3 that most of the observed criteria on which we compare the algorithms are more or less the same. The observed results show that we success-

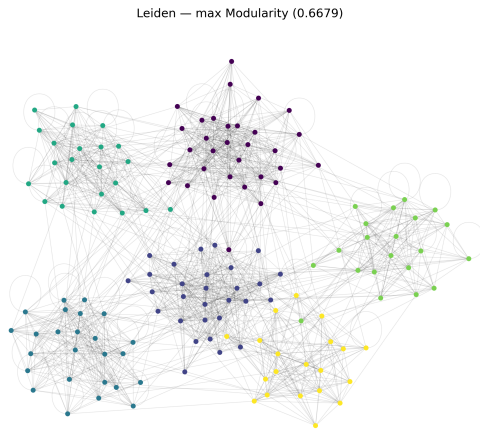

Figure 2: Detected communities in LFR network using Leiden algorithm

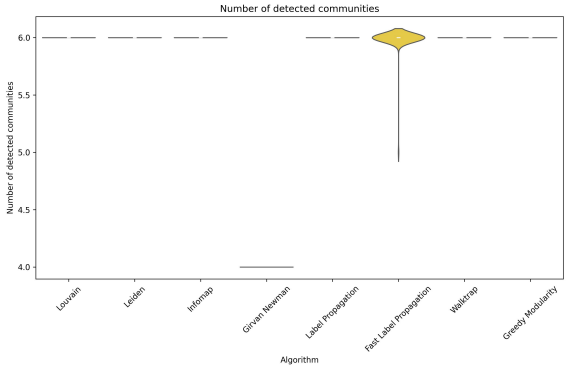

Figure 3: Detected communities by different algorithms in LFR network

fully identified communities in LFR benchmark network with exemption of the Girvan-Newman algorithm.

The first LFR benchmark network was in terms of size a small network, we now do the same analysis on a larger LFR benchmark network with 10.000 nodes. The parameters were as follows: n=10000, tau1=2.5, tau2=1.5, mu=0.2, min_degree=20,

Table 3: Comparison of community detection algorithms on the LFR network (n=150)

Method	**Num Communities**	**Max Size**	**Avg Size**	**Modularity**	**Execution Time (s)**
Louvain	6	32	25.0	0.668	0.0245
Leiden	6	32	25.0	0.668	0.0099
Infomap	6	32	25.0	0.668	0.0101
Girvan–Newman	4	84	37.5	0.476	6.91×10^{-6}
Label Propagation	6	32	25.0	0.668	0.0071
Fast Label Prop.	6	32	25.0	0.668	1.29×10^{-5}
Walktrap	6	32	25.0	0.668	0.0076
Greedy Modularity	6	31	25.0	0.659	0.1264

max_degree=80, min_community=100, max_community=300, seed=15, max_iters=300000.

As seen in Table 4 most of the algorithms also detected similar number of communities, we again see some variation with the Fast Label Propagation algorithm and Label Propagation as seen in Figure 4.

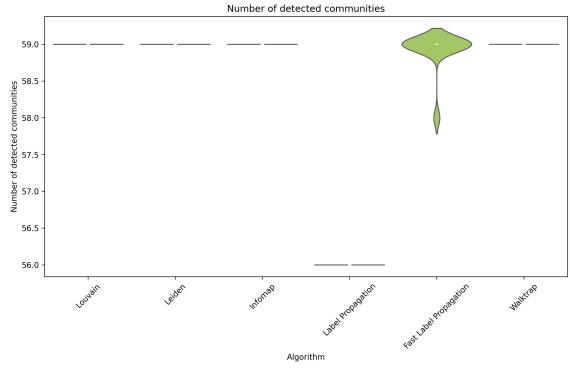

Figure 4: Detected communities by different algorithms in large LFR network

Table 4: Comparison of community detection algorithms on the LFR network (n = 10000)

Method	Num Communities	Max Size	Avg Size	Modularity	Execution Time (s)
Louvain	59	294	169.49	0.676	1.116
Leiden	59	294	169.49	0.676	0.918
Infomap	59	294	169.49	0.676	1.575
Label Propagation	56	501	178.57	0.675	0.230
Fast Label Propagation	59	294	169.49	0.676	6.20×10^{-6}
Walktrap	59	294	169.49	0.676	8.965

3.2 Benchmarking Community Detection Algorithms on different complex networks

In the following section we will now compare the algorithms on different complex networks and compare the results. Most of the work was originally done in master thesis research [14].

Table 5: Comparison of algorithms in terms of execution time (s)

Algorithm/Network	Zachary's karate club	Neuroscience Network	Synthetic Network	Social Network	E-mail Network	Citation Network
Louvain	0.0000999	0.0006111	0.0033162	0.4767473	8.3547758	4743.80048
Label Propagation	0.0001801	0.0002840	0.0012709	0.3739566	4.7337946	2291.033145
Fast Label Propagation	0.0000025	0.0000022	0.0000033	0.0000074	0.0000162	0.00231385231
Leiden	0.0010746	0.0009590	0.0033184	0.2472860	4.4387278	436.1735771
Infomap	0.0017194	0.0011369	0.0047150	1.9601560	104.2351924	/
Walktrap	0.0004096	0.0003394	0.0011552	0.8519429	/	/
Greedy Modularity	0.0018034	0.0010408	0.0094137	5.5470949	/	/
Girvan–Newman	0.0000017	0.0000017	0.0000030	/	/	/

In other networks the results show that the fastest algorithm in smaller networks is Girvan-Newman. In practice, there are no significant consequences regarding the choice of algorithm in smaller networks in terms of execution time. In medium-sized networks, the algorithms Girvan-Newman and Greedy modularity are not suitable, however the algorithms Leiden and Louvain give the best results.

Table 6: Comparison of algorithms in terms of modularity

Algorithm/Network	Zachary's karate club	Neuroscience Network	Synthetic Network	Social Network	E-mail Network	Citation Network
Louvain	0.58457	0.44077	0.76294	0.90862	0.79091	0.81151
Label Propagation	0.56520	0.30949	0.70496	0.81517	0.69716	0.57483
Fast Label Propagation	0.55826	0.37273	0.74724	0.80918	0.69585	0.61951
Leiden	0.58460	0.44490	0.76383	0.91144	0.80838	0.83249
Infomap	0.58300	0.43210	0.75960	0.84092	0.80796	/
Walktrap	0.58080	0.32316	0.75609	0.85014	/	/
Greedy Modularity	0.58280	0.41096	0.74609	0.90700	/	/
Girvan–Newman	0.57640	0.35806	0.64646	/	/	/

In larger networks, we did not obtain results for the algorithms Greedy modularity, Girvan-Newman and Walktrap, so they are not suitable for usage. The algorithms Louvain and Leiden perform most robustly in this case.

The algorithms detected different numbers of communities in all of complex networks. In LFR benchmark network the algorithms detected the same number of communities, with exception of Girvan-Newman algorithm. We observe deviations in the formation of community sizes (in Social Network, the Infomap algorithm creates an average community size of 175.54 nodes, while the Walktrap algorithm creates as many as 1613.00). In the largest network considered, the Leiden algorithm executes noticeably faster than Louvain (Louvain:4743.800s, Leiden: 436.173s). In medium and large networks, the Label Propagation algorithm creates more communities, while the Louvain algorithm creates fewer communities. In smaller networks, we did not find a pattern that could help us decide on the choice of algorithm based on the number of detected communities.

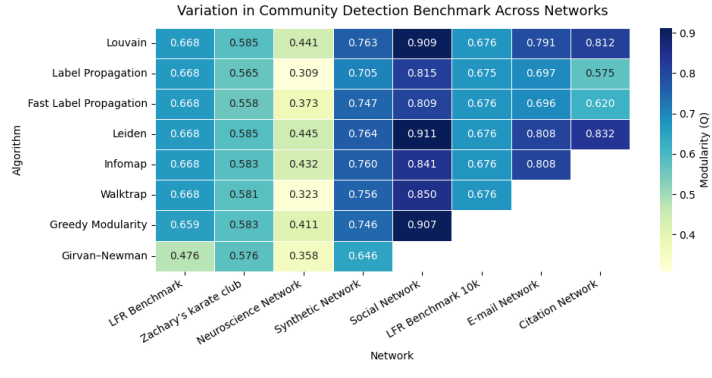

Figure 5: Variation of modularity in different networks

We observed that the algorithms are more or less very accurate in terms of analysis of LFR benchmark network, therefore it can be concluded that such algorithms are suitable for usage in networks of such structure. However we still need to take into account the comparison of algorithms in other complex networks where the results were very different.

It was shown that Leiden and Louvain algorithms had identical modularity values in LFR network, however the modularity was different on larger complex networks. This implies that the performance of this algorithms on synthetic networks cannot be implied on real networks.

In Figure 5 we see how the variation of modularity changes with different networks, especially noticeable is the comparison of LFR networks, the first LFR networks which is a lot smaller in terms of the LFR Benchmark 10k which has 10.000 nodes the modularity

is similar. However when we compare the results to real networks the modularity for algorithm changes when comparing different networks.

4 Discussion

We have shown that if we apply community detection on a LFR benchmark network most of the algorithms successfully detect the correct number of communities, however when we apply the algorithms on other complex networks, we see that the results are very much different due to the nature of specific algorithms and modularity optimization of algorithms.

We have also shown that if we benchmark algorithms on LFR benchmark network, we can successfully detect communities on all of the algorithms with exemption of Girvan-Newman, based on that we could conclude that most of the algorithms are suitable for community detection, however when we apply the same algorithms on complex networks we see that the results are very much different depending on each algorithm. Therefore we should benchmark algorithms in such way that we apply all algorithms on different networks, so that we can observe how specific algorithm acts in different network size and structure.

Results in this research are consistent with previous work on benchmarking community detection [11] where it was already noted that synthetic graphs are not reflecting the structure and characteristics of real networks. So further more algorithms act differently in LFR networks compared to real networks.

No algorithm performed optimal in all networks, each algorithm offers advantages in specific network size and structure. With that in perspective benchmarking community detection algorithms should have a multi dimensional approach using multiple networks with diverse structure.

References

[1] Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment, 2008*(10), P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008

[2] Christensen, A., Garrido, L., Guerra-Peña, K. & Golino, H. (2024). Comparing community detection algorithms in psychological data: A Monte Carlo Simulation. Behavior Research Methods, 56, pp. 1485–1505. https://doi.org/10.3758/s13428-023-02106-4

[3] Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Social Networks, 1*(3), 215–239. https://doi.org/10.1016/0378-8733(78)90021-7

[4] Newman, M. E. J. (2012). Communities, modules and large-scale structure in networks. *Nature Physics, 8*, 25–31. https://doi.org/10.1038/nphys2162

[5] Mrvar, A. (2014). *Analiza omrežij s programom Pajek* [Online resource]. Retrieved from `http://mrvar.fdv.uni-lj.si/sola/info4/uvod/mrv3a.pdf`

[6] Fortunato, S. (2010). Community detection in graphs. *Physics Reports, 486*(3–5), 75–174. https://doi.org/10.1016/j.physrep.2009.11.002

[7] Fortunato, S., & Hric, D. (2016). Community detection in networks: A user guide. *Physics Reports*, 1–44. https://doi.org/10.1016/j.physrep.2016.09.002

[8] Fagnan, J., Dubé, L. J., & Savaria, Y. (2018). Modular networks for validating community detection algorithms. *arXiv preprint arXiv:1801.01229*. https://arxiv.org/abs/1801.01229

[9] Newman, M. E. J. (2004). Fast algorithm for detecting community structure in networks. *Physical Review E, 69*(6), Article 066133. https://doi.org/10.1103/PhysRevE.69.066133

[10] Newman, M. (2006). Modularity and community structure in networks. Proc. Natl. Acad. Sci. U.S.A. 103, pp. 8577–8582. https://doi.org/10.1073/pnas.0601602103

[11] Lancichinetti, A., Fortunato, S., & Radicchi, F. (2008). Benchmark graphs for testing community detection algorithms. *Physical Review E, 78*(4), 046110. https://doi.org/10.1103/PhysRevE.78.046110

[12] Menczer, F., Fortunato, S., & Davis, C. A. (2020). *A first course in network science*. Cambridge: Cambridge University Press.

[13] Orman, G. K., Labatut, V., & Cherifi, H. (2013). Towards realistic artificial benchmark for community detection algorithms evaluation. *International Journal of Web Based Communities, 9*(3), 349–370. https://doi.org/10.1504/IJWBC.2013.054908

[14] Pritržnik, R. (2025). *Detekcija skupnosti v kompleksnih omrežjih* [Magistrsko delo]. Fakulteta za informacijske študije v Novem mestu.

[15] Pons, P., & Latapy, M. (2006). Computing communities in large networks using random walks. In *Computer and Information Sciences – ISCIS 2005*, Lecture Notes in Computer Science (Vol. 3733, pp. 284–293). Springer. `https://doi.org/10.1007/11569596_31`

[16] Rosvall, M., & Bergstrom, C. T. (2008). Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences, 105*(4), 1118–1123. https://doi.org/10.1073/pnas.0706851105

[17] Rossetti, G., Milli, L. & Cazabet, R. (2019). CDlib: a Python Library to Extract, Compare and Evaluate Communities from Complex Networks. Applied Network Science, 4, 52. https://doi.org/10.1007/s41109-019-0165-9

[18] Schult, D., Swart, P. & Hagberg, A. (2008). Exploring network structure, dynamics, and function using NetworkX. V Proceedings of the 7th Python in Science Conference (SciPy2008), str. 11–15.

[19] Tokala, S., Enduri, M., Lakshmi, J., & Hajarathaiah, K. (2025). Evaluating community detection algorithms: A focus on effectiveness and efficiency. *Journal of Scientometric Research, 14*(1), 62–74. https://doi.org/10.5530/jscires.14.1.6

[20] Traag, V. A., & Šubelj, L. (2023). Large network community detection by fast label propagation. *Scientific Reports, 13*, 1746. https://doi.org/10.1038/s41598-023-29610-z

Global Electric Circuit as a Driver of Space Weather Impacts: Cross-Sectoral Risks for Energy and Digital Infrastructures with a Spain Blackout Case Study

Valerij Grasic
Telekom Slovenije
Cigaletova 17, 1000 Ljubljana, Slovenia
valerij.grasic@telekom.si

Biljana Mileva Boshkoska
Faculty of Information Studies
Ljubljanska cesta 31A, 8000 Novo mesto, Slovenia
Jožef Stefan Institute
Jamova cesta 39, 1000 Ljubljana
biljana.mileva@fis.unm.si

Abstract: *Space weather is usually evaluated through large-scale geomagnetic disturbances, particularly coronal mass ejections (CMEs) and storm indices such as Kp and Dst. However, disruptive events can also arise when these parameters remain quiet, suggesting additional mechanisms. This paper introduces the Global Electric Circuit (GEC) as a framework to explain such cases, showing how changes in ionospheric conductivity, total electron content (TEC), and radiation flux can influence terrestrial infrastructures. The first contribution is to highlight the GEC as a driver of space weather impacts, extending existing models beyond CME and geomagnetic indices. The second is to develop a cross-sectoral risk perspective that traces how GEC-related disturbances affect both energy and digital infrastructures, creating cascading vulnerabilities. The approach is evaluated using the 2025 Spain blackout, when widespread disruptions occurred despite the absence of major CME activity. Observational data show anomalies in ionospheric and atmospheric conditions consistent with GEC-driven processes. These disturbances coincided with fluctuations in photovoltaic output, grid instability, and communication interruptions. The paper also proposes methodological guidelines, recommending multi-scale analysis windows (4 hours, 16 hours, 3–7 days) and the integration of multi-source datasets. These include upstream satellite observations at the L1 point, GNSS-derived TEC and ionosonde data, atmospheric reanalysis and pressure fields, ground magnetometer networks, and infrastructure-level energy and digital data. The findings demonstrate that incorporating GEC into space weather studies and explicitly linking energy and digital sectors provides a stronger basis for both scientific research and practical resilience planning.*

Key Words*: Smart City, Space weather, Global Electric Circuit (GEC), Energy sector, Digital sector, Spain blackout, resilience*

1 Introduction

Critical infrastructures such as energy and digital networks form the backbone of modern

societies, yet their interdependence makes them vulnerable to external disturbances. Space weather is one of the most prominent of these hazards, typically assessed through coronal mass ejections (CMEs), solar flares, and geomagnetic indices such as Kp and Dst. However, not all disruptions can be explained by these parameters. The 2025 Spain blackout is a case in point: major power and communication failures occurred during a period of low geomagnetic activity.

This paper advances two contributions. First, it introduces the Global Electric Circuit (GEC) as a mechanism by which atmospheric conductivity, ionospheric anomalies, and radiation flux affect terrestrial infrastructures even under "quiet" solar conditions. Second, it develops a cross-sectoral framework for analysing space weather impacts, showing how energy and digital infrastructures interact and amplify vulnerabilities.

The paper outlines a conceptual model of GEC and its pathways, evaluates the Spain blackout as a test case, and situates the findings within a broader discussion of risk assessment, resilience, and data requirements for critical infrastructure protection.

2 Problem definition

Smart cities integrate traditional networks with digital technologies to improve services [1], aiming for intelligent, adaptive, and flexible systems [2, 3, 4]. Examples such as intelligent call-handling [5] show benefits, but deeper interdependence also increases vulnerability. Failures in energy, health, or communications can cascade across sectors, as noted by the Endurance Project [6].

Infrastructure research identifies four interconnection types—physical, cyber, geographic, and logical—through which failures propagate [7]. Resilience theory later reframed this challenge, stressing robustness and adaptability [8, 9]. At the policy level, the CER Directive (EU 2022/2557) and NIS2 (EU 2022/2555) strengthen resilience and cybersecurity. ESA's Space Situational Awareness (SSA) initiative [10] positions Europe to build resilience by acting as an E2E space weather hub, linking solar activity with terrestrial impacts on critical infrastructure (CI), and making the GEC–CI connection explicit in operational forecasting.

Space weather is typically assessed via CMEs, Kp, and Dst, and their role in driving GICs [11], but this framework misses key risks: silent events during quiet conditions, overlooked parameters such as conductivity, TEC, or radiation flux [12], and sectoral isolation despite energy–digital interdependence [7].

The 2025 Spain blackout showed widespread failures in energy and communications under low geomagnetic activity [13]. This underscores the need for broader monitoring and a real-time model for smart cities that integrates space weather with cross-sector resilience.

3 Conceptual Framework: GEC and cross-sectoral risk

Earth's geomagnetic field extends from the core to the solar wind, shielding the atmosphere from charged particles [14]. It is shaped by both internal and external processes [15; 16], with the magnetosphere–solar wind interaction producing variable space weather [11]. Key monitoring parameters include solar wind velocity, IMF orientation, proton flux, plasma density, and geomagnetic indices such as Kp and Dst, but these do not capture all impacts.

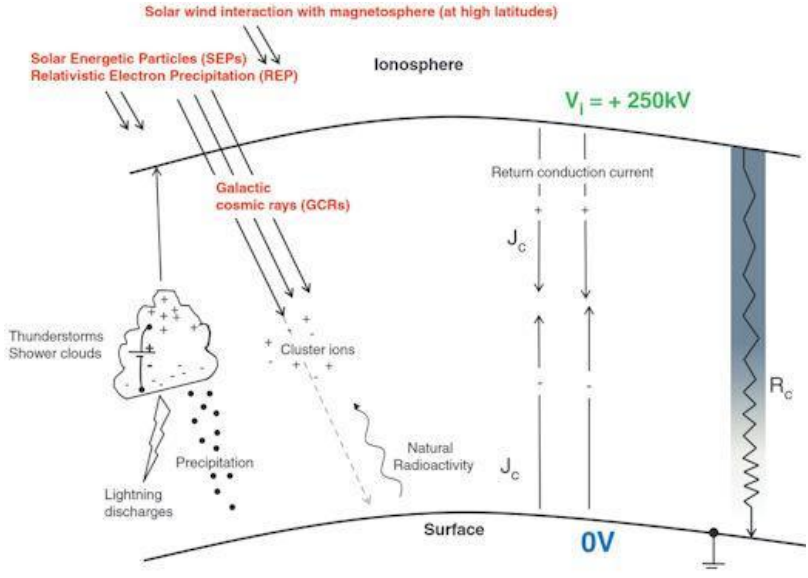

Figure 1: Global Electric Circuit (GEC) [12]

Figure 1 shows structure of Global Electric Circuit (GEC) [12]. GEC links ionosphere, atmosphere, and ground [12], influenced by thunderstorms, cosmic rays, and solar forcing [11]. Perturbations may arise even without CME, altering conductivity and atmospheric resistance. The GEC thus provides a transmission pathway for disturbances, propagating downward to affect ground electromagnetic conditions and infrastructures.

The F region of the ionosphere (split into F1 and F2 layers during daytime) plays a central role in conductivity and radio propagation. Variations in its critical frequency (foF2) and altitude (hmF2) reflect both solar and dynamical influences. Also, Total Electron Content (TEC) is the key diagnostics. Anomalies appear under both storm and quiet conditions: foF2 disappearance over Europe in May 2024 [17], weakened ionisation during Forbush decreases [18], enhanced atmospheric electric fields from proton events [19], and quiet-time fluctuations such as “space hurricanes” [20]. These findings highlight the need to consider GEC dynamics alongside traditional CME/Kp/Dst models.

3.1 Energy and digital infrastructures as coupled systems

Energy and digital infrastructures are deeply interdependent [7]: energy systems depend on digital control and communication, while digital systems require stable electricity. Failures propagate rapidly between them, producing cascading effects. To address this, the framework distinguishes sector-specific layers, their interdependencies, and a system-level view (Table 1).

Smart cities integrate energy, transport, water, and communication through digital platforms. This enhances efficiency but multiplies failure points. Space weather impacts mediated by the GEC can cascade across energy grids (transformers, PV, distribution), digital services (data centres, mobile networks), and dependent sectors such as transport and emergency systems. Smart cities thus act as amplifiers of vulnerability, making resilience planning essential.

A parameter-based monitoring model [21] shows how real-time magnetic data

combined with contextual and space weather inputs can capture and classify disturbances. Future work should develop high-resolution datasets of GEC exposure, establish correlations between infrastructure disruptions and heliophysical parameters, and apply AI-driven models for early warning and decision support.

Table 1: Sectoral and cross-sector impact mapping

Layer	Key Indicators	Data Sources	Infrastructure Layers
Energy	Grid outages, load fluctuations, and restoration time	Utility reports, sensor data, citizen feedback, flood maps, weather stations, and historical outage data	Power plants, substations, grid topology, priority loads (hospitals, data centres), backup systems
Digital	Network downtime, data anomalies, service availability and degradation	ISP updates, social media, app logs, unstructured reports, and historical downtime data	Fiber lines, base stations, copper/optical/wireless technologies, end-user service layers (TV, internet, phone)
Cross-Sector	Shared node failures, cascading disruptions	Combined event logs, linked outages, and infrastructure dependency maps	Interlinked assets (e.g., powered base stations), cross-sector workflows

4 Sectoral and cross-sector impact mapping: the Spain blackout case study

The analytical analysis proceeds in five steps:

1. Theoretical pathways of influence – Reviewing how CME, solar wind velocity, proton flux, plasma density, interplanetary magnetic field (IMF) orientation, Kp and Dst indices, and ground conductivity conditions translate into geomagnetically induced currents (GICs) and potential grid disruptions.
2. Event-specific data analysis – Gathering solar and geomagnetic parameter datasets for the period of the Spain blackout to assess potential triggers.
3. Propagation to infrastructures – Mapping how variations in space weather parameters could have influenced electricity and digital networks in Spain.
4. Cross-sectoral evaluation – Correlating space weather parameters with available operational data from grid operators and digital/telecommunication providers.
5. Resilience and risk implications – Deriving theoretical and practical insights for risk assessment, preparedness, and resilience enhancement.

4.1 Blackout event description

The blackout in Spain happened at 12:33 CEST on April 28, 2025. Since CEST is UTC plus 2 hours, that makes the blackout time at 10:33 UTC (April 28, 2025). In the rest of the paper, the UTC time is used.

The blackout sequence on the Iberian Peninsula (April 28, 2025) developed in

distinct phases, marking a period of instability. Figure 2 shows grid instability at that time:

- 08:30 UTC: first signs of voltage variability.
- 10:03–10:07 UTC: forced oscillation (~0.64 Hz) detected.
- 10:19–10:22 UTC: second oscillation period, stronger amplitude.
- 10:33:18–10:33:24 UTC: cascading failures and system collapse, initiating the blackout.

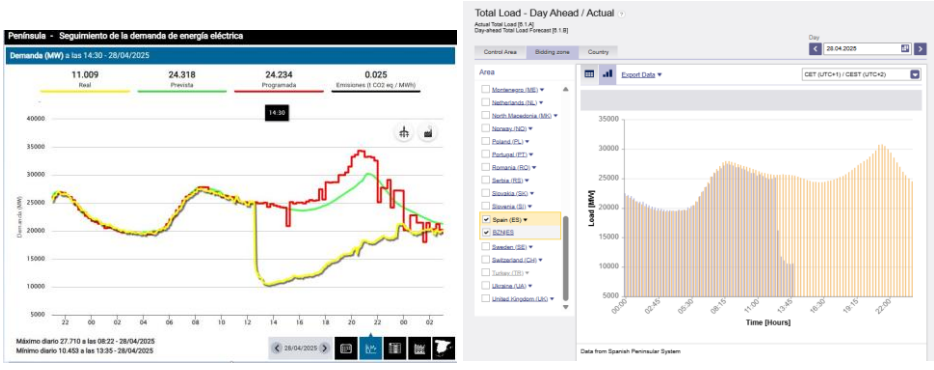

Figure 2: Grid instability

4.2 Hazard context, with energy and digital sector key points

The Table 2 shows up to 15 time points, from early grid stress to full restoration. It shows key impacts on the energy and digital/telco sectors, highlighting how disturbances evolved and cascaded across infrastructures. The information sources are retrieved from [22-27]; then processed using the LLM [28].

Table 2: Hazard context with energy and digital sector key points

Time (UTC)	Where / Region(s) Affected	Key Hazard Context / Trigger / Risk	Energy Key Points	Digital / Telco Key Points
~9:00-10:00	Spain (nation wide)	Low demand, high solar generation; system already under voltage stress in the transmission network.	Voltage levels in the 400 kV grid begin to fluctuate; synchronous generators are having trouble absorbing reactive power.	Digital networks are currently operational; backup systems have not yet been engaged. No significant reports have been received yet of digital outages.
10:03-10:07	Iberian grid (Spain & Portugal)	First low-frequency oscillation (0.6 Hz) local mode between generation clusters.	Operators applied countermeasures, including switching lines and changing HVDC links to fixed-power mode. Voltage oscillations observed.	No major digital disruptions have been reported yet; telecom infrastructure remains operational.

10:16-10:22	Spain / Iberian grid border with Europe	Inter-area oscillation (~0.2 Hz) between the Iberian Peninsula and the rest of Europe. Risk of instability rising.	Further reactive power instability, rising voltage, and fluctuations are visible in 400 kV and 220 kV lines.	Still functioning; telecoms are likely beginning to operate on backup in some critical sites, but no reports have yet been received of a widespread outage.
10:26	Southern Spain	Request to connect additional thermal power plant(s) for voltage control; scheduled too late.	The thermal generator has been asked to synchronise; however, the voltage is still rising, and the system's capacity for reactive support is low.	Telecom sites are still mostly operational; there may be increased stress on backup power, and monitoring is starting.
10:32:57-10:33:17	Southwestern Spain (Granada, Badajoz, Seville)	Sudden loss of generation: ~827-1000 MW from PV, wind, etc. Multiple generation trips in seconds.	Generation drop triggers frequency drop, voltage rise, loss of imports/export ties, and quick cascading failures.	Immediate digital disruption begins: many power-dependent telecom nodes lose supply; mobile/Internet users in affected zones lose service. Traffic drops sharply.
~10:33:19-10:33:24	The entire Iberian Peninsula grid	AC tie lines with France severed due to loss of synchronism; full grid collapse; automatic load shedding activates.	Complete collapse; power drops to zero in many zones; black-start process begins.	Widespread telecom failure: large portions of fixed & mobile networks are down; emergency services in some regions are cut off or operating on backup; internet traffic is ~25% of baseline in core services.
10:35	Affected zones begin black-start	Recovery starts in select plants; interconnections begin to be re-energised.	Morocco-Spain interconnector re-energised by ~13:04; other tie-lines gradually restored.	Some telecom infrastructure in regions where electricity is restored starts to come back online; mobile data/voice are partially

				restored in those zones.
11:00-12:00	Spain / Portugal	The power restoration phase involves bringing back black start plants, substations, and interconnectors.	The transmission system is largely restored, but with constraints and risk of voltage/frequency instability. High reactive power demands.	Digital networks are gradually recovering, but many users remain offline; battery backups are running out in rural / less critical sites.
Evening (~18:00)	Spain & Portugal broadly	Secondary disruptions from backup failures and telecommunications peak disruptions.	Grid much of the supply restored; critical areas operational; system still under monitoring.	Mobile networks had the worst impact around this time ("many areas reporting 50–100% of users offline") as backup power at sites depleted.
~19:00-21:00	Spain (major urban areas)	Evening surge in network demand; many telecom nodes still recovering; backup systems strained.	The energy system is mostly stable; the lights are back on, and full capacity is being restored.	Vodafone reports that ~50% of nodes are active & ~60% of mobile traffic has been restored by 11 pm.
Next morning / early hours (April 29)	Entire Spain & Portugal	Complete restoration of the electric grid and network infrastructure.	Transmission system restored (Spain by ~04:00, Portugal somewhat earlier) per ENTSO-E timeline.	Digital network traffic is nearly back to normal; fixed and mobile services are mostly restored, with some isolated longer outages remaining.

4.3 Key space weather parameters (Kp, Dst, AE)

All the data used for the blackout evaluation are present at GitHub portal [29].

Evaluation of the Kp index (Figure 3): There has been no high values for Kp index. However, (as seen by GFZ), there has been a 1-hour level 4 burst (Hp30, also Hp60) in the morning, and two 30-minute bursts (Hp30) at the evening a day before.

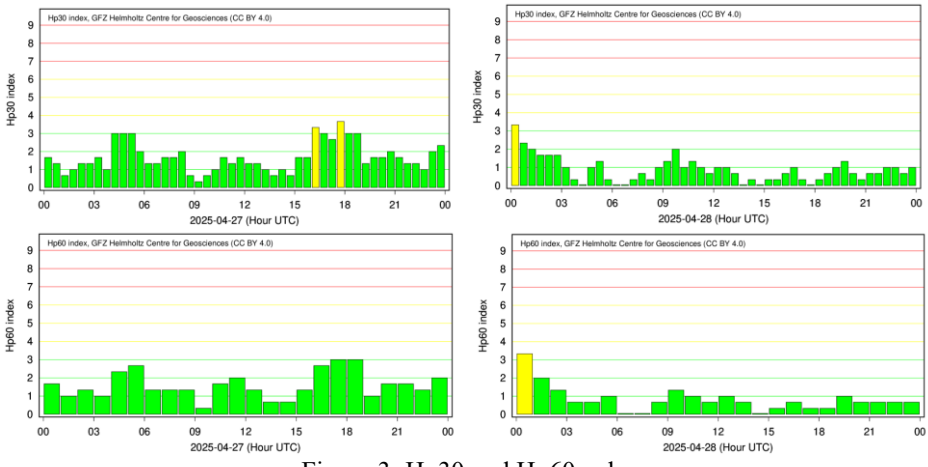

Figure 3: Hp30 and Hp60 values

Evaluation of Dst (Figure 4): There has been no major values regarding Dst (Kyoto).

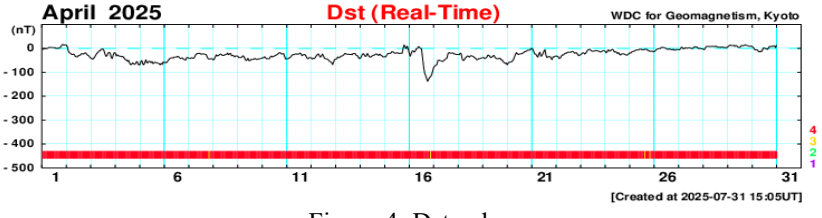

Figure 4: Dst value

Evaluation of AE/AL/AU/AO (Figure 5): There are seen no high values for AE indices (Kyto), moreover, they are minimal compared to the values days before or after event.

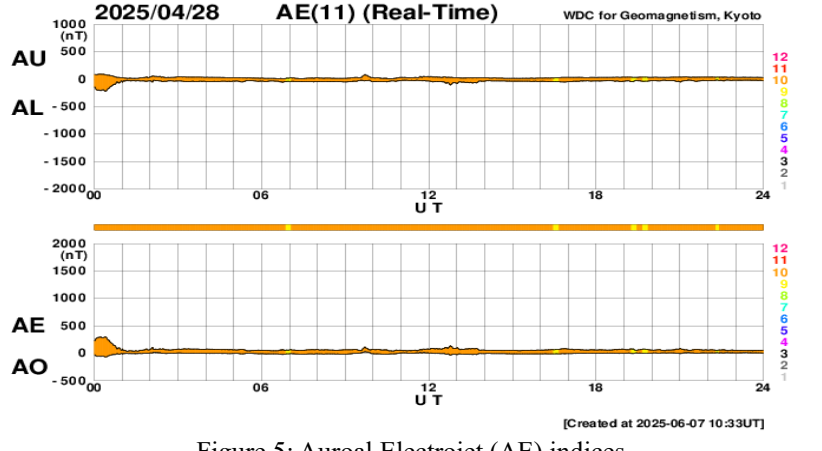

Figure 5: Auroal Electrojet (AE) indices

4.4 Other parameters

Magnetometers all over the globe (Figure 6): It can be seen some disturbance (Kyoto), affecting and moving around the globe.

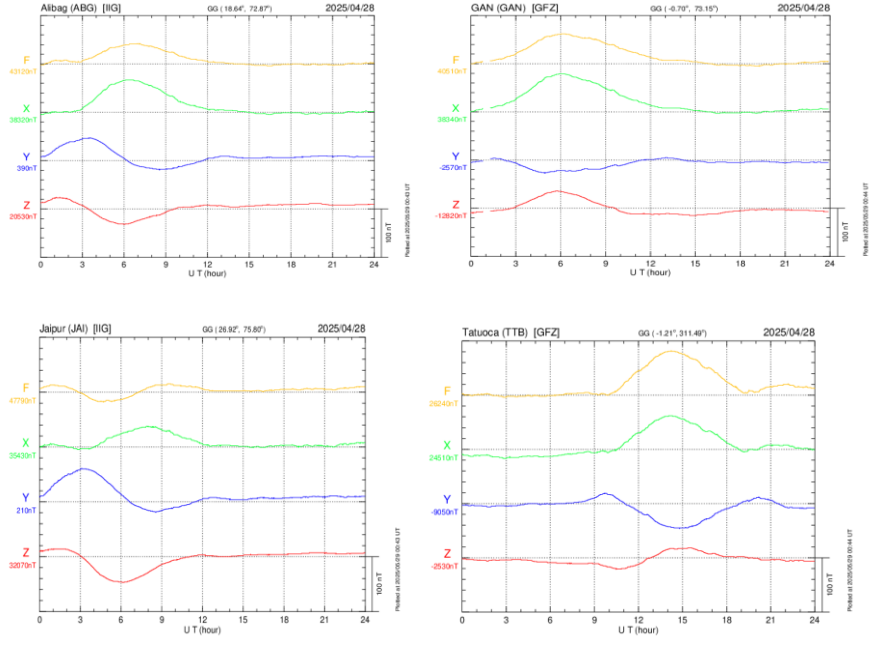

Figure 6: Magnetometers around the world

Magnetometers in Spain (Figure 7): Presented is magnetometer EBR in Spain, that is Ebre/Horta de Sant Joan. Instability is seen both at X and Y component of the magnetic field. At Y, from 08:30 UTC to 12:00 UTC the component decreases, by 60 nT (from 683 nT to 623 nT), and from 08:30 UTC to 10:30 UTC, by 40 nT (from 683 nT to 643 nT).

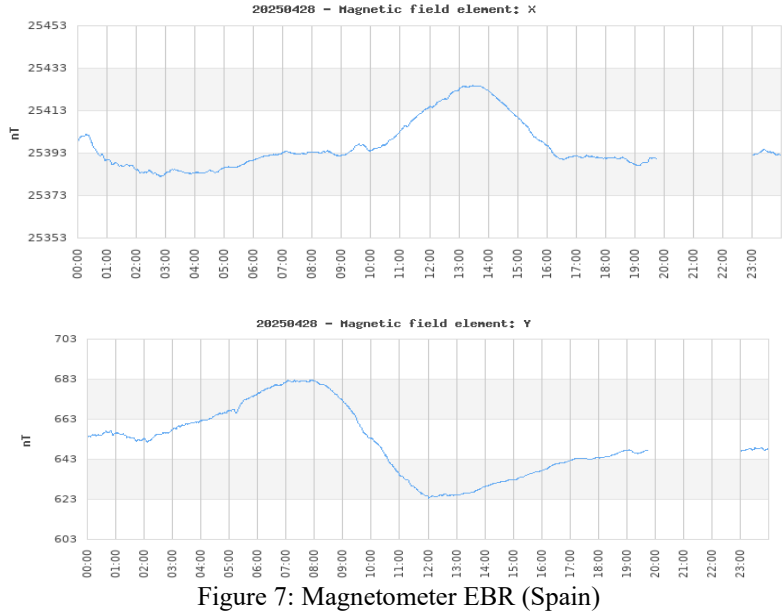

Figure 7: Magnetometer EBR (Spain)

Space Situational Awareness: Space Situational Awareness services (PROBA2 [30] satellites) show these events:

Date	Start	Peak	Stop	Flare class	Location	NOAA region
2025-04-28	05:46:00	06:02:00	06:21:00	C2.2	//	4069
2025-04-27	22:38:00	22:44:00	22:46:00	C2.1	//	4064
2025-04-27	22:20:00	22:29:00	22:38:00	C1.5	//	//
2025-04-27	18:37:00	18:49:00	19:04:00	C2.1	N08W72	4064
2025-04-27	17:17:00	17:24:00	17:30:00	C1.4	S14W26	4070
2025-04-27	08:18:00	08:28:00	08:39:00	C1.3	//	//

5 Discussion: risk and resilience

5.1 Lessons from the Spanish blackout

Applying the GEC framework to the April 2025 blackout in Spain shows how space-weather impacts can emerge even when global indices remain quiet. The framework brings together solar background activity (with daily sunspot numbers in the 110–130 range during April, and F10.7 flux near 150 sfu at the time of the blackout, reaching ~170 sfu in preceding days), persistent anticyclonic weather (with high pressure) over Iberia, and localized geomagnetic anomalies. This combination highlights pathways of influence that remain invisible in CME- or Kp/Dst-based approaches.

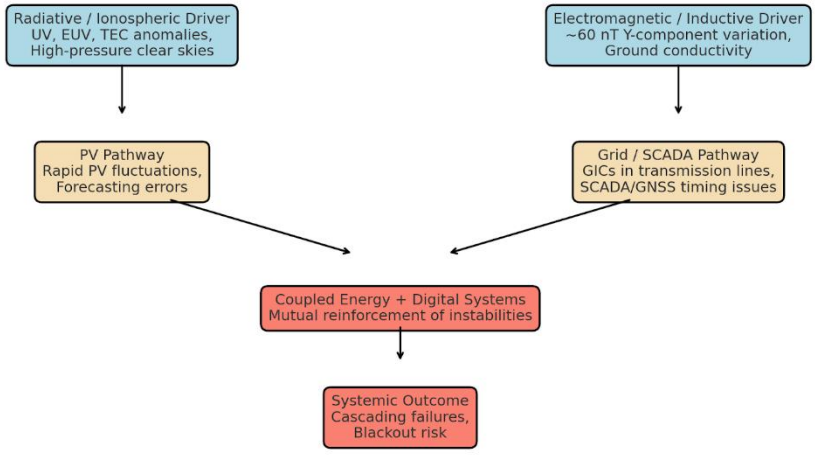

Figure 8: Dual Mechanisms of the Spain Blackout: Radiative/Ionospheric Effects on PV and Inductive/Electromagnetic Effects on Grid and SCADA

Preliminary inspection of available data suggests that two mechanisms may have coexisted. Radiative and ionospheric variability, amplified by clear-sky conditions and elevated solar activity, plausibly produced rapid fluctuations in photovoltaic generation. At the same time, localized geomagnetic variations of 40–60 nT in the horizontal field

(Figure 7) point to short bursts of **|dB/dt|**, sufficient to induce currents in long transmission corridors and disturb SCADA synchronization.

Magnetometer records (Figure 7) confirm rapid swings in the horizontal field—most notably a 60 nT excursion in the Y (eastward) component near the outage, with smaller X (northward) changes appearing later. Under Faraday's law, it is the time derivative of the field (**|dB/dt|**), not its absolute magnitude, that drives geoelectric induction at the surface. Directionally, a sharp change in Y tends to generate east–west electric fields, which couple most strongly into north–south conductors; conversely, X variations favor north–south electric fields that stress east–west lines. Because Spain's grid contains both N–S and E–W corridors, localized bursts in either component can plausibly drive quasi-DC geomagnetically induced currents (GICs), leading to transformer half-cycle saturation, harmonics, and timing disturbances in SCADA and GNSS-based systems.

These results do not constitute proof, but they reach a plausibility threshold: minute-scale **|dB/dt|** bursts in Y (and later X) coincide with the operational window and are of sufficient magnitude, given typical European impedances, to induce mid-latitude geoelectric fields. Combined with the independent PV/ionospheric pathway, this strengthens the case for a dual-mechanism interpretation of the Spain blackout. The implication is clear: future work should routinely compute **|dX/dt|** and **|dY/dt|**, assess the directional exposure of major corridors (N–S vs. E–W), and integrate these with PMU and neutral-current telemetry to evaluate grid sensitivity under "quiet-index" conditions.

Figure 8 illustrates these dual pathways—radiative/ionospheric impacts on PV and inductive/electromagnetic stress on the grid—showing how both are supported by observed anomalies and consistent with theoretical models of GEC coupling and geomagnetically induced currents.

In summary, the Spain blackout demonstrates how multiple mechanisms can operate simultaneously. Elevated solar activity and clear-sky conditions made PV output unusually sensitive to radiative fluctuations, while localized **|dB/dt|** bursts provided a plausible inductive pathway into the transmission grid and control systems. The coexistence of these two drivers offers a broader explanatory framework than storm indices alone. This lesson underscores the need to expand space-weather risk assessment beyond CME arrivals and global geomagnetic indices, toward multi-parameter, multi-sector monitoring that can capture hidden vulnerabilities under "quiet" conditions.

5.2 Methodological guidelines

To support structured evaluations, the Table 3 proposes four analysis windows (W1, W2, W3, W4), from short-term anomalies to week-scale contexts. Data sources can be from satellite observations (TEC, UV, conductivity), atmospheric reanalysis modes or operational logs (e.g., ENTSO-E).

Beyond temporal windows, comprehensive evaluation requires a multi-source parameter set integrating solar, geomagnetic, ionospheric, atmospheric, and infrastructure data. Table 4 shows multi-layer parameters and data sources for testing and monitoring space weather impacts across solar, ionospheric, atmospheric, ground, and infrastructure domains.

Solar and L1 point satellites (e.g., DSCOVR, ACE) provide upstream warnings with 30–60 minutes lead time. Ground magnetometers, such as INTERMAGNET stations and national observatories like the Ebro Observatory in Spain, are key for detecting local dB/dt variations. UV flux and solar duration influence both photovoltaic output and the ionospheric F1/F2 layers.

Table 3: Suggested analysis windows

W	Time scale	Purpose	Typical Parameters	Example Data Sources
W1	4 hours	Capture immediate anomalies and triggers	TEC fluctuations, UV spikes, and conductivity shifts	Satellite TEC data, GNSS error logs
W2	16 hours	Identify diurnal variations and daily cycles	Ionospheric conductivity, grid load data	Reanalysis models, ENTSO-E daily logs
W3	3 days	Assess short-term coupling between space weather and infrastructures	TEC trends, atmospheric pressure systems	Copernicus datasets, NOAA satellite records
W4	7 days (before and after)	Place events in a broader temporal context, detect preconditioning	Cumulative TEC anomalies, solar wind parameters	Multi-satellite datasets, geomagnetic observatories

Atmospheric pressure and the atmospheric electric field (AEF) link to the GEC and conductivity shifts. Satellite constellations and GNSS measurements track TEC, while ionosondes provide foF2 and hmF2. Global geomagnetic indices are available from Kyoto WDC.

Table 4: Multi-layer parameters and data sources for testing and monitoring space weather impacts

Layer	Parameters	Purpose / Typical Sources
Solar / Upstream	Solar wind velocity, plasma density, proton flux, IMF orientation (Bx, By, Bz), dynamic pressure, CME arrival, solar flares, X-ray/UV/EUV flux, solar irradiance, sunspot number	Early warning (15–60 min) from L1 satellites (DSCOVR, ACE, SOHO, Solar Orbiter); ground-based solar telescopes and radiometers for irradiance & sunspot monitoring
Geomagnetic	Kp, Dst, Ap indices; ground magnetic field variations (dB/dt)	Global indices (Kyoto WDC); regional/local magnetometer networks (INTERMAGNET, Ebro Observatory, IMAGE network)
Ionospheric	GNSS-derived TEC, foF2, hmF2, foE, foF1, Spread-F, scintillation indices (S4, σφ)	GNSS receiver networks, ionosondes, ESA/NOAA satellite instruments (Swarm, COSMIC-2)
Atmospheric / GEC	Atmospheric electric field (AEF), columnar conductivity, cosmic ray flux, Forbush decreases, surface pressure, temperature, humidity, cloud fraction, aerosol optical depth, sunshine duration, UV index at ground	Ground-based AEF sensors, neutron monitors (cosmic rays), meteorological networks, Copernicus/ECMWF reanalysis, pyranometers/UV sensors
Energy Sector	PV output and irradiance variability, grid load, frequency, ROCOF, voltage deviations,	TSO/DSO telemetry (ENTSO-E, Red Eléctrica, ELES, etc.), PMUs, PV fleet monitoring

	transformer/substation anomalies, GIC measurements (if available)	
Digital Sector	GNSS positioning/timing errors, communication latency, HF/VHF/UHF outages, packet loss, SCADA timing anomalies, data center performance logs	Telecom operators, GNSS monitoring services, critical infrastructure operators
Contextual / Ground	Ground conductivity/resistivity maps, geology, altitude, proximity to coasts or rivers (affecting GIC propagation)	Geological surveys, national mapping agencies, EPRI conductivity models

5.3 Implications for smart cities and integrated monitoring for the resilience framework

In smart cities, where energy, transport, healthcare, and communication systems are deeply interconnected, efficiency is paired with vulnerability. A disturbance in one sector can cascade quickly across others.

Space weather disturbances mediated by the Global Electric Circuit (GEC) exemplify this risk. Even without major geomagnetic storms, shifts in ionospheric conductivity or atmospheric electric fields can disrupt electricity, telecommunications, and transport at once—putting critical services such as hospitals and emergency response at risk.

The Spanish blackout revealed this dynamic: anomalies in photovoltaic generation and grid operations coincided with communication failures, producing mutually reinforcing disruptions.

Resilience therefore depends on integrated monitoring and governance. Forecasts must incorporate GEC parameters, and cross-sectoral data—energy, digital, and ground observations—should be combined to identify risks. Embedding space weather into urban resilience strategies is essential to maintain essential services in highly connected environments.

6 Conclusion

The Spanish blackout highlights the need to integrate space weather into infrastructure risk assessments. Impacts cannot be understood through CME events and geomagnetic indices alone. By incorporating the Global Electric Circuit, a fuller picture emerges of how ionospheric and atmospheric processes affect terrestrial systems.

Equally important is the cross-sector perspective: energy and digital infrastructures are tightly linked, and disturbances in one rapidly cascade to the other. The blackout illustrated this through simultaneous anomalies in photovoltaic generation, grid operation, and communications.

Although full quantitative modelling remains for future work, this paper offers initial guidelines for resilience analysis: multi-scale temporal windows and the combined use of satellite, atmospheric, and ground-based observations.

In sum, the paper advances two contributions: (1) the GEC as a framework for impacts beyond CME/Kp/Dst paradigms, and (2) a cross-sectoral model of risk and resilience spanning energy and digital domains. Both are essential for preparing societies to withstand complex space weather disruptions.

7 Acknowledgements

The research presented in this paper is part of the ENDURANCE project's research activity, which is funded by the European Union's Horizon Europe research and innovation program under Grant Agreement No. 101168007.

8 References

[1] European Commission. Smart Cities - Smart Living. Shaping Europe's digital future, 2019, https://ec.europa.eu/digital-single-market/en/smart-cities, downloaded: July 7th 2019.
[2] Komninos, N. The age of intelligent cities: smart environments and innovation-for-all strategies (First Edition). Routledge, Taylor & Francis Group, New York, 2015.
[3] Picon, A. Smart Cities: A Spatialised Intelligence. Wiley, Chichester, England, 2015.
[4] Department for Business innovation & Skills. Global Innovators: International Case Studies on Smart Cities Smart Cities Study (Bis research paper no. 135), 2013, https://www.gov.uk/government/publications/smart-cities-international-case-studies-global-innovators, downloaded: July 5th 2019.
[5] Grasic, V; Kos, A; Mileva-Boshkoska, B. Classification of incoming calls for the capital city of Slovenia smart city 112 public safety system using open Internet of Things data. International Journal of Distributed Sensor Networks, 14(9), 2018.
[6] Endurance Project. ENDURANCE: Strengthening the resilience of Europe's critical infrastructures (Horizon Europe, Grant Agreement No. 101168007), https://endurance-horizon.eu/, downloaded: July 7th 2025.
[7] Rinaldi, S. M., Peerenboom, J. P., & Kelly, T. K. (2001). Identifying, understanding, and analysing critical infrastructure interdependencies. IEEE Control Systems Magazine, 21(6), 11–25, https://doi.org/10.1109/37.969131, downloaded: July 7th 2025.
[8] Linkov, I., Trump, B. D., & Wagner, C. M. The frameworks of resilience planning and engineering, Environment Systems and Decisions, 34(4), 517–527, 2014.
[9] Linkov, I., & Trump, B. D. (2019). The science and practice of resilience, Springer, https://doi.org/10.1007/978-3-030-04565-4, downloaded: July 7th 2025.
[10] European Space Agency (ESA). *ESA Space Weather Awareness Portal*. SSA Space Situational Awareness Programme, https://swe.ssa.esa.int, downloaded: 30th 2025.
[11] Hapgood, M. (2017a). Space Weather. IOP Publishing, Bristol, 2017, doi: 10.1088/978-0-7503-1372-8, downloaded: September 20th 2024.
[12] Nicoll, K.A. (2014), Space weather influences on atmospheric electricity. Weather, 69: 238-241, 2014, doi: 10.1002/wea.2323, downloaded: July 20th 2024.
[13] ENTSO-E reports on grid disturbances.
[14] Cires. Introduction to geomagnetism. Cooperative Institute for Research in Environmental Sciences (Cires), University of Colorado, 2024, https://geomag.colorado.edu/magnetic-field-overview, downloaded: July 20th 2024.
[15] Olsen, N; Hulot, G; Sabaka, T. Sources of the Geomagnetic Field and the Modern Data That Enable Their Investigation. In: Freeden, W; Nashed, M.Z; Sonar, T. (eds) Handbook of Geomathematics. Springer, Berlin, Heidelberg, doi: 10.1007/978-3-642-01546-5_5, downloaded: July 20th 2024.
[16] HDGM. Modeling Earth's Geomagnetic Fields. National Oceanic and Atmospheric Administration (NOAA), May 29, 2019, Updated July 19, 2024, https://www.ncei.noaa.gov/news/HDGM, downloaded July 20th 2024.

[17] Paul, K. S., Haralambous, H., Moses, M., Oikonomou, C., Potirakis, S. M., Bergeot, N., & Chevalier, J.-M. Investigation of the ionospheric response on Mother's Day 2024 geomagnetic superstorm over the European sector, *Atmosphere, 16*(2), 180, 2025, https://doi.org/10.3390/atmos16020180, downloaded: July 7th 2025.
[18] Tacza, J., Li, G., & Raulin, J.-P. *Effects of Forbush decreases on the global electric circuit*. *Space Weather, 22*, Article e2023SW003852, 2024, https://doi.org/10.1029/2023SW003852, downloaded: July 7th 2025.
[19] Tacza, J., Raulin, J.-P., Mendonça, R. R. S., Makhmutov, V. S., Marun, A., & Fernandez, G. Solar effects on the atmospheric electric field during 2010–2015 at low latitudes, *Journal of Geophysical Research: Atmospheres, 123*(21), 11-970, 2018, https://doi.org/10.1029/2018JD029121, downloaded: July 7th 2025.
[20] Bershadskii, A. Chaotic variability of the magnetic field at Earth's surface driven by ionospheric and space plasmas, Journal of Atmospheric and Solar-Terrestrial Physics, 269, 106456, 2025, https://doi.org/10.1016/j.jastp.2025.106456, downloaded: July 7th 2025.
[21] Grasic, V, Mileva Boshoska, B, Parameters devised for modelling real-time monitoring of magnetic field for a smart city using a mobile phone, ITIS 2024, https://itis.fis.unm.si/wp-content/uploads/2024/11/ITIS-2024-Book-of-proceedings-web.pdf, pp, 39-50, downloaded: September 20th 2025.
[22] Mobile World Live. *Spanish operators power through blackouts,* https://www.mobileworldlive.com/telefonica/spanish-operators-power-through-blackouts/, downloaded: September 30th 2025.
[23] Federal Energy Regulatory Commission (FERC). *Iberian Peninsula blackout, April 2025: Final report,* https://www.ferc.gov/sites/default/files/2025-06/Iberian%20Peninsula%20Blackout%20April%202025_final.pdf, downloaded: September 30th 2025.
[24] NETSCOUT. *Iberian Peninsula blackout: Effects in cyberspace,* https://www.netscout.com/blog/asert/iberian-peninsula-blackout-effects-cyberspace, downloaded: September 30th 2025.
[25] Capacity Media. *Spain's telecom networks run on backup power as Iberian Peninsula goes dark,* https://www.capacitymedia.com/article/spains-telecom-networks-run-on-backup-power-as-iberian-peninsula-goes-dark, downloaded: September 30th 2025.
[26] JKempEnergy. *Spain's blackout blamed on poor voltage control,* https://jkempenergy.com/2025/06/19/spains-blackout-blamed-on-poor-voltage-control/, downloaded: September 30th 2025.
[27] ENTSO-E. *28 April 2025 Iberian blackout – Expert panel report,* https://www.entsoe.eu/publications/blackout/28-april-2025-iberian-blackout/, downloaded: September 30th 2025.
[28] OpenAI. (2025). ChatGPT (September 2025 version), ChatGPT-5, https://chat.openai.com, downloaded: September 20th 2025.
[29] SpainBlackout2025. SpainBlackout2025 SafeCity112, https://github.com/SafeCity112/SpainBlackout2025 , downloaded: September 30th 2025.
[30] Space Situational Awareness, services provided by PROBA2, https://proba2.sidc.be/ssa?date=2025-04-25, downloaded: September 20th 2025.

Time-resolved life cycle assessment for sustainable industry: integrating hourly analysis into smart infrastructure and energy management

Jelena Topić Božič[1,2], Andreja Dobrovoljc[1], Simon Muhič[1,2,3]
[1] Rudolfovo – Science and Technology Centre Novo mesto
Podbreznik 15, 8000 Novo mesto, Slovenia
[2] Faculty of Industrial Engineering Novo mesto
Šegova ulica 112, 8000 Novo mesto, Slovenia
[3] Institute for Renewable Energy and Efficient Exergy Use, INOVEKS d.o.o,
Cesta 2. grupe odredov 17, 1295 Ivančna Gorica, Slovenia

{jelena.topic.bozic, andreja dobrovoljc,
simon.muhic}@rudolfovo.eu
{jelena.topic-bozic, simon.muhic}@fini-unm.si

Abstract: *The role of data centers has intensified with the expansion of the digital economy and the advancement of information and communication technologies. Their environmental footprint is determined by the electricity mix, whose temporal and spatial variability is insufficiently addressed in the conventional life cycle assessment (LCA). In this study, a time-resolved environmental impact assessment was applied to electricity generation in Slovenia and Serbia in 2023. The focus was on three categories: climate change, resource use (minerals and metals), and water use. Hourly generation data from the ENTSO-E Transparency platform were linked with the Ecoinvent 3.11 datasets to generate hourly impact profiles and representative daily profiles for summer and winter.*

The study's results reveal clear differences primarily due to the distinct electricity mix structures of the two countries. Slovenia relies on nuclear, hydro, and photovoltaic power, while Serbia is predominantly coal-based. Photovoltaic generation in Slovenia reduces greenhouse gas emissions during daylight but increases the impacts related to the use of minerals and metals. Serbia exhibits higher climate change burdens yet lower variability in other categories. Seasonal and diurnal fluctuations influence emission intensities, underscoring the limits of static, annualized assessments.

The findings provide input for policy and smart infrastructure planning. Strategies for electric vehicle charging, data centers, and demand-side measures should integrate temporal profiles of environmental impacts. Tools such as environmentally differentiated tariffs or time-varying carbon pricing can help align energy use with periods of lower impact. More broadly, the results highlight trade-offs between greenhouse gas mitigation and other pressures, underscoring the need for holistic energy transition pathways.

Key Words*: data centers, life cycle assessment, electricity mix, climate change, temporal variability*

1 Introduction

The data center industry is being driven by the expansion of the digital economy and the demand for information and communication technology (ICT). The energy and

environmental costs of cryptocurrency mining and the use of data centers are an emerging issue [1]. The estimated total Bitcoin electricity use has been steadily rising. The data from the University of Cambridge Blockchain Network Sustainability Index (CBECI) shows that estimated yearly electricity use was 43.32 TWh in 2018 and has risen to 121.3 TWh in 2023, resulting in a 180.0% increase in electricity use [2]. It must be noted that mining activities also include other cryptocurrencies, which affect the overall electricity use associated with them. The study by Li et al. estimated that Monero mining may have used 645.62 GWh of electricity in 2018. If 5% of the mining activities are done in China, the electricity use would be at least 30.34 GWh, contributing to more than 19 thousand tons of carbon emission [3].

The nexus between digital technologies, machine learning workloads, and the environmental impact assessment has become a more prominent issue in recent years [4]. The global electricity demand increased by 4.3% in 2024 compared to 2023. The higher electricity demand was driven by the buildings sector, which grew four times faster than in 2023. The global electricity use in buildings increased by more than 600 TWh, with key drivers being rising demand for air conditioning and demand for power for new data centers [5]. Data centers are the backbone of ICT infrastructure; however, they consume 10 to 100 times more energy than conventional buildings while their lifespan is five times shorter [4]. It is estimated that data centers used 1.5% of electricity in 2024 [6]. Avgerinou et al. reported that the average annual electricity use of data centers participating in the EU Data Centre Code of Conduct (COC) was 13,684 MWh [7].

The growth in the number of large-scale data centers and their associated electricity demand should also be seen through the prism of rapidly changing power systems with an increasingly larger share of variable renewable energy sources (RES) [8].

Data centers' environmental impacts also differ based on factors such as location, energy sources, and efficiency measures [9]. As data centers require a significant amount of electricity, the carbon intensity of electricity will significantly impact the environmental footprint of data centers. One measure to decrease CO_2 emissions is by decarbonizing the electricity sector by adopting renewable and nuclear power technologies. Harnessing renewable energy sources like biomass, solar, wind, and geothermal energy to generate clean electricity is projected to lower global carbon emissions by approximately 70% [10]. The fluctuation in hourly emissions is influenced by the mix of technologies employed in electricity generation, as different generation technologies usually produce electricity. Since these technologies have significantly different emission factors, there is considerable potential for substantial variations in the emissions from electricity production [11]. Although using RES can result in lower environmental impacts and a reduction in greenhouse gas (GHG) emissions, other environmental impact categories, such as land use or mineral resource scarcity, may lead to higher impacts [11], [12].

To determine the environmental impact of products, services, and systems through the whole life cycle, life cycle assessment (LCA) can be used. LCA is an ISO-standardized methodology used to quantify the potential environmental impact of products, processes, or activities throughout their life cycle. It is defined by four main steps: establishing the study's goal and scope, conducting a life cycle inventory analysis, performing a life cycle impact assessment, and interpreting the results [13], [14]. LCA has emerged as a crucial tool for evaluating environmental performance across various sectors. In terms of determining the impact of electricity, most studies focus on static data inputs, which may not capture temporal variations on an hourly basis, which are associated with the specifics of electricity generation [14], [15], [16], [17].

The purpose of this study is to demonstrate the time-resolved environmental impact of the electricity mix in Slovenia and Serbia, providing context-specific data that can be applied in sustainability strategies for smart infrastructure and data centers, where aligning energy use with the environmental impact of electricity is crucial for reducing environmental impacts.

2 Materials and methods

2.1 Data collection

The data were obtained from the ENTSO-E Transparency platform, which is publicly accessible. This platform offers details on generation, load, transmission, and balancing data for European countries [18], [19] Data collection covered the year 2023 for Slovenia and Serbia. The datasets collected for creating time-resolved profiles are outlined below:

- Hourly-resolved actual generation by production type.
- Installed capacity by production type in annual resolution.

For each production type (j) and hour (t), the electricity production time series (P) was generated using Equation (1) as developed and published by [20].

2.2 Determination of hourly environmental impact of electricity produced

The hourly time series developed for electricity profiles were utilized to generate hourly environmental impact profiles for electricity production in Slovenia and Serbia by mapping them with the appropriate datasets of Ecoinvent 3.11. database for electricity generation technology [21]. For determination of environmental impact Environemntal Footprint (EF) life cycle impact assessment method was used [22]. Functional unit of 1 kWh of produced electricity was used.

Missing hourly values were interpolated by replacing the gap with the average of the immediately preceding and following data points, ensuring continuity of the time series.

For the purpose of clarity and consistent visualization, the analysis was restricted to three impact categories:

- Climate change
- Resource use, minerals and metals
- Water use.

Hourly impact values were aggregated to produce descriptive statistics: mean, standard deviation (SD), relative deviation (RSD), maximum and minimum.

3 Results and methods

3.1 Hourly Environmental Impact of Electricity Mix

Table 1 shows the average value of environmental impacts for 1 kWh of produced electricity in 2023 for Serbia and Slovenia in climate change, resource use, minerals and metals, and water use impact categories. The generation profiles of Slovenia and Serbia vary significantly from each other. Slovenia has three predominant sources of electricity: nuclear, hydro (run-of-river), and coal (lignite) power. Nuclear and hydroelectric power account for more than 70% of the total generation, followed by coal with an

approximately 20% share [23]. On the other hand, the electricity in Serbia is mainly produced from thermal power plants (63.2%), followed by hydro power plants (31.3%). The wind has 2.8% share [24].

Due to significant share of coal-based electricity generation technologies, Serbia has 103.1% higher average value in impact category climate change compared to Slovenia. Due to the lower share of intermittent RES, Serbia has lower RSD in hourly values of studied environmental impact categories.

Table 1: LCA results of the Slovenian and Serbian 2023 average hourly electricity generation mix with mean, minimum (min), and maximum (max) values, standard deviation (SD) and relative standard deviation (RSD) for selected impact categories. Dev is abbreviation for deviation.

Impact category		Unit	Mean	SD	RSD (%)	Max	Min
Climate change	Slovenia	kg CO_2 eq.	0.2812	0.1323	47.06	1.02	0.0203
	Serbia	kg CO_2 eq.	0.8790	0.1136	12.93	1.17	0.6109
Resource use, minerals and metals	Slovenia	kg Sb eq.	2.303×10^{-7}	2.063×10^{-7}	89.56	1.992×10^{-7}	5.124×10^{-7}
	Serbia	kg Sb eq.	2.1756×10^{-7}	4.793×10^{-7}	22.03	3.774×10^{-7}	1.371×10^{-7}
Water use	Slovenia	m^3 depriv.	0.06486	0.0127	19.56	0.0916	0.0078
	Serbia	m^3 depriv.	0.1136	0.0115	10.13	0.1427	0.0794

A considerable variation in the impact category of climate change throughout the average year is shown in Fig 1. The lower values in the impact category can be associated with a higher share of RES in the electricity mix, consistent with the available data for other countries, such as Germany, Poland, and Italy [20].

The scenarios in which increases in the share of PV electricity were modeled for the use cases of Slovenian and Italian yearly electricity mixes showed an impact reduction in the climate change category and an increase in the impact category of mineral resource scarcity [13], [25].

In the climate change category, higher values were observed in summer in Serbia, possibly due to the lower shares of hydro generation. A lower impact can be observed on summer days in Slovenia, with the lowest impact occurring during daylight hours owing to the production of solar PV, which is consistent with available data [11], [20]. Conversely, winter is associated with less output from solar PV, resulting in lower shares of photovoltaics in the mix and a higher impact in the case of Slovenia [11].

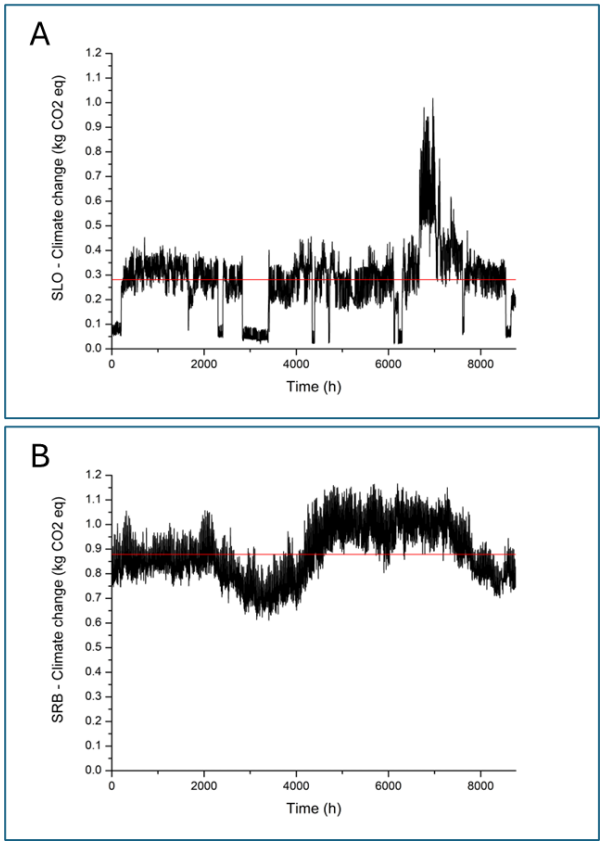

Figure 1: Time series of environmental impact category climate change for Slovenia (A) and Serbia (B) electricity production mix in 2023. The average value is indicated by a red line.

The results demonstrate that the highest daily values of GHG emissions can vary depending on the electricity mix composition, which requires the best demand-side management adoption strategies for off-peak emissions charging in the case of BEV or smart infrastructure managing. The best time period for charging BEVs from the GHG emission reduction perspective could differ between countries, although for a more accurate assessment, electricity trading should also be considered when analyzing the environmental impact of the electricity system [11].

4 Conclusions

This study highlights how the results can be applied to assess sector-specific emission factors and improve the accuracy of GHG accounting in transportation, buildings, and other end use sectors, such as data centers as they use significant amount of electricity. The findings may be used in policy development, as most existing strategies depend on static or demand-driven time-of-use tariffs and might not accurately reflect the actual environmental impact of electricity production.

For systems with a significant share of intermittent RES, i.e., Slovenia, seasonality and diurnal variations in renewable generation and resulting environmental impact should be incorporated into charging and scheduling policies. Environmentally differentiated tariffs or time-varying carbon pricing can potentially serve as effective tools to encourage consumers to adjust their behavior, promoting energy use during periods of lower environmental impact.

5 References

[1] M. Manganelli, A. Soldati, L. Martirano, and S. Ramakrishna, "Strategies for Improving the Sustainability of Data Centers via Energy Mix, Energy Conservation, and Circular Energy," Sustainability 2021, Vol. 13, Page 6114, vol. 13, no. 11, p. 6114, May 2021, doi: 10.3390/SU13116114.

[2] University of Cambridge, "Cambridge Blockchain Network Sustainability Index: CBECI." Accessed: Sep. 11, 2025. [Online]. Available: https://ccaf.io/cbnsi/cbeci

[3] J. Li, N. Li, J. Peng, H. Cui, and Z. Wu, "Energy consumption of cryptocurrency mining: A study of electricity consumption in mining cryptocurrencies," Energy, vol. 168, pp. 160–168, Feb. 2019, doi: 10.1016/J.ENERGY.2018.11.046.

[4] C. Bux, R. L. Rana, M. Lombardi, P. Giungato, and C. Tricase, "A critical analysis of global warming potential of data centers in the digital era," International Journal of Life Cycle Assessment, pp. 1–13, Jan. 2025, doi: 10.1007/S11367-024-02419-2/TABLES/2.

[5] IEA, "Electricity – Global Energy Review 2025 – Analysis - IEA." Accessed: Sep. 16, 2025. [Online]. Available: https://www.iea.org/reports/global-energy-review-2025/electricity

[6] IEA, "Understanding the energy-AI nexus – Energy and AI – Analysis - IEA." Accessed: Sep. 11, 2025. [Online]. Available: https://www.iea.org/reports/energy-and-ai/understanding-the-energy-ai-nexus

[7] M. Avgerinou, P. Bertoldi, and L. Castellazzi, "Trends in Data Centre Energy Consumption under the European Code of Conduct for Data Centre Energy Efficiency," Energies 2017, Vol. 10, Page 1470, vol. 10, no. 10, p. 1470, Sep. 2017, doi: 10.3390/EN10101470.

[8] C. Koronen, M. Åhman, and L. J. Nilsson, "Data centres in future European energy systems—energy efficiency, integration and policy," Energy Effic, vol. 13, no. 1, pp. 129–144, Jan. 2020, doi: 10.1007/S12053-019-09833-8/TABLES/1.

[9] T. Murino, R. Monaco, P. S. Nielsen, X. Liu, G. Esposito, and C. Scognamiglio, "Sustainable Energy Data Centres: A Holistic Conceptual Framework for Design and Operations," Energies 2023, Vol. 16, Page 5764, vol. 16, no. 15, p. 5764, Aug. 2023, doi: 10.3390/EN16155764.

[10] F. Wang et al., "Do 'green' data centres really have zero CO2 emissions?," Sustainable Energy Technologies and Assessments, vol. 53, p. 102769, Oct. 2022, doi: 10.1016/J.SETA.2022.102769.

[11] J. T. Božič, A. Čikić, and S. Muhič, "Environmental Impact of Slovenian and Croatian Electricity Generation Using an Hourly Production-Based Dynamic Life Cycle Assessment Approach," Energies 2025, Vol. 18, Page 4826, vol. 18, no. 18, p. 4826, Sep. 2025, doi: 10.3390/EN18184826.

[12] M. Farghali et al., "Social, environmental, and economic consequences of integrating renewable energies in the electricity sector: a review," Environ Chem Lett, vol. 21, no. 3, pp. 1381–1418, Jun. 2023, doi: 10.1007/S10311-023-01587-1/TABLES/2.

[13] J. Dimnik, J. Topić Božič, Čikić. Ante, and S. Muhič, "Impacts of High PV

Penetration on Slovenia's Electricity Grid: Energy Modeling and Life Cycle Assessment," Energies 2024, Vol. 17, Page 3170, vol. 17, no. 13, p. 3170, Jun. 2024, doi: 10.3390/EN17133170.
[14] R. Turconi, A. Boldrin, and T. Astrup, "Life cycle assessment (LCA) of electricity generation technologies: Overview, comparability and limitations," Renewable and Sustainable Energy Reviews, vol. 28, pp. 555–565, Dec. 2013, doi: 10.1016/J.RSER.2013.08.013.
[15] B. Kiss, E. Kácsor, and Z. Szalay, "Environmental assessment of future electricity mix – Linking an hourly economic model with LCA," J Clean Prod, vol. 264, p. 121536, Aug. 2020, doi: 10.1016/J.JCLEPRO.2020.121536.
[16] D. Vuarnoz and T. Jusselme, "Temporal variations in the primary energy use and greenhouse gas emissions of electricity provided by the Swiss grid," Energy, vol. 161, pp. 573–582, Oct. 2018, doi: 10.1016/J.ENERGY.2018.07.087.
[17] J. N. Louis, S. Allard, V. Debusschere, S. Mima, T. Tran-Quoc, and N. Hadjsaid, "Environmental impact indicators for the electricity mix and network development planning towards 2050 – A POLES and EUTGRID model," Energy, vol. 163, pp. 618–628, Nov. 2018, doi: 10.1016/J.ENERGY.2018.08.093.
[18] L. Hirth, J. Mühlenpfordt, and M. Bulkeley, "The ENTSO-E Transparency Platform – A review of Europe's most ambitious electricity data platform," Appl Energy, vol. 225, pp. 1054–1067, Sep. 2018, doi: 10.1016/J.APENERGY.2018.04.048.
[19] ENTSOE, "Electricity Market Transparency." Accessed: Jul. 07, 2025. [Online]. Available: https://www.entsoe.eu/data/transparency-platform/
[20] G. Naumann, J. Famiglietti, E. Schropp, M. Motta, and M. Gaderer, "Dynamic life cycle assessment of European electricity generation based on a retrospective approach," Energy Convers Manag, vol. 311, p. 118520, Jul. 2024, doi: 10.1016/J.ENCONMAN.2024.118520.
[21] Ecoinvent, "ecoinvent v3.11 - ecoinvent." Accessed: Jul. 07, 2025. [Online]. Available: https://ecoinvent.org/ecoinvent-v3-11/
[22] B. S. ANDREASI et al., "Updated characterisation and normalisation factors for the Environmental Footprint 3.1 method," Publications Office of the European Union, vol. JRC130796, 2023, doi: 10.2760/798894.
[23] IEA, "Slovenia - Countries & Regions - IEA." Accessed: Jul. 07, 2025. [Online]. Available: https://www.iea.org/countries/slovenia/electricity
[24] J. Topić Božič, S. Muhič, M. S. Komatina, M. M. Perić, and J. Dimnik, "LIFE CYCLE ASSESSMENT OF ENERGY GREEN TRANSITION GOALS IN SLOVENIA AND SERBIA Heat Pump Example," Thermal Science, vol. 28, no. 6, pp. 4709–4721, 2024, doi: 10.2298/TSCI240618222T.
[25] A. Gargiulo, M. L. Carvalho, and P. Girardi, "Life Cycle Assessment of Italian Electricity Scenarios to 2030," Energies 2020, Vol. 13, Page 3852, vol. 13, no. 15, p. 3852, Jul. 2020, doi: 10.3390/EN13153852.

Digital Twin and Machine Learning for Industrial Measurement and Diagnostics in Industry 4.0

Mare Srbinovska, Zivko Kokolanski, Vladimir Dimcev, Dimitar Taskovski, Marija Markovska Dimitrovska, Maja Celeska Krstevska
Ss Cyril and Methodius University in Skopje
Faculty of Electrical Engineering and Information Technologies,
Ruger Boskovik 18, 1000 Skopje, Macedonia
{mares, kokolanski, vladim, dtaskov, marijam, celeska}@feit.ukim.edu.mk

Abstract: *This paper presents the conceptual stage of the DTML-I4.0 project, establishing a framework that defines the methodological, architectural, and validation foundations for integrating Digital Twin (DT) and Machine Learning (ML) in industrial environments to monitor, analyse, and optimize industrial processes in real time. The framework leverages real-time sensor data to enhance process efficiency, enable early anomaly detection, and support data-driven decision-making. By combining deterministic DT models with predictive ML algorithms, the framework can accurately replicate system behaviour, forecast performance, and improve operational reliability. The proposed DT-ML framework introduces a unified architecture that connects physical modelling, IoT-based data acquisition, and intelligent analytics for industrial measurement and diagnostics. Although the work is currently conceptual, it outlines a clear validation roadmap for pilot implementation, defining metrics for evaluating predictive accuracy and maintenance efficiency.*

Key Words*: Digital twin; Machine learning; Industry 4.0; Forecast performance.*

1 Introduction

The rapid evolution of Industry 4.0 has transformed traditional manufacturing and process industries by introducing advanced digital technologies. Among the most influential tools in this transformation are Digital Twins (DT), which serve as virtual replicas of physical systems, enabling real-time monitoring, simulation, and optimization. At the same time, Machine Learning (ML) has emerged as a powerful method for extracting patterns, predicting failures, and enhancing decision-making based on large volumes of industrial data. The integration of DT and ML offers a unique opportunity to improve measurement and diagnostic methods, bridging the gap between physical processes and intelligent analytics.

Industrial processes rely heavily on accurate measurements and diagnostics to ensure reliability, safety, and efficiency. Conventional approaches often struggle to deal with complex, dynamic environments where data variability and uncertainty are high. By embedding sensors and IoT devices into machines, continuous streams of operational data can be captured and fed into the digital twin for real-time analysis. ML models can then process this information to identify anomalies, predict potential breakdowns, and optimize performance. This approach not only reduces downtime and maintenance costs but also enhances the sustainability and resilience of industrial systems.

In recent years, the integration of DTs with unsupervised learning approaches, such as autoencoders combined with Long Short-Term Memory (LSTM) networks, has shown remarkable promise in predictive maintenance scenarios, achieving anomaly detection accuracies up to 98% in industrial machinery applications [1]. Another practical implementation is found in food processing plants, where digital twin models coupled with ML algorithms (regression, neural networks, clustering) successfully identify system anomalies in real time [2]. A conveyor belt case study further demonstrates the efficacy of a DT framework using ML for predictive maintenance in manufacturing environments [3]. Similarly, AI-driven DT systems in smart manufacturing have improved predictive accuracy by 35%, reduced unplanned downtimes by 40%, and optimized maintenance costs by 25% [4].

The architecture of predictive maintenance solutions, leveraging IoT sensors, cloud computing, and AI-enabled digital twins, has been documented to boost operational reliability by 30% and reduce outages by 25% [5]. Recent pilot implementations in production environments highlight practical challenges and lessons learned when deploying DTs for maintenance, emphasizing the need for ongoing calibration and resilience in real-world settings [6]. Beyond applied use cases, conceptual studies investigate the co-evolution of AI, IoT, and DT technologies, mapping how these interdependencies enhance system resilience and human-machine interaction in Industry 4.0 [7]. State-of-the-art reviews also underscore the pivotal role of machine learning in enabling multi-scale, cross-domain digital twin functionalities such as modeling, monitoring, optimization, and visualization [8].

Despite the growing adoption, challenges remain in areas such as cybersecurity, privacy, and the integration of trustworthy AI within DT systems [9]. These limitations indicate that while DT and ML are already reshaping industrial practices, their combined application in measurement and diagnostic methods is still in its early stages and requires further exploration.

This paper presents the conceptual phase of a project, in which a unified Digital Twin–Machine Learning (DT-ML) architecture is designed for industrial measurement and diagnostics. The framework integrates deterministic physical modeling with data-driven learning and introduces an adaptive calibration loop between the physical and virtual layers. Such implementations are still rare in the region, and this initiative establishes the first integrated DT-ML testbed for industrial diagnostics at a national university in North Macedonia. The scientific contribution of this work lies in defining the architecture, data flow, and validation methodology that connect IoT-enabled sensing with predictive analytics for process optimization.

The objective of this paper is to present a conceptual and practical framework integrating Digital Twin and Machine Learning techniques for industrial processes, aimed at enhancing diagnostic accuracy, enabling predictive maintenance, and supporting the broader objectives of Industry 4.0.

2 Methodology

The methodology of this work is closely aligned with the planned activities of the DTML-I4.0 project and is structured in several phases:

i) Development of DT models via mathematical modeling and simulation,

ii) Integration of IoT sensors and data acquisition systems,

iii) Implementation of ML algorithms (i.e. Random Forest, LSTM) for predictive diagnostics with target accuracy >85%,

iv) Validation in a real industrial environment, and

v) Deployment in pilot applications with SCADA/IoT platforms.

Since the DTML-I4.0 project is in its conceptual stage, this section defines the system architecture, data flow, and measurable objectives, such as prediction accuracy, latency reduction, and diagnostic reliability that will guide and evaluate subsequent implementation. The interaction between the DT and ML layers is modeled as a continuous feedback loop, where live sensor data update the digital-twin state, and the trained ML algorithms generate predictive diagnostics that are fed back to support early warnings, fault detection, and maintenance decisions in the industrial environment.

2.1 System architecture

The proposed framework integrates DT and ML into a unified architecture tailored for industrial measurement and diagnostic applications. The system consists of four main layers: physical process layer, sensing and IoT layer, DT layer, and ML anf analytics layer as it is shown in Figure 1.

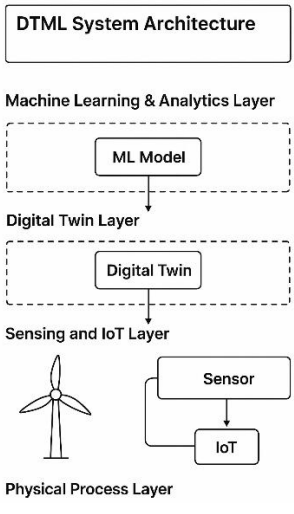

Figure 1 DTML System Architecture

At the physical process layer, industrial equipment such as compressors, HVAC (Heating, Ventilation and Air Conditioning) units, or production lines generate operational data during real-time operation. The sensing and IoT layer captures this information using

vibration, temperature, pressure, and air quality sensors, which are connected via IoT gateways to ensure continuous monitoring and secure data transmission.

The DT layer provides a continuously updated virtual replica of the physical system, enabling real-time visualization, simulation of operational scenarios, and comparison of measured and simulated behavior for deeper analysis.

Finally, the ML and analytics layer applies advanced ML algorithms, including supervised methods (regression, random forest, neural networks) and unsupervised methods (clustering, anomaly detection), to detect abnormal patterns, predict failures, and optimize process performance. Cloud-based and edge-computing solutions are considered to ensure scalability and real-time processing capability.

The communication between layers follows a modular architecture based on standard data protocols to ensure interoperability. Feature-extraction and preprocessing modules are designed to manage sensor-data heterogeneity and noise. During future implementation, model performance will be evaluated through cross-validation and benchmarking against traditional threshold-based diagnostics.

2.2 Measurement and diagnostic methods

Accurate measurement and diagnostics are essential to validate and continuously improve the proposed DT–ML framework. Multiple measurement streams are acquired through IoT-enabled sensors, including:

- Vibration signals for detecting mechanical imbalance, wear, or misalignment.
- Temperature measurements for identifying overheating and thermal stress.
- Pressure and flow data for assessing process efficiency and stability.
- Particulate concentration (PM2.5, PM10) for ensuring environmental and safety compliance.

Collected data undergo preprocessing (noise reduction, normalization, and feature extraction) before being fed into the ML model. Diagnostic methods rely on hybrid approaches where deterministic models (derived from system physics) are combined with data-driven ML models for enhanced accuracy. Predictive maintenance is achieved by ML algorithms trained on both historical and live sensor data, enabling early-fault detection and minimizing unplanned downtime.

2.3 Planned Validation and Evaluation Roadmap

Future validation of the DT-ML framework will be carried out in a real industrial environment where process data are continuously measured and transmitted through an integrated metering platform to a secure web interface. The DT environment will mirror key operational parameters for comparison with real measurements, while the ML module will analyze streaming data in real time to detect anomalies and forecast performance.

Model evaluation will rely on cross-validated experiments using recorded time-series data, with performance quantified by Precision–Recall AUC (PR-AUC), F1-score, and Mean Absolute Percentage Error (MAPE). Expected results include improved early-fault detection and reduced diagnostic latency relative to conventional threshold-based

approaches. Pilot validation in the industrial environment will confirm the framework's accuracy, robustness, and adaptability, providing the basis for quantitative assessment in later project phases.

3 Discussion

The proposed DTML framework (Figure 2) represents the conceptual foundation of an ongoing initiative to integrate Digital Twin and Machine Learning technologies for industrial measurement and diagnostics. It defines a unified architecture and data flow logic designed for real-time, data-driven decision support in industrial environments.

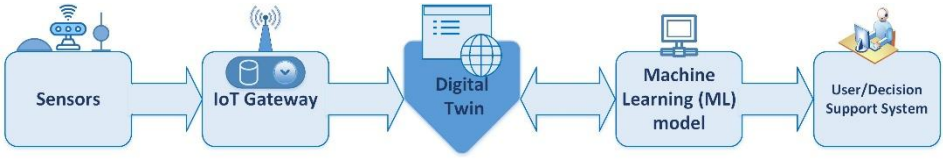

Figure 2 Block diagram of DTML proposed model

Through continuous acquisition of sensor data, the digital-twin environment replicates process dynamics and provides a synchronized virtual representation for simulation and scenario testing without interrupting real operations. This coupling between live data and virtual modeling enables the analysis of process deviations as they occur, supporting anomaly detection and predictive maintenance decisions. Compared to conventional threshold-based monitoring, the DT-ML approach is expected to deliver more accurate diagnostics and earlier detection of performance degradation.

A key contribution of the framework is its methodological integration of model-agnostic ML algorithms within the twin loop, supported by standardized data exchange and evaluation metrics (PR-AUC, F1, MAPE). This structure enables transparent and reproducible validation, facilitating comparison across different industrial contexts. In practice, the framework provides operators with visualization and decision-support capabilities that transform raw data into actionable insights, thereby improving response time and maintenance planning.

The design also considers critical Industry 4.0 challenges, including data interoperability, computational scalability, and reliability under real-time constraints. Although initial implementation requires investment in sensing and infrastructure, the long-term benefits in reliability, cost reduction, and energy efficiency are expected to be substantial. Moreover, the framework is inherently modular and can incorporate explainable AI, edge analytics, or federated learning in future phases, ensuring adaptability to evolving industrial ecosystems.

Overall, the DT–ML system advances the integration of DTs and ML by establishing an operational feedback framework where machine learning continuously informs the DT and supports data-driven diagnostics. The forthcoming pilot validation will provide empirical evidence on accuracy, robustness, and scalability, confirming its applicability to real industrial processes.

4 Conclusion

This paper presented the conceptual development of a unified Digital Twin and Machine Learning (DT-ML) framework for real-time monitoring, diagnostics, and optimization of industrial processes. By linking virtual process models with live sensor data, the framework enables synchronized simulation, predictive analytics, and data-driven decision support aimed at improving reliability and operational efficiency.

The proposed methodology defines the architectural, analytical, and validation foundations of the DTML-I4.0 project, establishing a clear pathway from conceptual design to industrial implementation**.** ML algorithms trained on historical and live data will support predictive maintenance and provide recommendations for process adjustment, while the DT environment will provide real-time visualization and comparison between measured and simulated performance.

The study emphasizes the importance of accurate data acquisition, sensor integration, and signal processing as prerequisites for trustworthy predictions and diagnostics. Designed as modular and platform-agnostic, the framework can be applied across diverse industrial domains, ensuring scalability and interoperability.

Future work will focus on pilot validation in a real industrial environment to assess accuracy, robustness, and responsiveness under operational conditions. These results will inform subsequent integration of advanced predictive algorithms, real-time feedback, and explainable AI mechanisms.

Overall, the DT-ML framework establishes a flexible and transparent foundation for bridging physical and digital industrial systems. Its combination of DT modeling and ML analytics provides a pathway toward predictive, adaptive, and sustainable industrial operations.

5 Acknowledgment

The authors gratefully acknowledge the support of the Ministry of Education and Science of the Republic of North Macedonia for funding the project *"Integration of Digital Twin and Machine Learning for Measurement and Diagnostic Methods in Industrial Processes within Industry 4.0."* The Ministry's commitment to fostering scientific excellence and innovation in higher education has been instrumental in enabling the successful development of this research.

6 References

[1] Kerkani, R., Mhalla, A., Bouzrara, K. Unsupervised Learning and Digital Twin Applied to Predictive Maintenance for Industry 4.0. Journal of Electrical and Computer Engineering, 2025. https://doi.org/10.1155/jece/3295799

[2] Tancredi, G. P., Vignali, G., Bottani, E. Integration of Digital Twin, Machine-Learning and Industry 4.0 Tools for Anomaly Detection: An Application to a Food Plant. Sensors, 22(11), 4143, 2022. https://doi.org/10.3390/s22114143

[3] Pulcini, V., Modoni, G. Machine learning-based digital twin of a conveyor belt for predictive maintenance. The International Journal of Advanced Manufacturing

Technology, 133, 6095–6110, 2024. https://doi.org/10.1007/s00170-024-14097-3

[4] Prabu, S., Senthilraja, R., Mudassar Ali, A., Jayapoorani, S., Arun, M. AI-Driven Predictive Maintenance for Smart Manufacturing Systems Using Digital Twin Technology. International Journal of Computational and Experimental Science and Engineering, 11(1), 2025. https://doi.org/10.22399/ijcesen.1099

[5] Malik, J. A., Akhtar, M., Naz, N. S., Rani, A. Digital Twin Technology for Predictive Maintenance in Industrial Systems. Southern Journal of Engineering & Technology, 3(2), 2025.

[6] Hassan, M., Svadling, M., Björsell, N. Experience from implementing digital twins for maintenance in industrial processes. Journal of Intelligent Manufacturing, 35, 875–884, 2024. https://doi.org/10.1007/s10845-023-02078-4

[7] Omiyale, B. O., Odeyemi, J., Ogbeyemi, A., Olorunsogbon, F., Zhang, W. C. Impact of cyber physical systems on enhancing robotic system autonomy: a brief critical review. The International Journal of Advanced Manufacturing Technology, 138, 3925-3942, 2025.https://doi.org/10.1007/s00170-025-15828-w

[8] Nele, L., Mattera, G., Yap, E.W., Vozza, M., Vespoli, S. Towards the application of machine learning in digital twin technology: a multi-scale review. Discover Applied Science, 6, 502, 2024. https://doi.org/10.1007/s42452-024-06206-4

[9] Zemskov, A. D., Fu, Y., Li, R., Wang, X., Karkaria, V., Tsai, Y.-K., Chen, W., Zhang, J., Gao, R., Cao, J., Loparo, K. A., Li, P. Security and Privacy of Digital Twins for Advanced Manufacturing: A Survey. arXiv preprint, arXiv:2412.13939, 2024. pp

Establishing a Data-Driven Feedback Loop for the Optimization of Production Processes

Andrej Dobrovoljc
Faculty of Information Studies
University of Novo mesto
Ljubljanska cesta 31A, 8000 Novo mesto, Slovenia
andrej.dobrovoljc@fis.unm.si

Abstract: *This paper presents the design and implementation of a digital feedback loop for optimizing material consumption in a manufacturing environment. The study focuses on small and medium-sized enterprises (SMEs) that often lack access to costly Manufacturing Execution Systems (MES). We demonstrate how commonly available tools such as Microsoft Excel, Power Query, and open-source solutions can be combined. We created a functional feedback mechanism linking ERP data, CNC machine outputs, and production logs. The proposed solution was developed and tested in a woodworking company producing custom furniture components. By integrating heterogeneous data sources, we established a real-time overview of material usage and waste, reducing manual work and increasing process transparency. The study highlights the role of simple Extract Transform Load (ETL) tools in supporting smart manufacturing, data-driven decision-making, and continuous process improvement.*

Key Words: *Digital feedback loop, Smart manufacturing, Power Query, ERP integration, Data transformation, Material optimization, ETL process*

1 Introduction

The competitiveness of manufacturing enterprises increasingly depends on their ability to digitalize and automate core business processes. However, many organizations face challenges in implementing full-scale digital transformation. Incomplete digitalization results in information silos, manual interventions, and inefficient workflows. Employees often lack the necessary digital literacy to exploit the potential of available technologies, and integration between systems such as Enterprise Resource Planning (ERP). Therefore production control is frequently inadequate. For small and medium-sized enterprises (SMEs), these challenges are further amplified by limited financial resources and insufficient IT support. While large companies can afford dedicated MES software, SMEs must seek alternative, low-cost solutions for data integration and process monitoring.

The aim of this paper is to explore how a digital feedback loop can be established using affordable tools like Microsoft Excel, Power Query and open sources solutions to optimize material usage in manufacturing processes. The approach demonstrates how even basic tools can enable smart manufacturing principles, linking real-time data capture, analysis, and feedback.

2 Related Work

To maintain competitiveness in the digital economy, effective change management and the optimization of operational performance are crucial. These factors demand high organizational agility. Traditional, periodic feedback methods are too slow to address real-time challenges. Organizations are therefore shifting to Continuous Feedback Loops (CFL). These act as a constant mechanism for capturing, analyzing, and responding to events in the production process in real time. This dynamic approach enables leaders to make data-driven adjustments and proactively address emerging issues. This consequently increases agility and reduces resistance to change [1].

Key to operational management and performance improvement are Business Intelligence (BI) systems, which transform raw data into knowledge for decision-making. The Microsoft Power BI platform allows for the creation of real-time dashboards, facilitating the tracking of Key Performance Indicators (KPIs) and enabling timely, data-driven decisions. Such affordable solutions are particularly important for Small and Medium-sized Enterprises (SMEs) [5]. The Action Design Research (ADR) methodology is often used in research to develop and evaluate these solutions in a real-world context [5, 7].

Establishing a digital feedback loop requires a reliable ETL (Extraction, Transformation, Load) process to capture and transform data from various sources. In the modern environment using predictive analytics, traditionally static ETL pipelines are being transformed into dynamic feedback infrastructures. These operate as two-way channels: they supply data for training (Forward Flow) and capture feedback from model outputs or user responses (Backward Flow). This is essential for continuous model learning and adaptation, enabling automated monitoring, retraining, and redeployment [2].

The Power Query tool (integrated into Power BI and Excel) is recognized as suitable for performing the ETL process even for less skilled users. It allows for data preparation, processing, and cleaning, and addresses issues related to insufficient knowledge and poor digital literacy of employees [6].

In real-time systems, data quality is the most critical success factor, because incomplete or inconsistent data leads to inaccurate predictions and operational inefficiency. Data quality assurance must be an integrated and continuous process throughout the entire ETL lifecycle [3].

Beyond the technical challenges of digitalization, there is also the risk of degenerative feedback loops in human-AI interactions. Interacting with biased AI systems alters human judgment processes and consequently amplifies biases in people, triggering a snowball effect [4]. This highlights the necessity of moving towards objective, data-driven systems.

3 Research context

The research was conducted in a medium-sized woodworking company specializing in the production of custom furniture for automotive interiors. Each product consists of numerous wooden components manufactured from large wooden panels using CNC machines. The production process starts with raw boards that are cut based on Computer-Aided Design (CAD) drawings. The CAD software generates two files: a

visual layout in PDF file format and a binary file for machine control. The PDF file contains the cutting layout, element identifiers, and quantities, while the binary file serves as input for CNC machines.

The estimated standard for material consumption of the selected production item was recorded in the bill of materials (BOM) within the ERP system. One of the main challenges was the inconsistency between CAD and ERP identifiers, as the same drawing could correspond to multiple ERP item codes. To enable integration, a mapping function between CAD and ERP codes had to be developed. Data from the production process was not stored in one system, and integration was missing. This led to inefficient manual data reconciliation. Consequently, real-time monitoring of material consumption was hindered.

4 Methodology

We developed the digital feedback loop following the Action Design Research (ADR) framework, combining practical intervention with systematic evaluation. The process consisted of four main stages: (1) identification of data sources, (2) integration and transformation of data, (3) creation of a consolidated data model, and (4) validation and visualization of results.

At first, all relevant data sources were identified, including ERP exports, PDF and CSV files generated by CNC machines, and Excel records maintained by warehouse staff. Each source used different file formats, encodings, and naming conventions. In the second step, by using Power Query in Excel, we implemented an Extract-Transform-Load (ETL) process to automatically gather and clean these data sources. Power Query was chosen due to its accessibility, its ability to import multiple files simultaneously, and its strong transformation capabilities. The ETL pipeline standardized column names, removed inconsistencies, and merged records based on shared identifiers. In third step, we established a mapping table to link CAD item identifiers with ERP material codes, ensuring that all material usage data could be aggregated per product and per production order. Finally, validation was performed through manual comparison of calculated and recorded material consumption, and results were visualized through Excel dashboards and summary tables.

5 Results

The implemented Power Query solution successfully connected data from multiple heterogeneous sources into a unified view. By combining CNC-generated files, ERP issue records, and Excel-based warehouse logs, the system provided an integrated overview of material flow through production.

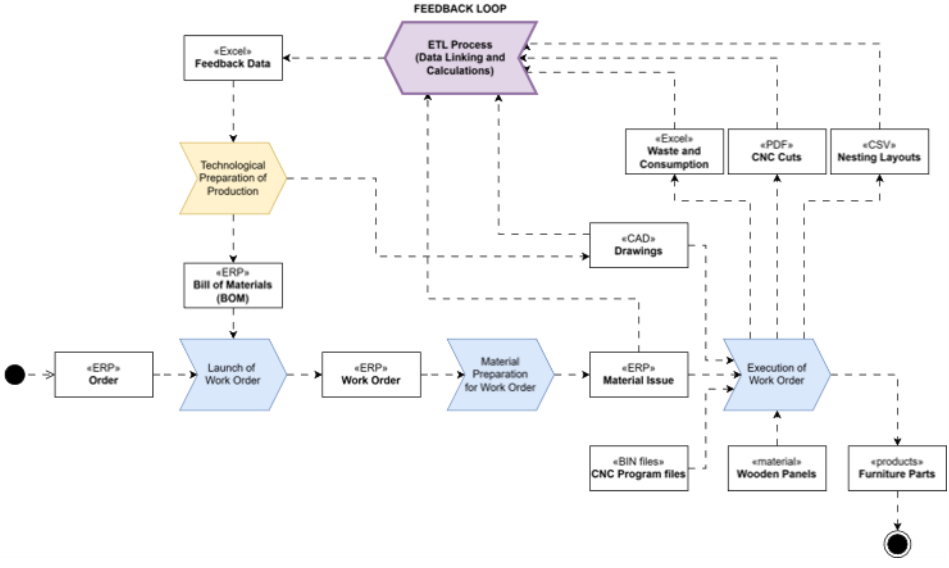

Figure 1 – Production Process Overview and Feedback Loop

Figure 1 illustrates the manufacturing process flow from CAD design to production feedback. Each stage generates digital data that feed into the feedback loop.

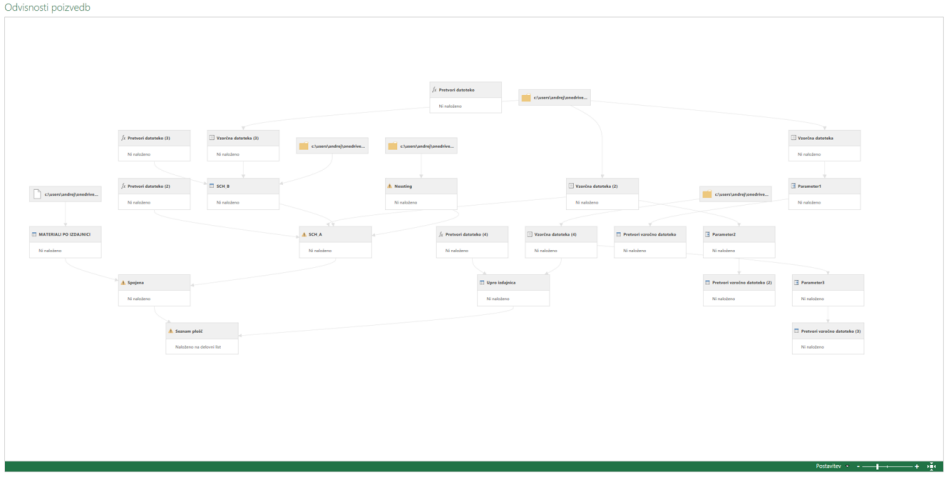

Figure 2 – Data Flow Diagram

Figure 2 presents the data flow architecture, highlighting sources, their connection and dependencies and outputs.

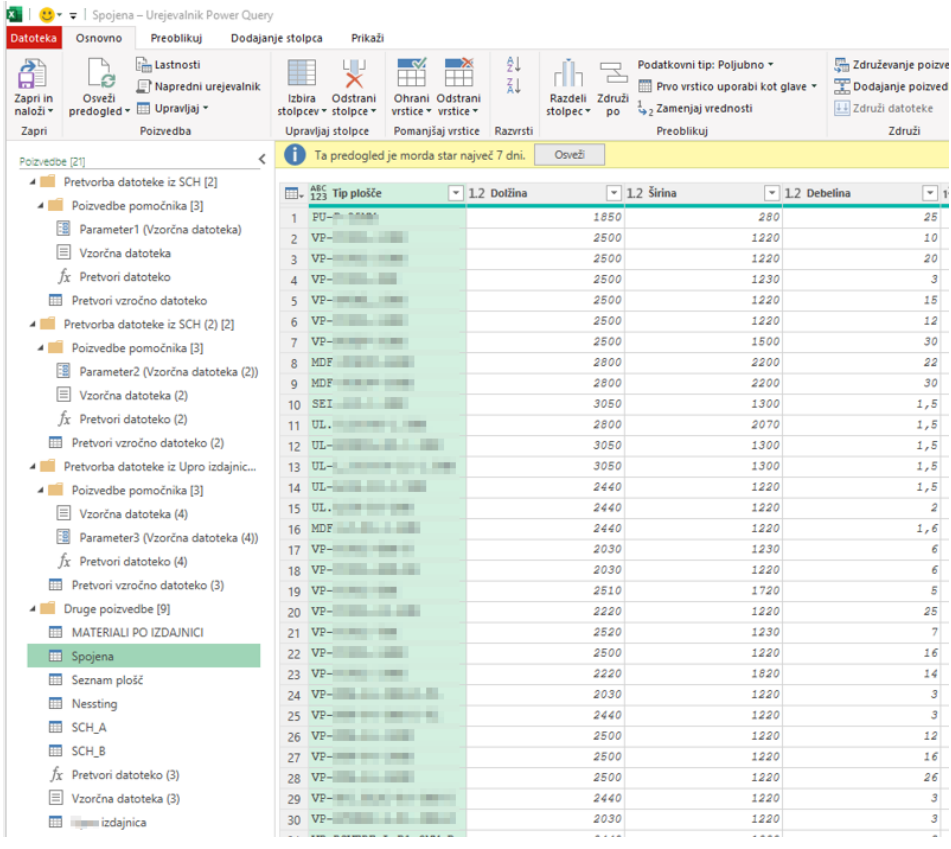

Figure 3 – ETL process supported by Power Query

Figure 3 presents the sources and steps of ETL process conducted in Power Query module within Microsoft Excel application. The final results were exported as Excel table.

The resulting system enabled real-time tracking of material usage per product, identification of waste, and deviation analysis between normative and actual consumption. Employees were able to verify discrepancies quickly, reducing manual reporting and increasing awareness of material efficiency. The project confirmed that even cost-effective software like Power Query with ETL process support (available in Microsoft Excel), when properly configured, can replicate the essential functionalities of MES. The costs of realizing such a system are also significantly lower. The approach thus bridges the gap between traditional manufacturing and smart factory principles.

6 Discussion

The results demonstrate that SMEs can achieve substantial benefits from digital feedback loops without investing in expensive IT infrastructure. The proposed solution fosters collaboration between production, warehouse, and development departments by providing shared visibility of operational data. However, challenges remain in ensuring data quality and employee engagement. Training proved essential, as limited digital literacy among staff initially hindered adoption. The study also highlights that

organizational readiness is as critical as technological capability in successful digital transformation. From a technical perspective, Power Query and Python libraries such as *pandas* can complement each other (Power Query for accessibility and Python for advanced analytics). The next logical step involves extending this approach with automated updates of normative in ERP application, predictive models for waste reduction, and integration with Power BI dashboards.

7 Conclusion and Further Work

This research confirmed that a digital feedback loop can be implemented using standard office tools like Microsoft Excel and Power Query, making smart manufacturing principles achievable even in resource-constrained environments. The integration of heterogeneous data sources improved transparency, material efficiency, and data-driven decision making. The case study validated the feasibility of using low-cost digital tools to replicate key MES functionalities.

Future research will focus on expanding automation, introducing real-time data synchronization between ERP and production equipment, and developing predictive analytics models to further optimize material usage and reduce production waste.

8 References

[1] Carreno, A. M. (2024). Building a Continuous Feedback Loop for Real-Time Change Adaptation: Best Practices and Tools. White Paper. DOI: 10.5281/zenodo.14051466.

[2] Edward, E. (2025). Data Quality Assurance in Real-Time Predictive ETL. ResearchGate.

[3] Edward, E. (2025). Model Retraining and Continuous Learning Through ETL-Oriented Feedback Loops. ResearchGate.

[4] Glickman, M., & Sharot, T. (2024). How human–AI feedback loops alter human perceptual, emotional and social judgements. Nature Human Behaviour. https://doi.org/10.1038/s41562-024-02077-2

[5] Nabil, D. H., Rahman, M. H., Chowdhury, A. H., & Menezes, B. C. (2023). Managing supply chain performance using a real time Microsoft Power BI dashboard by action design research (ADR) method. Cogent Engineering, 10(2), 2257924. DOI: 10.1080/23311916.2023.2257924

[6] Nazarov, D. M., Nazarov, A. D., & Kondratenko, I. S. (2020). Digital Education Competencies: Power Query. Ural State University of Economics.

[7] Sein, M. K., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action design research. MIS Quarterly, 35(1), 37–56. https://doi.org/10.2307/23043488

Comparative Analysis of Machine Learning Models for Telecommunications Churn Prediction

Maja Cerjan[1], Leo Mršić[2], Kornelije Rabuzin[1], Biljana Mileva Boshkoska[3]
[1]University of Zagreb Faculty of Organization and Informatics
Pavlinska ulica 2, 42000 Varaždin, Croatia
[2]Algebra Bernays University
Gradišćanska 24, 10000 Zagreb, Croatia
[3]Faculty of Information Studies
Ljubljanska cesta 31A, 8000 Novo mesto, Slovenia
biljana.mileva@fis.unm.si
macerjan@foi.unizg.hr, krabuzin@foi.unizg.hr
Leo.Mrsic@algebra.hr

Abstract:

Customer retention is a major problem in the telecommunications industry. This study develops and evaluates models to identify possible churners. Machine learning techniques ("Decision Trees", "Random Forests", "Logistic Regression" and "Neural Networks (multilayer perceptron MLP)") were applied through Python and R to analyze the "Telco Customer Churn" Kaggle dataset, based on customer assests and service usage. The data pre-processing compiled missing data and then standardized it. Evaluation used nested 10-fold cross-validation with an inner loop for hyperparameter tuning and mutual-information top-K feature pruning, with pre-processing confined to training folds. In Python, RF and LR achieve F1(~0.629), with Logistic Regression accuracy ~0.75. In R, Logistic Regression performed best (F1 ≈ 0.60 ± 0.03, Accuracy ≈ 0.80 ± 0.01). Metrics derived from pooled confusion matrices averaged over folds equal outer-fold means, confirming generalization across folds and between Python and R. Research offers empirical evidence for transferring and testing churn prediction models across Python and R in telecommunications analytics, with fully reproducible evaluation and results.

Keywords*: Customer churn, telecommunications, churn prediction, machine learning, random Forest, decision tree, logistic regression, neural networks (MLP), Python, R, Nested cross-validation*

1 Introduction

For the providers of telecommunications, high customer churn means severe financial losses since their customers leave for a variety of reasons that can range from unhappiness to receiving better offers elsewhere or due to personal changes as also observed in works on dynamic churn behavior [1]. The knowledge of predicting customer churn is important in the process of retaining the customers [11]. Loss of customers means lost revenue for businesses, because keeping existing clients is cheaper than acquiring new ones. Being able to predict churn, allows companies to reduce expenses and increase revenue. Organizations that can predict churn, enabling them to offer personalized rewards and better service offerings to improve customer satisfaction and retention, can lead the market and continue revenue generation for innovation. This paper compares the

performance of four supervised techniques: Decision Tree, Random Forest, Logistic Regression and Neural Network (MLP) for predicting the telecom churn on public "Telco Customer Churn" dataset [4]. A nested 10-fold model is used for unbiased estimation and fair model selection. Accuracy, Precision, Recall and F1 score are reported as the outer-fold mean ± standard deviation performance using both Python (*scikit-learn*) and R (*caret with rpart/randomForest/nnet*) implementations.

In contrast to almost all the studies in churn prediction, our analysis provides a cross-language comparison (Python and R) of two of the most popular analytic languages in industry that provide empirical evidence supporting reproducibility of estimation. In addition, the study considers for a nested cross-validation (which is disregarded largely in churn literature) to exclusively test the predictive model. These contributions not only increase much-needed academic rigor (in the form of reproducible and methodologically sound work), but also make participants' efforts more valuable.

2 Literature Review

Customer churn analysis is an important domain in telecommunications that affects revenue and customer loyalty. Several studies focus on churn models and the application of machine learning and deep learning for predicting churn reduction in the telecommunications sector, including data pre-processing and evaluation metrics, and surveys that provide an overview of various methodologies [8]. This paper discusses the current works, methodologies and results on churn prediction, reviewing advances and most common approaches.

2.1 Overview of Existing Research on Churn Prediction

Various literature sources investigate the use of machine learning for churn prediction. For instance, in [23] a data balancing was introduced, [5] emphasized the efficiency of techniques and in [2] learning techniques were suggested to improve capacities. Churn prediction has evolved from statistics to machine learning. Among the datasets used in studies are those provided by telecom providers over platforms such as Kaggle for example, "Telco Customer Churn" dataset.

2.2 Commonly Used Techniques and Models in Churn Prediction

The various machine learning algorithms used in churn prediction techniques, such as SVM, k-NN, and other neural networks, have been shown to have varying benefits and drawbacks in real-world churn prediction systems [14].

Ensemble techniques, like Random Forest and Gradient Boosting, are well-known tools to improve the precision of models [5], [3], [26]. Highly imbalanced classes was one of the challenges in churn prediction, for which SMOTE, and resampling have been used to employ [23], [19], [20], [28], [9]. Feature selection is also an essential part to improve the accuracy [2], [22], [25], [24]. Recent papers [17], [12], [7] have proved that more powerful algorithms, especially the algorithms in LSTM and CNN, can be adopted to describe complex temporal patterns and customer behavior. Model performance would be generally in the form of numbers like accuracy, precision, recall, F1-score and AUC score. Furthermore, graph-based social network analysis has been applied in churn prediction to understand more about customer relationships [13]. There is also a new tendency towards integrated frameworks of churn prediction together with customer

segmentation [28], [21], [30], while clustering methods are often combined with other models for improved performance [29], [15]. In addition, the cloud-based ETL frameworks are also useful for efficiently processing large-scale data extraction, transformation, and loading [31].
This paper presents a cross-language benchmark on the Telco dataset. The models implemented are Decision Tree, Random Forest, Logistic Regression, and a neural network (MLP) in both Python and R. The ranks of the algorithms and their absolute performance is reported along with Fold-aggregated confusion matrices for reproducibility.

3 Dataset Description

This section describes the dataset used in the study, and its source.

3.1 Source of the Data

The study uses the publicly available "Telco Customer Churn" dataset from Kaggle [4]. The dataset contains 21 columns in total, from which 19 are predictors, 1 target (*churn*) and 1 identifier (*customerID*) describing customer demographics, service usage, and account details from a telecommunications provider. The target is moderately imbalanced. In the cleaned data (n=7032), Churn = Yes: 1869 (26.6%) and No: 5,163 (73.4%). Table 1 lists all attributed with brief descriptions and their measurement scales.

Table 1. Summary table of the variables

Variable	*Description*	*Type / Values*
customerID	Unique customer identifier	String (excluded from modeling)
gender	Customer gender	Categorical: *Female, Male*
SeniorCitizen	Senior status	Binary: *0 = No, 1 = Yes*
Partner	Whether customer has a partner	Categorical: *Yes, No*
Dependents	Whether customer has dependents	Categorical: *Yes, No*
tenure	Number of months customer has stayed with company	Integer: *1 – 72*
PhoneService	Whether customer has a phone service	Categorical: *Yes, No*
MultipleLines	Multiple line service	Categorical: *Yes, No, No phone service*
InternetService	Internet service provider	Categorical: *DSL, Fiber optic, No*
OnlineSecurity	Online security add-on	Categorical: *Yes, No, No internet service*
OnlineBackup	Online backup add-on	Categorical: *Yes, No, No internet service*
DeviceProtection	Device protection plan	Categorical: *Yes, No, No internet service*
TechSupport	Technical support add-on	Categorical: *Yes, No, No internet service*
StreamingTV	Streaming TV service	Categorical: *Yes, No, No internet service*
StreamingMovies	Streaming movies service	Categorical: *Yes, No, No internet service*
Contract	Contract type	Categorical: *Month-to-month, One year, Two year*
PaperlessBilling	Whether customer uses paperless billing	Categorical: *Yes, No*

PaymentMethod	Payment method	Categorical: *Electronic check, Mailed check, Bank transfer (automatic), Credit card (automatic)*
MonthlyCharges	Monthly charges billed during tenure	Continuous (float): *18 – 118 USD*
TotalCharges	Total charges billed during tenure	Continuous (float): *19 – 8685 USD*
Churn	Whether customer discontinued service during observation	Categorical: *Yes, No*

These attributes capture key aspects of customer demographics, account setup, and service usage that are commonly associated with churn behavior.

4 Methodology

In this section, approaches and steps taken in building and evaluating churn prediction models were described. It shows the preparation of data and model, their training and evaluation. Three methods were used to reduce the effects of class imbalance: stratified folds, precision-recall-F1 evaluation reports, and class weights during tuning for Logistic Regression and Random Forest.

4.1 Decision Trees

Decision Trees, a non-parametric method for classification and regression, iteratively divide data based on input feature values [16]. Splits are determined by metrics like Gini impurity:

$$Gini = 1 - \sum_{i=1}^{n} p_i^2$$

where pi is the probability of an element being classified for a particular class.

Decision trees provide advantages through their easy-to-understand structure and their capability to process both numerical and categorical variables without needing normalization or feature scaling. The interpretability of Decision Trees makes them appropriate for applications that require transparency. The high degree of variance in Decision Trees makes them sensitive to small input data perturbations, which results in unstable tree structures. The model tends to overfit complex datasets when not limited by regularization techniques such as pruning or maximum depth restriction.

4.2 Random Forest

Random Forest, an ensemble method, combines multiple decision trees to improve accuracy [26]. Predictions are made by majority voting:

$$y = \text{mode}(y_1, y_2, \ldots, y_n)$$

where *yi* is the prediction from the tree, and the term *mode* refers to the most frequently occurring prediction among the individual decision trees.

The Random Forest models reduce overfitting in decision trees by training multiple trees on different subsets of data and feature spaces. The ensemble method improves both generalization performance and model robustness especially when dealing with high-dimensional datasets [27]. Random Forests demonstrate the ability to detect complex non-linear patterns, and they offer built-in feature importance scoring. The improved predictive capabilities of Random Forests result in decreased interpretability because the decision-making process becomes obscure through the combination of multiple trees. The implementation of Random Forests demands additional computational resources which include memory and processing time requirements especially when working with large datasets or numerous trees.

4.3 Logistic Regression

Logistic Regression, widely used for classification, has been shown effective in telecom churn prediction, including both traditional logistic models [10] and broader reviews of churn prediction techniques. The logistic function is:

$$P(y = 1|X) = \frac{1}{1 + \mathrm{e} - (\beta_0 + \beta_1 \mathrm{X}_1 + \cdots + \beta_n \mathrm{X}_n)}$$

where $\beta_0, \beta_1, \ldots, \beta_n$ are the coefficients of the model and $\mathrm{X}_1, \ldots, \mathrm{X}_n$ are the feature values.

Logistic Regression is commonly used in binary classification problems due to its simplicity, interpretability, and well-established statistical theory. Logistic Regression also assumes that the independent variables are linearly related to the log-odds of the dependent variable and that it allows for direct evaluation of feature contributions through model coefficients. This aspect of Logistic Regression makes it best suited to situations where model explainability is necessary. Nevertheless, its linearity limits its ability to model complex, non-linear relationships between data. Furthermore, it can perform sub optimally when there are correlated features that are of interest or when class distributions are unbalanced unless additional pre-processing or regularization is employed.

4.4 Neural Network

A feedforward multilayer perceptron (MLP) was used to model non-linear relationships [6][19]. Hyperparameters were tuned by inner cross-validation. The MLP used Adam in Python and a quasi-Newton optimizer with weight decay in R.

Neural Networks are highly flexible models, capable of learning complex non-linear relationships within data via hierarchical representation, and have also been optimized with evolutionary and rain-inspired algorithms [17] [18]. This flexibility makes them outperform traditional algorithms on tasks with large and intricate datasets, especially where the interactions among features aren't well defined. However, such flexibility typically comes at the expense of increased risk of overfitting, especially when data are sparse or noisy. Also, Neural Networks generally require painstaking hyperparameter tuning and longer training time, and their internal representations are generally "black box," so they are less interpretable in practice than more transparent models like logistic regression or decision trees.

4.5 Data Preparation

The dataset was prepared in Python and R programming environments and pre-processing process were completed in multiple steps.

First, dependent variables: *MonthlyCharges* and *TotalCharges* were standardize replacing decimal commas by periods (keeping only numbers). Next, 11 rows with NA charge values and the *customerID* variable were removed since it's a primary key without analytical value. Categorical features were then processed using language-dependent strategies, with python's *one-hot encoding* and R's *factor levels*. Independent variables were also scaled for comparability of feature representation (in Python, by calling the *StandardScaler* function; in R *scale*).

Pre-processing steps together enabled training and testing models on both programming environments. It is also interesting to see how the combined effect of all steps can lead from the raw data to a reduced sample size and what impact individual pre-processing steps have on this.

The summary effect in Table 2. shows how many entries remain after big stages of pre-processing.

Table 1. Dataset Size after Each Pre-processing Step

Pre-processing Step	*Rows Remaining*	*Description*
Original dataset	7043	Original data from Kaggle
After numeric conversion and NaN removal in TotalCharges	7032	11 rows removed due to missing TotalCharges values
After dropping customerID	7032	No data loss
After encoding and standardization	7032	No data loss

Although tree-based models do not require scaling, a unified pipeline was used for consistency and because scaling benefits scale-sensitive models such as logistic regression and the multilayer perceptron.

4.6 Model Training and Evaluation

The construction and validation of predictive models require disciplined procedures for training and evaluation. In this study, Decision Trees, Random Forests, Logistic Regression, and a Neural Network (MLP) were implemented in Python and R and trained within a nested cross-validation protocol. The outer 10-fold stratified split provided unbiased test folds, while an inner stratified split was used for hyperparameter tuning and mutual-information top-K feature pruning. All pre-processing (encoding and standardization) was fit on the training portion of each fold only and applied to the corresponding validation/test portion to prevent leakage.

Hyperparameters were optimized in the inner loop with grid search to be tested finally on the outer test folds. For the Decision Trees parameters maximum depth, minimum samples per split/leaf and cost-complexity were tuned. For Random Forest number of

trees, max depth and subsampling of features were tuned. Logistic Regression models were tuned for the choice of penalty type and strength. For the multilayer perceptron (MLP), hyperparameters were hidden-layer size, L2 penalty (weight decay), learning rate, and batch size. The F1-score was preferred metric and class weighting of Logistic Regression and Random Forest was another method to deal with class imbalance.

Feature processing formed part of the modeling pipeline. After localized-number cleaning and removal of rows with missing charges, the identifier (*customerID*) was dropped, categorical predictors were encoded, and numeric features were standardized. Feature ranking by mutual information was computed on the training split only, and atop-K subset (K ∈ {10, 20, 30, all} was selected jointly with the model hyperparameters inside the inner loop.

To evaluate models fairly under class imbalance, multiple metrics were reported for the outer test folds. Overall accuracy measured the proportion of correctly classified instances:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

where TP (True Positives), TN (True Negatives), FP (False Positives), and FN (False Negatives).

The measure Precision reports how correct are the positive predictions by describing what fraction of the predicted positive instances are actually positive. It is defined as:

$$Precision = \frac{TP}{TP + FP}$$

Contrastingly, Recall measures how well the model identifies all actual positive cases, calculated as the ratio of true positives to the sum of true positives and false negatives. It is calculated as:

$$Recall = \frac{TP}{TP + FN}$$

In order to trade-off precision and recall not particularly important in this specific application, we used the F1-Score as their harmonic mean:

$$F1 = 2\frac{Precision * Recall}{Precision + Recall}$$

These metrics together offer a more nuanced interpretation of model performance, considering both the correctness of positive predictions (precision) and the completeness of the identified positives (recall). For each model, the outer-fold mean ± standard deviation was reported for these metrics. In addition, predictions were aggregated across all outer test folds to form a pooled confusion matrix, from which pooled Accuracy, Precision, Recall, and F1 were recomputed as a robustness check.

In customer churn prediction, besides the previously defined and explained basic metrics, it is also useful to include the ROC curve (Receiver Operating Characteristic) together with its related AUC value. The ROC curve shows how the model relates correctly identified churners (true positive rate) to wrongly labeled cases (false positive rate) for several decision thresholds. The AUC helps to see how well the model separates users who are likely to leave from those who will stay. This kind of visual overview is often helpful when presenting results to a broader audience.

In this paper, the AUC and ROC curves were not chosen as the main comparison criteria, because the focus was mainly on the balance between precision and recall, and this can best be seen through the F1 measure. Since the dataset is not strongly imbalanced (around 27% of the cases are labeled "Churn = Yes"), the F1 score was considered a good indicator of how well the model detects users who might leave, without producing too many false alarms. In future work, we plan to include an additional analysis of model behavior using ROC and Precision–Recall curves.

To enable the repetition of the modelling process, the following subsection provides additional implementation details.

4.7 Implementation Details and Reproducibility

For reproducibility, all models were implemented using the scikit-learn library in Python and the caret package in R. In Python, for the RandomForestClassifier and LogisticRegression models, the parameter class_weight='balanced' was used to handle class imbalance by adjusting the weight of each class according to its frequency.

The training data from each inner fold received resampling and rebalancing operations but the outer test folds remained unmodified to stop data leakage. This procedure ensured that preprocessing and parameter tuning were "fold-safe," providing a fair and unbiased model evaluation.

5 Results and Discussion

This section reports performance for Decision Tree, Random Forest, Logistic Regression, and a Neural Network (MLP) under the nested 10-fold protocol. Metrics are computed on the outer test folds only; values are shown as mean ± standard deviation across the 10 folds.

5.1 Evaluation of the Decision Tree Model

The pooled (across the 10 outer folds) confusion matrix shows the counts of true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN). The positive class is Churn = Yes. The convention used is [TP FP; FN TN].

Table 2. Pooled confusion matrix for Decision Tree

Environment	*TP*	*FP*	*FN*	*TN*
Python	991	878	626	4537

R	907	962	500	4663

The classification report is providing many metrics, including F1 score, precision recall, and overall accuracy for positive class (Churn = Yes).

Table 3. Classification report for Decision Tree

Environment	*Accuracy*	*Precision*	*Recall*	*F1*
Python	0.786	0.530	0.613	0.569
R	0.792	0.485	0.645	0.554

Both languages show similar accuracy (≈0.79) with the typical tree trade-off: moderate recall for churners but lower precision. Python's tree has slightly higher precision, while R's has slightly higher recall.

5.2 Evaluation of the Random Forest Model

Table 4. Pooled confusion matrix for Random Forest

Environment	*TP*	*FP*	*FN*	*TN*
Python	1303	566	971	4192
R	980	889	529	4634

Table 5. Classification report for Random Forest

Environment	*Accuracy*	*Precision*	*Recall*	*F1*
Python	0.781	0.697	0.573	0.629
R	0.798	0.524	0.649	0.580

Random Forest reduces variance relative to a single tree. Python emphasizes precision (fewer false positives), while R attains higher recall (more churners found). Overall F1 is competitive in both.

5.3 Evaluation of the Logistic Regression Model

Table 6. Pooled confusion matrix for Logistic Regression

Environment	*TP*	*FP*	*FN*	*TN*
Python	1497	372	1397	3766
R	1030	839	542	4621

Table 7. Classification report for Logistic Regression

Environment	*Accuracy*	*Precision*	*Recall*	*F1*
Python	0.748	0.801	0.517	0.629
R	0.804	0.551	0.655	0.599

A valid baseline is still Logistic. Python favors precision (very few false positives), while R balances recall and accuracy better. Across languages, this model is among the top two models by F1.

5.4 Evaluation of the Neural Network (MLP) Model

Table 8. Pooled confusion matrix for Neural Network (MLP)

Environment	*TP*	*FP*	*FN*	*TN*
Python	975	894	499	4664
R	920	949	537	4626

Table 9. Classification report for Neural Network (MLP)

Environment	*Accuracy*	*Precision*	*Recall*	*F1*
Python	0.802	0.522	0.662	0.583
R	0.789	0.492	0.631	0.553

The MLP captures non-linear patterns and attains accuracy around ~0.79-0.80, with recall higher than precision. Performance is slightly below Logistic Regression and Random Forest on F1, consistent with tabular churn tasks without heavy rebalancing.

5.5 Overall comparison of models

Performance across Python and R for all four models is summarized in Table 11 below, displaying the outer-fold mean results for Accuracy, Precision, Recall and F1-score. Although complete results for every model are given in the previous subsections, this summary facilitates a comparison of strengths. Logistic Regression had the best balance of overall accuracy and F1-score, followed closely by Random Forest. Decision Trees and Neural Networks (MLP) got less F1 but maintained similar accuracies (~0.79–0.80). These findings demonstrate that even interpretable, simple models can achieve competitive performance in tabular churn prediction.

Table 10. Summary of model performance across Python and R

Model	*Environment*	*Accuracy*	*Precision*	*Recall*	*F1*
Decision Tree	Python	0.786	0.530	0.613	0.569

Decision Tree	R	0.792	0.485	0.645	0.554
Random Forest	Python	0.781	0.697	0.573	0.629
Random Forest	R	0.798	0.524	0.649	0.580
Logistic Regression	Python	0.748	0.801	0.517	0.629
Logistic Regression	R	0.804	0.551	0.655	0.599
Neural Network MLP	Python	0.802	0.522	0.662	0.583
Neural Network MLP	R	0.789	0.492	0.631	0.553

6 Conclusion

This study has systematically construct and empirically examined each of four churn prediction models (Decision Tree, Random Forest, Logistic Regression and Neural Network (MLP). A nested 10-fold cross validation with inner-loop hyperparameter optimization, mutual-information based top-K feature selection and leakage-safe pre-processing was implemented.

The results were validated on multiple programming contexts. Best performance was achieved by Logistic Regression, where R has F1 ≈ 0.60 ± 0.03 and Accuracy ≈ 0.80 ± 0.01. In Python it scores F1 ≈ 0.63 with similar Accuracy ≈ 0.75. Random Forest came very close behind (Python F1 ≈ 0.63, Accuracy ≈ 0.78). Decision Tree and Neural Network (MLP) achieved lower (≈ 0.55) F1 scores however still retained high accuracy (~0.79–0.80). The aggregated metrics derived from the confusion matrix of the outer folds exhibit minimal variation from their respective means, indicating that the estimates are both stable and unbiased.

Code for the proposed distance metric is provided in both Python and R. The comparison across programming languages through fold-safe computation results is presented, opening possibilities for further research. The parameters are carefully tuned by validation data and the robustness of these models is ensured with strict evaluation of modeling process, without test set leakage.

From a methodological perspective, regularized Logistic Regression is an interpretable and powerful baseline for this dataset. Random Forest and several tree-based models are still helpful in approximating the non-linear associations and relative importance of features. All predictions will need to be calibrated for probabilities (probability calibration) and thresholds aligned with real-world costs before the model can enter a business workflow for decision. It has been observed that usage of class weights and fold-

safe resampling improves recall. It is recommended that 10-fold cross-validation be used for good model selection and interpretation.

7 Contributions and Future work

This work is novel in two ways. First, a comprehensive cross-language comparison of churn prediction models was conducted using data from multiple companies, whether of being coded by Python or R, all of which reaches equivalent performance, and an approach based on leakage-safe nested cross-validation, performing robust and unbiased evaluation. These contributions enlarge the methodological frontiers of churn prediction research and offer a practical contribution that is useful to both researchers and industry professionals willing to approach reproducible and transferable workflows.

Future research should explore several directions. The propsed pipeline needs to be tested with different datasets, operators and timelines for external and temporal validity. Additionally, a fold-safe rebalancing algorithms and threshold and cost sensitive optimization if reasonable for the real retention rules should be studied.

8 References

[1] Alboukaey, N., Joukhadar, A., & Ghneim, N. (2020). Dynamic behavior based churn prediction in Mobile Telecom. *Expert Systems with Applications*, *162*, 113779. https://doi.org/10.1016/j.eswa.2020.113779

[2] Amin, A., Adnan, A., & Anwar, S. (2023). An adaptive learning approach for customer churn prediction in the telecommunication industry using evolutionary computation and Naïve Bayes. *Applied Soft Computing*, *137*, 110103. https://doi.org/10.1016/j.asoc.2023.110103

[3] Beeharry, Y., & Tsokizep Fokone, R. (2021). Hybrid approach using machine learning algorithms for customers' churn prediction in the telecommunications industry. *Concurrency and Computation: Practice and Experience*, *34*(4). https://doi.org/10.1002/cpe.6627

[4] BlastChar. (2018, February 23). *Telco Customer Churn* [Data set]. Kaggle. https://www.kaggle.com/datasets/blastchar/telco-customer-churn

[5] Bogaert, M., & Delaere, L. (2023). Ensemble methods in customer churn prediction: A comparative analysis of the state-of-the-art. *Mathematics*, *11*(5), 1137. https://doi.org/10.3390/math11051137

[6] Customer churn prediction in telecommunication industry using Deep Learning. (2022). *Information Sciences Letters*, *11*(1), 185–198. https://doi.org/10.18576/isl/110120

[7] Dalli, A. (2022). Impact of hyperparameters on Deep Learning model for customer churn prediction in telecommunication sector. *Mathematical Problems in Engineering*, *2022*, 1–11. https://doi.org/10.1155/2022/4720539

[8] Geiler, L., Affeldt, S., & Nadif, M. (2022). A survey on machine learning methods for churn prediction. *International Journal of Data Science and Analytics*, *14*(3), 217–242. https://doi.org/10.1007/s41060-022-00312-5

[9] Imani, M., & Arabnia, H. R. (2023). Hyperparameter optimization and combined data sampling techniques in machine learning for customer churn prediction: A

comparative analysis. *Technologies*, *11*(6), 167. https://doi.org/10.3390/technologies11060167

[10] Jain, H., Khunteta, A., & Srivastava, S. (2020a). Churn prediction in telecommunication using logistic regression and logit boost. *Procedia Computer Science*, *167*, 101–112. https://doi.org/10.1016/j.procs.2020.03.187

[11] Jain, H., Khunteta, A., & Srivastava, S. (2020b). Telecom churn prediction and used techniques, datasets and performance measures: A Review. *Telecommunication Systems*, *76*(4), 613–630. https://doi.org/10.1007/s11235-020-00727-0

[12] 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). https://doi.org/10.1038/s41598-023-44396-w

[13] Kostić, S. M., Simić, M. I., & Kostić, M. V. (2020). Social network analysis and churn prediction in telecommunications using graph theory. *Entropy*, *22*(7), 753. https://doi.org/10.3390/e22070753

[14] Lalwani, P., Mishra, M. K., Chadha, J. S., & Sethi, P. (2021). Customer churn prediction system: A machine learning approach. *Computing*, *104*(2), 271–294. https://doi.org/10.1007/s00607-021-00908-y

[15] Liu, R., Ali, S., Bilal, S. F., Sakhawat, Z., Imran, A., Almuhaimeed, A., Alzahrani, A., & Sun, G. (2022). An intelligent hybrid scheme for customer churn prediction integrating clustering and classification algorithms. *Applied Sciences*, *12*(18), 9355. https://doi.org/10.3390/app12189355

[16] Manzoor, A., Atif Qureshi, M., Kidney, E., & Longo, L. (2024). A review on machine learning methods for customer churn prediction and recommendations for Business Practitioners. *IEEE Access*, *12*, 70434–70463. https://doi.org/10.1109/access.2024.3402092

[17] Pustokhina, I. V., Pustokhin, D. A., Nguyen, P. T., Elhoseny, M., & Shankar, K. (2021). Multi-objective rain optimization algorithm with WELM model for customer churn prediction in telecommunication sector. *Complex & Intelligent Systems*, *9*(4), 3473–3485. https://doi.org/10.1007/s40747-021-00353-6

[18] Pustokhina, I. V., Pustokhin, D. A., RH, A., Jayasankar, T., Jeyalakshmi, C., Díaz, V. G., & Shankar, K. (2021). Dynamic customer churn prediction strategy for business intelligence using text analytics with evolutionary optimization algorithms. *Information Processing & Management*, *58*(6), 102706. https://doi.org/10.1016/j.ipm.2021.102706

[19] Saha, L., Tripathy, H. K., Gaber, T., El-Gohary, H., & El-kenawy, E.-S. M. (2023). Deep churn prediction method for telecommunication industry. *Sustainability*, *15*(5), 4543. https://doi.org/10.3390/su15054543

[20] Saha, S., Saha, C., Haque, Md. M., Alam, Md. G., & Talukder, A. (2024). ChurnNet: Deep Learning Enhanced Customer Churn prediction in telecommunication industry. *IEEE Access*, *12*, 4471–4484. https://doi.org/10.1109/access.2024.3349950

[21] Saleh, S., & Saha, S. (2023). Customer retention and churn prediction in the telecommunication industry: A case study on a danish university. *SN Applied Sciences*, *5*(7). https://doi.org/10.1007/s42452-023-05389-6

[22] Sana, J. K., Abedin, M. Z., Rahman, M. S., & Rahman, M. S. (2022). A novel customer churn prediction model for the telecommunication industry using data transformation methods and feature selection. *PLOS ONE*, *17*(12). https://doi.org/10.1371/journal.pone.0278095

[23] Sikri, A., Jameel, R., Idrees, S. M., & Kaur, H. (2024). Enhancing customer retention in telecom industry with Machine Learning Driven Churn prediction. *Scientific Reports*, *14*(1). https://doi.org/10.1038/s41598-024-63750-0

[24] Sina Mirabdolbaghi, S. M., & Amiri, B. (2022). Model optimization analysis of customer churn prediction using machine learning algorithms with focus on feature reductions. *Discrete Dynamics in Nature and Society*, *2022*, 1–20. https://doi.org/10.1155/2022/5134356

[25] Sudharsan, R., & Ganesh, E. N. (2022). A swish RNN based customer churn prediction for the telecom industry with a novel feature selection strategy. *Connection Science*, *34*(1), 1855–1876. https://doi.org/10.1080/09540091.2022.2083584

[26] Usman-Hamza, F. E., Balogun, A. O., Capretz, L. F., Mojeed, H. A., Mahamad, S., Salihu, S. A., Akintola, A. G., Basri, S., Amosa, R. T., & Salahdeen, N. K. (2022). Intelligent Decision Forest models for customer churn prediction. *Applied Sciences*, *12*(16), 8270. https://doi.org/10.3390/app12168270

[27] V. Kavitha, G. Hemanth Kumar, S. V Mohan Kumar, & M. Harish. (2020). Churn prediction of customer in telecom industry using machine learning algorithms. *International Journal of Engineering Research And*, *V9*(05). https://doi.org/10.17577/ijertv9is050022

[28] Wu, S., Yau, W.-C., Ong, T.-S., & Chong, S.-C. (2021). Integrated churn prediction and customer segmentation framework for telco business. *IEEE Access*, *9*, 62118–62136. https://doi.org/10.1109/access.2021.3073776

[29] Xiahou, X., & Harada, Y. (2022). B2C e-commerce customer churn prediction based on K-means and SVM. *Journal of Theoretical and Applied Electronic Commerce Research*, *17*(2), 458–475. https://doi.org/10.3390/jtaer17020024

[30] Xu, T., Ma, Y., & Kim, K. (2021). Telecom churn prediction system based on ensemble learning using feature grouping. *Applied Sciences*, *11*(11), 4742. https://doi.org/10.3390/app11114742

[31] Zdravevski, E., Lameski, P., Apanowicz, C., & Ślęzak, D. (2020). From Big Data to business analytics: The case study of churn prediction. *Applied Soft Computing*, *90*, 106164. https://doi.org/10.1016/j.asoc.2020.106164

Transforming Generic Flyers into Tailored Promotions: A Case Study in AI-Powered Grocery Retail

Dr. Tomaž Aljaž, Uroš Kosanovič
Spar Slovenia
Letališka cesta 26, 1000 Ljubljana, Slovenia
tomaz.aljaz@spar-ics.com, uros.kosanovic@spar.si

Abstract: *Retailers traditionally rely on generic promotional catalogues that provide identical offers to all customers, regardless of their purchase histories or preferences. This "one-size-fits-all" approach often results in low engagement and inefficient allocation of marketing resources. This study investigates how cloud computing and artificial intelligence (AI) can automate the generation of personalized promotional offers, improving both customer relevance and retailer efficiency.*
Using a mixed-methods design science approach, requirements were gathered through interviews with marketing and IT staff, and a recommender system was developed by integrating ERP data, transaction histories, and product attributes into a private cloud environment. From a weekly pool of over 150 promotions, including more than 600 products in promotion, the system generated personalized lists of 6–20 promotions per customer, delivered via email, mobile applications, and print-ready flyers. A 12-week randomized controlled trial involving 400,000 loyalty program members evaluated the solution's impact.
Results show that customers receiving personalized offers generated 10% higher sales compared with the control group. Moreover, a positive correlation was observed between exposure frequency and both shopping frequency and basket growth, ranging from 6% to 36%. Marketing processes were standardized, improving campaign quality and consistency while reducing manual variability. Qualitative feedback confirmed that customers perceived the offers as more relevant and convenient, while highlighting the importance of clear communication about data usage, even though privacy is formally managed through the loyalty program.
This study provides empirical evidence of the measurable business impact of hybrid recommender systems in grocery retail and emphasizes the importance of transparency, governance, and trust for successful AI-driven personalization.

Key Words: *artificial intelligence, cloud computing, grocery retail, personalization, recommender systems*

1 Introduction

Retailers have traditionally relied on weekly catalogues and price promotions to stimulate consumer demand [1] [2]. These catalogues, whether printed or distributed via email, typically present a broad selection of over 150 promotions, representing more than 600 products, designed to appeal to the widest possible audience, as shown in Figure 1. While effective in generating short-term revenue, such generic promotions risk creating promotional fatigue when offers lack relevance to individual customers, leading to

diminished engagement and inefficient marketing expenditure. Moreover, generic promotions can weaken the perceived value of loyalty programs, which are intended to foster behavioral loyalty through relevant benefits [13].

Figure 1: Weekly leaflet

The increasing digitalization of commerce has shifted customer expectations toward more personalized and data-driven experiences. Advances in information technology (IT), particularly in cloud computing and artificial intelligence (AI), now enable retailers to mine large volumes of transactional data and deliver individualized promotions at scale. Affordable computing power and elastic infrastructure in the cloud have lowered barriers for retailers to deploy sophisticated recommendation systems that learn from purchase histories, demographic characteristics, and seasonal patterns [3].

Despite these technological possibilities, many grocery retailers continue to rely on manual, generic campaigns. These catalogues require considerable staff effort to compile yet often include products irrelevant to segments of the customer base. This inefficiency not only wastes marketing resources but may also undermine the perceived value of loyalty programs. The COVID-19 pandemic further accelerated the shift to online and mobile shopping, fundamentally altering consumer expectations for timely and personalized promotional content [4], [14]. Although recommender systems are well established in academic literature, few studies document their deployment and evaluation in grocery retail practice, where weekly catalogues remain dominant.

This paper examines how a grocery retailer in Slovenia digitally transformed its promotional practices through the deployment of a AI-based recommender system. The study contributes to the literature on personalized marketing, cloud architecture, and AI by demonstrating how these techniques can be combined into an end-to-end solution in a real-world context.

The remainder of the paper is structured as follows. Section 2 reviews related work on personalized marketing and enabling technologies. Section 3 defines the research problem and objectives. Section 4 details the methodology and system architecture. Section 5 presents the results of the case study. Section 6 discusses the implications and limitations, and Section 7 concludes with directions for future research.

2 Literature review

Early attempts to personalize promotions relied on rule-based expert systems and market basket analysis (MBA). The seminal work of Agrawal, Imieliński, and Swami [5] introduced association rule mining, enabling retailers to identify frequently co-purchased item pairs and recommend complementary products. Building on this foundation, Olson and Lauhoff [6] describe how MBA techniques can be applied in retail analytics to uncover actionable cross-selling opportunities.

With the growth of data availability, machine learning has driven the development of data-driven recommender systems capable of uncovering complex patterns in customer preferences. Collaborative filtering approaches, which infer a user's interests based on similarities with other users, have become widely adopted in e-commerce platforms [7]. In contrast, content-based methods rely on product attributes and individual user profiles. To overcome the limitations of these approaches, particularly the cold-start problem, hybrid recommender systems that integrate collaborative and content-based features have been proposed and empirically validated [8]. This study builds on this stream of research by implementing and evaluating a hybrid model in a real-world grocery retail setting, thereby providing empirical evidence of its business impact.

The rise of cloud computing has been critical for scaling these systems. Public, private, and hybrid cloud infrastructures provide on-demand computing and storage resources for processing vast transactional datasets and delivering real-time recommendations. Zhang, Yao, Sun, and Tay [9] emphasize that cloud-based and distributed architectures make it feasible for retailers to deploy sophisticated personalization engines. Furthermore, edge computing has been explored to reduce latency in mobile applications, although core recommendation algorithms generally remain centralized in data centers.

Evidence from practice underscores the business value of personalization. For instance, Park, Park, and Schweidel [10] found that mobile coupon promotions in a supermarket context significantly increased purchase likelihood and basket value, with free-sample coupons delivering enduring sales effects. Similarly, Kim, Choi, and Li [11] demonstrate that deep learning–based recommendation models balancing accuracy and diversity achieve higher levels of customer satisfaction compared with traditional algorithms. More recent findings suggest that personalized recommendations enhance the trust–satisfaction–loyalty chain in digital commerce [12].

Despite these successes, challenges persist. Issues such as data sparsity, seasonality, privacy concerns, and the need for continuous model maintenance remain central research topics [9]. Proposed solutions include incremental learning techniques, federated learning, and the integration of contextual signals (e.g., time of day, location, device) into recommendation models.

Our case study contributes to this evolving body of knowledge by presenting an end-to-end solution that integrates a private cloud environment, an ERP system, and a machine learning platform for delivering scalable and secure personalized promotions in the grocery retail sector.

3 Problem definition

The retailer examined in this study currently distributes a weekly catalogue containing over 150 promotions, including more than 600 products. The catalogue is created manually by selecting products from the enterprise resource planning (ERP) system and formatting a leaflet that is subsequently printed and distributed to loyalty program

members via email and mobile application. This process results in generic, undifferentiated promotional content, whereby all customers receive the same offers regardless of their purchase histories or preferences. Consequently, mismatches occur. For example, vegetarian customers may receive offers for meat products, or households that have recently purchased bulk quantities of household staples may be targeted with redundant promotions. Such inefficiencies lead to low customer engagement and suboptimal allocation of marketing resources.

This situation highlights a significant research gap: although the retailer possesses rich transactional and product data, these assets remain underutilized in guiding marketing decisions. The primary research question addressed in this study is therefore:

To what extent can cloud computing and AI be leveraged to automate the generation of personalized promotional offers, and how do such offers impact customer engagement and marketing efficiency compared with traditional generic campaigns?

To address this question, the study sets the following objectives:

- To design a system architecture that integrates the retailer's ERP system, historical transaction database, and product catalogue with a private cloud platform capable of running AI algorithms.
- To develop a recommendation model that learns from customers' purchase patterns and product attributes to generate individualized offers based on weekly promotions.
- To implement a multi-channel delivery mechanism (email, mobile application, and printed leaflets) to distribute personalized promotions.
- To evaluate the effectiveness of the solution by comparing engagement metrics and marketing efficiency with baseline generic campaigns.

4 Methodology

To evaluate the proposed solution and ensure its practical relevance, a structured research process was adopted. The methodology follows the principles of design science research, which emphasizes building and evaluating artefacts that address real organizational challenges. This approach was particularly appropriate given the dual objectives of this study: to design a technically robust recommender system and to assess its measurable impact on customer behavior and sales performance. The following subsections outline the research approach, system architecture, and evaluation procedure in detail.

4.1 Research approach

This study employed a mixed-methods design science approach, combining qualitative design activities with quantitative evaluation. This approach was selected to ensure that the solution was both scientifically rigorous and organizationally relevant. The research process followed four main phases:

- Knowledge base review: A structured review of the literature and best practices in recommender systems, cloud-based analytics, and digital personalization (Section 2) to establish a theoretical foundation.
- Requirement elicitation: Semi-structured interviews with marketing, IT, and data analytics teams to capture functional and non-functional requirements for the system, with particular emphasis on integration with existing ERP data and privacy-compliant handling of loyalty program information.
- System design and implementation: Development of a prototype recommender system that generates personalized promotional lists from a pool of over 150

promotions, covering more than 600 products. The system was deployed in a private cloud environment to ensure scalability and data sovereignty.

- Evaluation: A 12-week randomized controlled trial involving 400,000 loyalty program members was conducted. The primary quantitative metric was sales growth, comparing customers who received and engaged with personalized promotions against those who did not view them. In addition, the analysis examined the relationship between exposure frequency and customer behavior, measuring changes in shopping frequency and basket growth rate.

This approach ensured that the system was developed to meet real organizational needs and that its effectiveness was measured based on business impact rather than intermediate digital engagement metrics.

4.2 System architecture

The proposed solution integrates existing enterprise systems with cloud-based analytics and machine learning components.

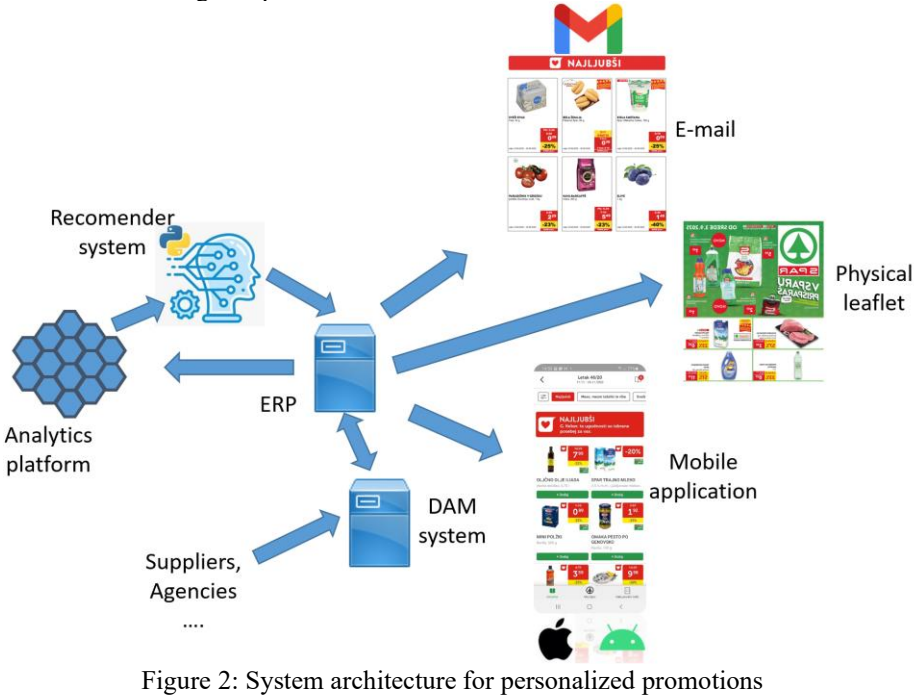

Figure 2: System architecture for personalized promotions

It comprises four main layers, as shown in Figure 2:

- ERP and transaction layer: The retailer's ERP system stores product information, inventory levels and transaction records. A data extraction module periodically transfers purchase histories and product attributes to the analytics platform.
- Data Asset Management (DAM) system maintains a repository of product images for marketing use.
- Analytics platform: A private cloud environment hosts a scalable analytics platform. The platform performs data cleaning and feature engineering. The choice of a private

cloud balances scalability with data sovereignty, ensuring sensitive customer information remains under retailer control.

- Recommendation engine – A hybrid recommendation model was implemented, combining collaborative filtering with content-based filtering.
    1. Collaborative filtering: User–item interactions were represented in a sparse matrix, with matrix factorization used to learn latent factors capturing similarities among customers and products.
    2. Content-based filtering: Product attributes such as category, price range, and nutritional tags were encoded as feature vectors.
    3. Hybrid model: A neural network combined latent and content features to predict customer–product relevance scores. Cold-start issues for new products were mitigated by weighting content-based features more heavily until sufficient interaction data accumulated.
    4. Bias mitigation: Historical campaign performance and A/B testing results were used to fine-tune the model and reduce systematic biases.
- Delivery channels: From the pool of over 600 weekly promotional products, the system generated tailor-made lists of 6-20 promotions per customer. These lists represented a new service introduced by the retailer, shifting from a generic catalogue toward individualized offers. Personalized lists were assembled into HTML emails, mobile application, and print-ready flyers via a template engine. The campaign management module handled scheduling, channel preferences, and event tracking (open rates, exposures, purchases).

4.3 Evaluation procedure

The evaluation followed a randomized controlled trial (RCT) design over a 12-week period. A total of 400,000 loyalty program members were randomly assigned into two groups:

- Control group: continued receiving the generic weekly catalogue with over 150 promotions, including more than 600 products.
- Treatment group: received personalized lists of 6–20 promotions generated by the recommender system, delivered through their preferred channels (email, mobile application, or print-ready flyer).

Randomization was stratified by demographic and behavioral attributes (e.g., household size, past purchase frequency) to ensure the two groups were comparable at baseline.

The primary outcome measure was sales growth, calculated as the percentage difference in total revenue per customer between the treatment and control groups. Two secondary metrics were evaluated:

- Shopping frequency: change in the number of shopping trips per customer during the evaluation period.
- Basket growth rate: change in the average value of a customer's basket compared to baseline.

In addition, we analyzed the relationship between exposure frequency and customer behavior by grouping treatment customers based on the number of times they viewed personalized promotions. Correlation analysis was conducted to determine whether increased exposure was associated with higher purchase frequency and sales growth.

Sales, frequency, and basket data were aggregated at the customer level. Independent samples tests were applied to compare treatment and control groups, and Pearson

correlation coefficients were used to evaluate the association between exposure frequency and behavioral outcomes. Statistical significance was set at $p < 0.05$.

4.4 Limitations

While the study applied a rigorous mixed-methods approach, several limitations should be acknowledged.

First, the evaluation was conducted with a single retailer in the grocery sector, which may limit the generalizability of findings to other retail formats such as fashion, electronics, or specialty goods. Future studies should validate the proposed architecture across diverse retail contexts.

Second, the 12-week timeframe restricted the ability to capture long-term customer behavior shifts and seasonality effects. Extended deployments would provide deeper insights into customer loyalty, repeated use of personalized promotions, and the sustainability of observed benefits.

Third, while privacy and data protection are fully managed through the retailer's loyalty program, where customers explicitly consent to the use of their personal data for promotional purposes, future research could explore additional privacy-preserving machine learning techniques (e.g., differential privacy, federated learning). These approaches could further strengthen customer trust and offer insights into how to balance personalization accuracy with enhanced data protection.

By recognizing these limitations, the study provides a transparent basis for interpreting the findings and highlights opportunities for further research.

5 Results

The following sections present the results of the 12-week randomized controlled trial, structured around the research objectives. The analysis covers system performance, sales impact, behavioral changes, and process improvements.

5.1 System architecture and efficiency

One of the key objectives of this project was to design and implement a system architecture that integrates the retailer's enterprise systems with a private cloud analytics platform. The implemented solution achieved this goal and, most importantly, standardized the process of leaflet and campaign creation. By embedding product data, layouts, and images into a unified workflow, the system ensures that both the generic weekly catalogue and the newly introduced personalized lists are generated using the same templates and procedures.

This standardization reduced inconsistencies, eliminated manual rework, and improved the overall quality and branding consistency of campaign materials. Marketing staff reported that the automated workflow provided greater clarity and predictability in leaflet preparation, making the campaign process more efficient and less error prone. They highlighted that this freed time for focusing on creative and strategic aspects of campaign design rather than repetitive formatting tasks. At the same time, they emphasized the importance of ongoing monitoring to ensure that the automatically generated content remains aligned with branding guidelines and current marketing priorities.

5.2 Sales Impact

The evaluation demonstrated a significant positive impact on sales performance as a result of introducing personalized promotions. During the 12-week randomized controlled trial,

customers in the treatment group—who received personalized lists of 6–20 promotions—generated 10% higher sales compared with customers in the control group who did not view personalized promotions.
This sales lift provides robust evidence that tailoring offers from the existing pool of over 150 promotions (600+ products) weekly promotional products can meaningfully influence purchasing behavior. The result confirms that personalization not only engages customers but also drives measurable business outcomes in terms of revenue growth.

5.3 Shopping Frequency and Exposure Analysis

Beyond overall sales growth, the analysis revealed a clear behavioral shift among customers exposed to personalized promotions. A positive correlation was observed between the number of times customers viewed personalized promotions and their subsequent shopping frequency and basket growth rate.
Specifically, customers with the highest exposure to personalized lists demonstrated increases in shopping frequency and basket value ranging from 6% to 36% relative to their baseline purchasing behavior. This finding suggests that repeated exposure to relevant, targeted offers does more than generate one-time purchases — it encourages more frequent store visits and gradually increases overall spending per visit.
These results reinforce the potential of personalization as a long-term engagement driver, demonstrating that the system not only produces an immediate sales uplift but also contributes to more sustained customer participation in promotional campaigns.

5.4 Multi-channel delivery performance

The multi-channel delivery component performed reliably throughout the evaluation period. Personalized lists were successfully distributed via email, mobile application notifications, and print-ready flyers, ensuring broad customer coverage regardless of preferred channel.
The campaign management module correctly applied channel preferences, guaranteeing that customers received offers through their chosen medium. This not only improved customer experience but also provided the data granularity needed to track exposure frequency at the individual level. Such tracking was essential for the correlation analysis that linked repeated promotion views to increased shopping frequency and sales growth.
Marketing staff highlighted that the unified campaign management interface simplified scheduling and allowed them to monitor delivery performance across all channels in a consistent manner. This improved visibility enabled faster detection of issues and ensured that campaigns remained synchronized across digital and physical touchpoints.

5.5 Overall impact

Taken together, the results demonstrate that the implementation of the personalized promotion system achieved both operational and business objectives.
On the operational side, the system successfully standardized the campaign creation process, improving quality, consistency, and collaboration between marketing and IT teams. This unification of workflows for generic catalogues and personalized lists reduced manual rework and allowed marketing staff to focus more on strategic planning and creative design.
On the business side, the introduction of personalized lists delivered a measurable 10% sales uplift and a clear behavioral shift among customers, with shopping frequency and basket values increasing in proportion to the number of times personalized promotions were viewed (6–36% growth range). These results confirm that personalization not only

enhances campaign relevance but also drives sustainable revenue growth when deployed at scale.

Collectively, these outcomes show that the solution is a viable, scalable approach for transforming a retailer's traditional promotion process into a data-driven, customer-centric service, generating value for both the business and its customers.

5.6 Summary of findings

The evaluation confirms that the personalized promotion system successfully met the research objectives. The solution not only standardized and improved the campaign preparation process but also delivered measurable business outcomes, including a 10% increase in sales and significant growth in shopping frequency and basket size among highly exposed customers.

Table 1: Summary of research objectives, evidence, and outcomes

Research Objective	Key Evidence	Outcome
Design and implement a system architecture integrating ERP, transaction data, and private cloud	Four-layer architecture deployed in private cloud; ensured data sovereignty and integration with loyalty program consent process.	Achieved Functional, secure and compliant
Generate personalized promotions from 600+ weekly offers	Automated production of 6–20 item personalized lists per customer using unified workflow.	Achieved Personalization service fully operational
Implement multi-channel delivery (email, app, print)	Campaign engine supported email, app, and print channels; honored customer preferences; provided exposure tracking.	Achieved Reliable multi-channel execution
Measure business impact	Sales lift of 10% compared to control group; positive correlation between exposure frequency and shopping frequency/basket growth (6–36%).	Achieved Clear evidence of revenue and engagement impact
Standardize campaign preparation process	Unified templates and workflows reduced manual variability and improved brand consistency.	Achieved Enhanced process quality and predictability

6 Discussion

The findings of this study provide clear evidence that personalized promotions, enabled by a cloud-based recommender system, deliver a tangible business impact in grocery retail. Customers who received personalized lists of 6–20 items drawn from the weekly pool of 150+ promotions generated 10% higher sales compared with customers in the control group. Moreover, analysis revealed a positive, exposure-dependent behavioral shift, with shopping frequency and basket growth improving between 6% and 36% among customers who repeatedly viewed personalized offers.

From an operational perspective, the solution standardized the campaign preparation process, ensuring consistent templates, layouts, and workflows across both generic catalogues and personalized lists. This process of improvement reduced variability and improved quality, allowing marketing teams to focus more on strategic planning and creative campaign development.

These findings confirm prior research demonstrating the advantages of data-driven personalization over generic approaches [5], [7], [8] and extend the literature by providing evidence of sales uplift and behavioral change in a real-world grocery retail environment. By showing that personalization can both increase revenue and standardize internal workflows, the study suggests that such systems create dual value: enhanced customer outcomes and improved operational discipline.

Privacy and data protection were fully managed through the retailer's loyalty program, where customers explicitly consent to the use of their data for promotional purposes. Nevertheless, the results highlight the importance of transparent communication and algorithmic governance to maintain customer trust. Regular monitoring and bias detection remain essential to ensure that personalization remains fair and aligned with customer expectations [9], [12].

The following subsections elaborate on the theoretical implications, managerial implications, and directions for future research.

6.1 Theoretical implications

This research contributes to the recommender systems literature by providing empirical evidence that hybrid recommendation models can deliver measurable business impact in a real-world grocery retail setting. While hybrid recommenders—combining collaborative filtering and content-based methods—have been widely studied in theory [8], relatively few studies report large-scale deployments that quantify revenue impact and behavioral change. The observed 10% sales uplift and the positive, exposure-dependent growth in shopping frequency and basket size strengthen the argument that hybrid models are not only effective in overcoming cold-start problems but also capable of driving sustainable business outcomes when applied to high-frequency retail environments.

In addition, the study advances the design science perspective by demonstrating how ERP data, transaction histories, and product attributes can be integrated within a private cloud architecture to satisfy both functional requirements (accuracy, scalability) and non-functional requirements (data protection, compliance, and operational integration). The proposed architecture serves as a reference design for practitioners and researchers, illustrating how theoretically sound recommender models can be operationalized into an enterprise-grade, privacy-compliant system that generates consistent, repeatable promotional content.

6.2 Managerial implications

From a managerial perspective, the findings provide a compelling business case for investing in personalization technologies. The implementation of the system standardized campaign preparation, reducing variability and manual rework while improving brand consistency. This process improvement allows marketing teams to operate more efficiently and shift their focus toward higher-value activities such as campaign design, creative content development, and strategic planning.

The 10% sales uplift observed during the trial demonstrates that personalization directly contributes to revenue growth, making it a high-impact initiative for retailers seeking to improve return on marketing expenses. The positive correlation between repeated exposure to personalized offers and increased shopping frequency (6–36%) further indicates that personalization can foster longer-term engagement and strengthen customer loyalty over time.

To fully capture these benefits, managers must also prioritize customer trust and transparency. Although privacy is already managed through the loyalty program, it is essential to communicate clearly how customer data is used and safeguarded. Transparent messaging helps maintain confidence in the personalization service and ensures continued customer participation.
Finally, algorithmic governance is critical. Without proper oversight, recommendation models may unintentionally favor high-margin products or create repetitive patterns that reduce customer satisfaction. Establishing regular audits, bias detection processes, and ethical guidelines for personalization strategies will be essential for sustaining trust and ensuring the system remains aligned with both customer needs and business goals.

6.3 Limitations and Future Research

Several limitations of this study must be acknowledged. First, the evaluation was conducted with a single grocery retailer, which may limit the generalizability of the findings to other retail formats such as apparel, electronics, or specialty goods. Future research should replicate this approach across different retail sectors to test its transferability and explore whether similar personalization effects occur in lower-frequency or higher-margin purchase environments.
Second, the evaluation period of 12 weeks constrained the ability to observe long-term impacts such as sustained loyalty, seasonality effects, and customer lifetime value. Extending the duration of future experiments would provide deeper insights into whether the observed sales uplift and increased shopping frequency are maintained over time.
Third, while this study relied entirely on behavioral data rather than self-reported survey measures, strengthening objectivity, it did not investigate the psychological mechanisms driving the observed behavioral changes (e.g., perceived relevance, habit formation). Future research could complement behavioral analysis with customer interviews or experiments designed to measure motivational factors and attitudinal shifts.
Finally, while privacy and data protection are fully managed through the retailer's loyalty program, where customers explicitly consent to the use of their personal data for promotional purposes, future work could examine additional strategies for reinforcing trust. For example, research could investigate the impact of transparency dashboards, opt-out controls, and privacy-preserving machine learning techniques such as federated learning on customer acceptance and perceived fairness of personalization.
By addressing these areas, future studies could deepen understanding of both the short- and long-term effects of personalization and provide further guidance for retailers seeking to implement ethically responsible and customer-centric AI solutions.

7 Conclusion

This study set out to address the limitations of generic, undifferentiated promotions in grocery retail and to explore how cloud computing and artificial intelligence (AI) can be leveraged to generate personalized promotional offers. By designing, implementing, and evaluating a recommender system, the research demonstrated that personalization can meaningfully transform the retailer's promotional strategy from a static weekly catalogue into a customer-centric, data-driven service.
The results show that, from a weekly catalogue of more than 600 discounted products, the system successfully generated tailor-made lists of 6–20 promotions per customer, delivered through email, mobile applications, and print-ready flyers. Customers who received personalized offers generated 10% higher sales compared with the control group and exhibited a positive, exposure-dependent increase in shopping frequency and basket

size ranging from 6% to 36%. On the operational side, campaign preparation processes were standardized and streamlined, improving consistency and enabling marketing teams to focus on creative and strategic tasks rather than manual compilation.

This research makes several contributions. Theoretically, it validates the effectiveness of hybrid recommender models that combine collaborative and content-based filtering in a real-world grocery retail setting, demonstrating their ability to deliver measurable business outcomes. Methodologically, it shows how a design science approach can integrate ERP, transaction, and product data within a private cloud infrastructure while maintaining compliance with privacy requirements through loyalty program consent. Practically, it offers a reference implementation that demonstrates how retailers can achieve dual benefits: measurable revenue growth and improved internal process efficiency.

At the same time, the study highlights ongoing considerations for responsible AI deployment. While privacy is formally managed through the retailer's loyalty program, transparent communication about data usage remains essential to maintaining customer trust. Furthermore, regular monitoring and bias detection should be institutionalized to ensure recommendations remain fair, relevant, and aligned with customer expectations.

Looking ahead, future research should replicate these findings across diverse retail contexts, extend the evaluation period to capture seasonal effects and long-term loyalty impacts, and explore advanced privacy-preserving techniques such as federated learning or differential privacy. Such efforts would further strengthen the case for scalable, ethically sound personalization systems.

In conclusion, this research demonstrates that cloud-based AI recommender systems can transform grocery retail marketing by delivering measurable improvements in sales, customer engagement, and process efficiency. Their full potential will be realized when technological innovation is matched with robust governance, transparency, and sustained commitment to customer trust.

8 References

[1] *Aguilar-Barrientos, S. (2021). Pricing and promotion: A literature review. Aibi Revista De Investigación, Administración e Ingeniería. https://doi.org/10.15649/2346030X.2587*

[2] *Grewal, D. (2011). Innovations in Retail Pricing and Promotions. Journal of Retailing. https://doi.org/10.1016/J.JRETAI.2011.04.008*

[3] *López Echeverry, A.M., Velásquez Isaza, J.M., López-Flórez, S., De la Prieta, F. (2025). AI-Based Recommendation System for the Retail Industry. In: Novais, P., et al. Ambient Intelligence – Software and Applications – 15th International Symposium on Ambient Intelligence. ISAmI 2024. Lecture Notes in Networks and Systems, vol 1279. Springer, Cham. https://doi.org/10.1007/978-3-031-83117-1_6*

[4] *Dabija, DC., Câmpian, V., Philipp, B. et al. How did consumers retail purchasing expectations and behaviour switch due to the COVID-19 pandemic?. J Market Anal (2024). https://doi.org/10.1057/s41270-024-00344-9*

[5] *R. Agrawal, T. Imieliński, and A. Swami, "Mining association rules between sets of items in large databases," Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, pp. 207–216, 1993. https://doi.org/10.1145/170035.170072*

[6] *D. L. Olson and G. Lauhoff, Market Basket Analysis. In Descriptive Data Mining, pp. 31–44. Springer, 2019. https://doi.org/10.1007/978-981-13-7181-3_3*

[7] X. Su and T. M. Khoshgoftaar, "A survey of collaborative filtering techniques," Advances in Artificial Intelligence, vol. 2009, Article ID 421425, 2009. https://doi.org/10.1155/2009/421425

[8] R. Burke, "Hybrid recommender systems: Survey and experiments," User Modeling and User-Adapted Interaction, vol. 12, no. 4, pp. 331–370, 2002. https://doi.org/10.1023/A:1021240730564

[9] S. Zhang, L. Yao, A. Sun, and Y. Tay, "Deep learning based recommender system: A survey and new perspectives," ACM Computing Surveys, vol. 52, no. 1, pp. 1–38, 2019. https://doi.org/10.1145/3285029

[10] C. H. Park, Y.-H. Park, and D. A. Schweidel, "The effects of mobile promotions on customer purchase dynamics," International Journal of Research in Marketing, vol. 35, no. 3, pp. 453–470, 2018. https://doi.org/10.1016/j.ijresmar.2018.05.001

[11] J. Kim, I. Choi, and Q. Li, "Customer satisfaction of recommender system: Examining accuracy and diversity in several types of recommendation approaches," Sustainability, vol. 13, no. 11, 6165, 2021. https://doi.org/10.3390/su13116165

[12] N. Hassan, M. Abdelraouf, and D. El-Shihy, "The moderating role of personalized recommendations in the trust–satisfaction–loyalty relationship: An empirical study of AI-driven e-commerce," Future Business Journal, vol. 11, no. 66, 2025. https://doi.org/10.1186/s43093-025-00476-z

[13] J. Leenheer, T. H. A. Bijmolt, H. J. van Heerde, and A. Smidts, "Do loyalty programs really enhance behavioral loyalty? An empirical analysis accounting for self-selecting members," International Journal of Research in Marketing, vol. 24, no. 1, pp. 31–47, 2007. https://doi.org/10.1016/j.ijresmar.2006.10.005

[14] H. R. Abbu, D. Fleischmann, and P. Gopalakrishna, "The Digital Transformation of the Grocery Business—Driven by Consumers, Powered by Technology, and Accelerated by the COVID-19 Pandemic," in Trends and Applications in Information Systems and Technologies, AISC, vol. 1367, WorldCIST 2021, Springer, 2021, pp. 329–339. https://doi.org/10.1007/978-3-030-72660-7_32

Qualitative User Evaluation of an Intelligent HR Analytics Prototype for Absence Data

Information Technologies and Information Society (ITIS)

Peter Zupančič, Panče Panov
Faculty of Information Studies
Ljubljanska cesta 31A, 8000 Novo mesto, Slovenia
peter.zupancic, pance.panov}@fis.unm.si

Abstract: *This paper presents an qualitative evaluation of a prototype tool developed as part of a doctoral research project. The main goal was to assess how effectively the tool supports users in interpreting employee data and making informed HR decisions. The evaluation followed a multi-stage process, starting with an initial survey to gather user impressions, identify issues, and collect improvement suggestions. Based on these findings, additional targeted surveys and semi-structured interviews were conducted to gain deeper qualitative insights into user perceptions and practical use. This approach provided a comprehensive understanding of the tool's usefulness and relevance. The results indicate which data visualizations and analytical outputs users found most helpful for their daily work, while also identifying areas where further refinement could improve clarity, interpretability, and decision support.*

Key Words*: empirical evaluation, decision support, human resource analytics*

1 Introduction

In modern organisations, the effective management of human capital has become one of the most critical factors for achieving competitive advantage and long-term sustainability.

In this paper, which forms part of a broader doctoral research project, we present the evaluation of the developed prototype solutions and discuss the results from the perspective of end users. The purpose of the evaluation was to examine how the prototype tool performs in practice, with particular attention to its usability and the overall user experience. The process began with a survey [1], which served not only to collect initial feedback but also to generate new ideas and suggestions for improvement. The results of this first stage informed the design of the subsequent evaluation steps. To gain deeper insights, additional surveys and semi-structured interviews were carried out with selected users. This combination of methods provided both a broader overview and a more detailed understanding of user expectations, needs, and perceptions, allowing for a comprehensive assessment of the prototype tool.

2 Employee absenteeism data

In this section, we provide an overview of the system MojeUre [2] on which the entire

analytical prototype is built for different analysis scenarios, such as predictive and descriptive analysis.

In our case, we use a relational database with 25 tables. In a relational database [3] data is organised into tables with rows and columns, and primary and secondary keys are used to define relationships between tables and to ensure the integrity of the data. Predictive and descriptive analysis.

We are outlining the data sources and the underlying system architecture, with a focus on key tables such as CheckIn, Employee, Company, and EmployeeVacation. Particular attention is given to the subset of data related to employee presence and absence, which forms the core of this research. The dataset is characterized by strong demographic diversity, spanning various regions, industries, and employment types, thereby enabling meaningful cross-sectional analyses.

For predictive tasks [4], such as absence forecasting, a sliding-window methodology was applied to capture historical context and temporal dynamics. These datasets included more than 140 attributes, covering demographic information, lagged absence profiles, seasonal and holiday indicators, as well as region- and company-specific features.

For descriptive tasks, such as clustering [5] and anomaly detection [6], a uniform dataset was prepared for the year 2019. It contains 365 binary daily presence/absence attributes per employee, enriched with aggregate and demographic features. Clustering was employed to identify distinct absence behaviour patterns, while anomaly detection was used to highlight unusual or unexpected attendance behaviours through multiple unsupervised algorithms.

3 Evaluation design and User feedback

3.1 Evaluation Process

The evaluation was carried out through interviews with key personnel, such as HR managers and company directors, with a particular focus on employee-related data. The primary aim of the evaluation was to understand how employee absences affect workflows and to identify opportunities for improving work organisation. To achieve this, open-ended questions were carefully prepared to encourage detailed responses, for example: how current systems support organisational processes, and what difficulties sudden absences create for management.

The interviews followed a structured process:

- Selection of participants: representatives from different roles and company sizes were chosen to ensure diverse perspectives.
- Interview conduct: each interview began with a short introduction and an explanation of confidentiality, followed by a structured but flexible dialogue.
- Documentation: responses were recorded and transcribed to accurately capture emerging themes.
- Analysis: data were analysed to identify common issues, gaps, and potential solutions.

- Reporting: findings were summarised into key challenges and actionable recommendations, which were then shared with participants to demonstrate how their input influenced the research.

In this doctoral study, interviews were conducted with representatives from four companies, covering the year 2023. The organisations varied in size, from small enterprises (0–10 employees) to medium-sized ones (10–50 employees), and included different organisational structures, such as department heads and directors. This diversity enabled a wide range of perspectives on absence-related challenges. The interviews were conducted in Slovenian, with one taking place online and three in person, depending on the interviewee's preference.

Nine main questions, supplemented by sub-questions, were prepared to explore specific aspects in more depth. Graphical representations of the analysis results were also presented to the participants, who provided feedback on how they interpreted these outputs. Their reflections offered valuable insights into how the software prototype could be further refined. The discussion focused on the evaluation of outputs generated by the prototype, including visualisations such as graphs. The collected feedback was carefully documented, highlighting both common themes across organisations and unique challenges faced by individual companies.

The findings provided actionable insights for improving absence management and informed recommendations for further development of the prototype tool.

3.2 Interview Structure

The interview process was conducted in a clear and organised manner to ensure consistent and meaningful data collection. It began with an introduction, during which the purpose of the interview was explained, and confidentiality assurances were provided to the participants. The central part of the process involved asking open-ended questions, allowing participants to elaborate on their experiences, challenges, and needs related to absence management. The graphs generated from the analysis were also presented, and participants provided feedback on these visualisations. The graphs presented to participants illustrated weekly absence patterns (stacked bar charts), comparisons of actual versus predicted absences for selected weeks (grouped bar charts), and long-term absence trends over time (line charts). We aimed to show how effectively the visualisations highlighted peaks, deviations, and patterns in absences, as well as the alignment between actual and predicted values. Participants were asked to provide feedback on the clarity, interpretability, and usefulness of these visualisations for supporting absence management decisions. Active listening was employed throughout, with follow-up questions to dive deeper into the responses. The interview concluded by summarising the key points and inviting any additional feedback. This structured approach ensured that all relevant topics were addressed while enabling participants to share their insights freely.

3.2.1 Question design

The list of questions was developed to address how companies manage work organisation and predict absences, especially in cases of sudden or frequent employee absences. It was designed in line with the research goals, focusing on the main challenges organisations face in scheduling, planning work, and managing staff availability. The aim was to collect

feedback on current systems and expectations for analytical tools that could enhance absence prediction and support better workforce management.

In preparing the questions, we identified key topics such as the use and effectiveness of existing attendance systems, the need for improvements, and experiences with work reorganisation during absences. The questions were presented in a table to provide a clear overview and allow for easier comparison and analysis, helping to identify potential improvements and guide the development of more effective predictive solutions.

3.2.1 Execution of the interview

The interview process followed a clear structure:

The interviews followed a structured process:

- Introduction - Participants were informed about the purpose of the interview and reassured regarding confidentiality.
- Opening Questions - The discussion started with broad, open-ended questions to encourage participants to share their experiences, challenges, and needs around absence management.
- Interactive Segment - After around seven questions, the session became more interactive. Participants were shown visual materials based on their own data analysis and asked to provide feedback.
- Active Listening - Follow-up questions were used throughout to explore responses in greater depth and ensure clarity.
- Conclusion - The interview ended with a summary of the main points and an invitation for any final remarks.

This approach allowed all key areas to be covered while still giving participants the space to contribute their perspectives openly. Interview questions are presented in Table 1.

Table 1: List of questions for interview

No.	Question
1	How does your current time-tracking system help with work organisation?
2	What expectations do you have for a developed analytical tool for absence prediction?
3	How does absence prediction impact work organisation in your company?
4	Does a sudden employee absence affect your work organisation? Do you know of any experiences from others?
5	What kind of information would be useful for you when predicting absences, and why?
6	How is a specific job position reorganised in the event of a sudden absence? What problems, challenges, or obstacles arise?
7	How accurate is the current absence prediction? Do the current data suffice for easier decision-making?
8	What are your expectations regarding absence prediction with an analytical tool? What potential problems do you foresee?
9	Are you familiar with any existing solutions that offer similar functionalities for predicting absences?

4 Results

4.1 Company A

Company A, located in the Posavska region, operates with up to 10 employees and develops innovative web applications and advanced business systems. Their solutions help companies streamline operations and improve digital infrastructure.

The manager explained that their time-tracking system efficiently records arrivals, departures, and absences (e.g., vacations, sick leave). However, the absence of key staff can delay projects, especially during critical phases. They would welcome an analytical tool for predicting absences to improve work organisation and prevent staff shortages. Currently, short-term absences are managed through direct communication, depending on task urgency.

From the presented charts they found the visuals clear and useful for interpreting absence patterns. Absences are typically higher during summer, which aligns with model predictions. They noted that the model effectively captures general trends but struggles with sudden, unpredictable absences.

In Figure 1 bellow we show the weekly distribution for the year 2023 of predicted absence durations (0–5 days) in case of a one-week-ahead prediction for vacation leave for the whole company. During most of the year, the model predicts 0-day absences, while a more diverse range of predicted absence lengths appears in the summer weeks due to increased vacation leaves.

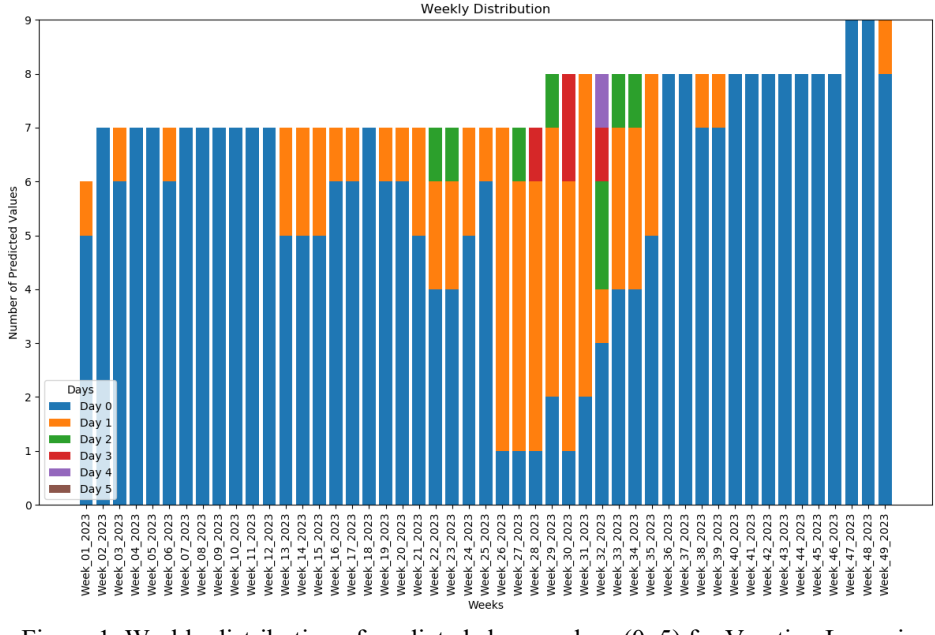

Figure 1: Weekly distribution of predicted absence days (0–5) for Vacation Leave in Company A, with the Y-axis showing the number of absent employees and colours on

the X-axis indicating absence days (e.g., blue = 0 days, red = 3 days).

Company A interviewees found the first graph clear, as it effectively showed absences by weeks. The second graph was also understandable, presenting absences for a specific week. The third graph was less clear due to deviations between predicted and actual values.

4.2 Company B

Company B, with around 300 employees and several branches across Slovenia, sells agricultural tools, equipment, and its own produce such as fruits and wines. The company operates in multiple shifts and maintains a strong presence in Slovenia's agricultural sector.

The HR representative explained that their time-tracking system covers basic attendance and absence monitoring, while scheduling is still done manually. A predictive tool for forecasting absences would help identify unreliable staff and support planning during critical periods like harvests. Although unexpected absences are usually manageable thanks to flexible staff coverage, information on absence duration and type would improve planning. The company sees potential in absence prediction but would first test its practical performance. They are not aware of similar existing solutions.

Figure 2 illustrates the weekly distribution of predicted absences across days in 2023 for the entire company, revealing a consistent concentration on Day 0 and Day 1. This suggests that the model effectively identifies immediate short-term vacation leave absences. The distribution remains stable throughout the observed period, with fewer predictions for longer absences.

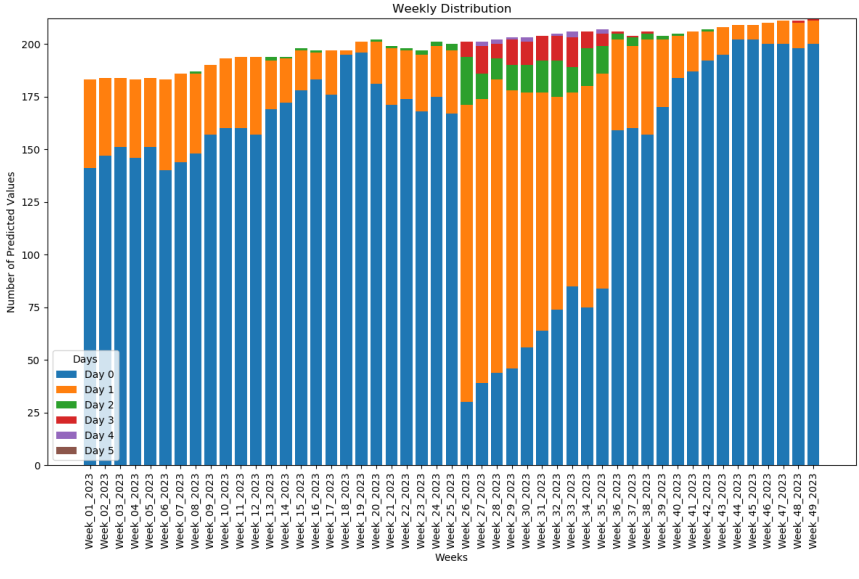

Figure 2: Weekly distribution of predicted absence days (0–5) for Vacation Leave in Company B, with the Y-axis showing the number of absent employees and colours on the X-axis indicating absence days (e.g., blue = 0 days, red = 3 days).

The interviewees from Company B found the first graph clear, while the second graph was considered less useful due to inaccurate predictions. The third graph was also perceived as unclear, and overall, the visualisations were not deemed particularly useful because the predictions lacked sufficient accuracy.

4.3 Company C

Company C, a healthcare institution in the Spodnjeposavska region with around 50 employees, provides high-quality medical services in two shifts from Monday to Friday. It plays an important role in the local community and uses modern technology to support patient care.

The HR representative explained that they use separate tools for time tracking and shift scheduling. Sudden absences, especially of key medical staff, cause operational disruptions and delays in patient care. They would welcome an integrated system that combines scheduling, absence tracking, and prediction to better plan replacements and reduce last-minute issues. Finding substitutes is often difficult and increases administrative work. The company is interested in a user-friendly, accurate prediction tool to improve workforce management, though they are not aware of similar existing solutions.

Figure 3 shows the weekly distribution of predicted vacation leave days for company C in year 2023 for whole company. Most predictions are concentrated around 0 and 1 days, indicating a high share of employees expected to be present or absent for only one day. A visible drop in predictions occurs during mid-year weeks, which corresponds to the summer holiday period.

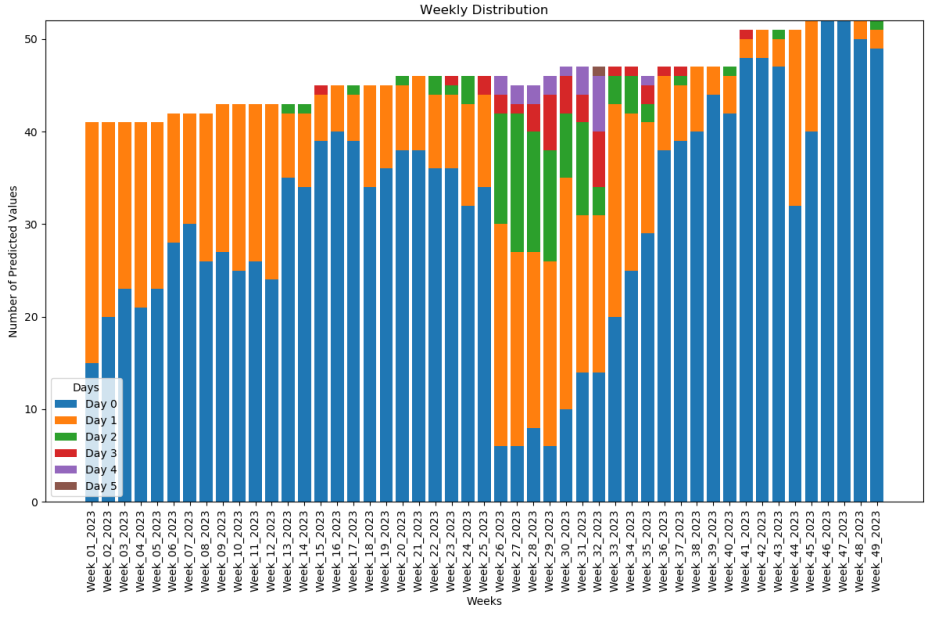

Figure 3: Weekly distribution of predicted absence days (0–5) for Vacation Leave in

Company C, with the Y-axis showing the number of absent employees and colours on the X-axis indicating absence days (e.g., blue = 0 days, red = 3 days).

The interviewees from Company C reported that they had no difficulties interpreting any of the graphs, as all visualisations were clear and easy to follow. However, they emphasised that the accuracy of the predictions was not sufficient, which limited the practical usefulness of the presented results.

4.4 Company D

Company D, an educational institution in the Osrednjeslovenska region with around 50 employees, operates mainly in morning shifts from Monday to Friday. Staff work varied schedules depending on their roles, supporting the institution's educational activities.

The manager explained that their time-tracking system efficiently records attendance, working hours, and payroll. They are interested in an analytical tool for predicting absences to improve planning and ensure sufficient staffing, especially in childcare, where absences cause major disruptions. Frequent absences and staff shortages make coverage difficult, often requiring reassignments across locations. They seek a more accurate prediction tool, as current systems are either imprecise or too costly and lack forecasting capabilities. Expectations are high, as such a tool could significantly improve work organisation.

Figure 4 presents the weekly distribution of predicted vacation leave days for employees in company D throughout the year 2023 for the whole company. The data is categorised from 0 to 5 leave days per employee. A visible peak of multi-day absences occurs during the summer months, while shorter absences dominate in other periods.

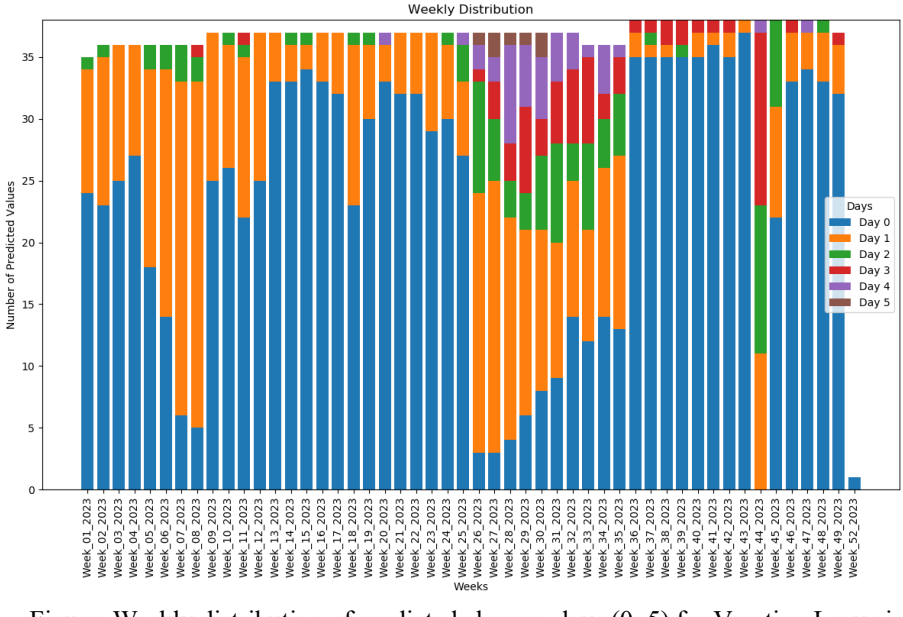

Figure: Weekly distribution of predicted absence days (0–5) for Vacation Leave in Company D, with the Y-axis showing the number of absent employees and colours on

the X-axis indicating absence days (e.g., blue = 0 days, red = 3 days).

The interviewees from Company D stated that all graphs were clear and understandable, as they could easily interpret the meaning of each visualisation. They found the second graph, which presents weekly absence predictions, to be the most useful since it allows them to plan for specific predicted absences. Nevertheless, they noted that the accuracy of predictions could be improved to increase the practical value of the visualisations.

5 Conclusion

The first company, A, located in the Posavska region, employs up to ten people and specialises in developing web applications and advanced business systems. While their time-tracking system is effective, unforeseen absences during critical project phases cause delays. They seek an analytical tool for absence prediction to improve task reorganisation, though expectations are modest, as current reorganisation relies on direct communication. Their feedback on the graphs showed that the first and second visualisations were clear and understandable, while the third was less valuable due to inaccurate predictions.

The second company, B, with up to 300 employees, sells agricultural tools and products across Slovenia. Their time-tracking system supports attendance monitoring, but manual scheduling remains key for shift organisation. A predictive absence tool would aid planning during critical periods, such as harvests, though current systems generally suffice. They are open to exploring predictive solutions. Regarding the graphs, they found the first one clear, while the second and third were less useful and lacked the accuracy needed for practical application.

The third company, C, is a healthcare institution in the Posavska region with 50 employees that operates on a two-shift schedule. Sudden key staff absences disrupt operations, highlighting a need for an integrated predictive tool to improve workforce management, reduce delays, and maintain patient care. They are open to exploring predictive solutions. They had no problems understanding the graphs, but noted that the prediction accuracy was insufficient, limiting their usefulness.

The last company, D, is an educational institution in the Central Slovenian (Osrednjeslovenska) region, with up to 50 staff. They face challenges from frequent absences, particularly in childcare. Their time-tracking system manages attendance but lacks predictive forecasting. Previous attempts with costly tools like PlanDela were insufficient. They seek a precise, affordable absence prediction tool for better planning and reduced disruptions. Their feedback indicated that all graphs were clear, and they found the second graph most useful for weekly absence predictions. However, they also pointed out a need for higher prediction accuracy.

The findings show that all four companies have effective attendance systems but seek better tools for predicting absences to improve planning and reduce disruptions. Smaller firms focus on quick reorganisation during key projects, while larger ones prioritise coverage of critical roles and reducing administrative work. The educational institution, facing frequent absences, needs a more accurate and affordable solution. Overall, all companies express interest in reliable, user-friendly analytical tools for better workforce planning. While the graphs were clear and easy to understand, their practical value was limited by low prediction accuracy.

6 References

[1] Zupančič, P., Klisara, J., & Panov, P. (2023). Razvoj orodja za napovedovanje odsotnosti zaposlenih: Analiza potreb uporabnikov. Journal of Universal Excellence (JUE)/Revija za Univerzalno Odličnost (RUO), 12(3).

[2] 1A Internet. (2023). Evidenca delovnega časa MojeUre. Retrieved from Aplikacija MojeUre: https://mojeure.si

[3] Harrington, J. L. (2016). Relational database design and implementation. Morgan Kaufmann.

[4] Zupančič, P., & Panov, P. (2024). Predicting employee absence from historical absence profiles with machine learning. Applied Sciences, 14(16), 7037.

[5] Zupančič, P., & Panov, P. (2024). Clustering of employee absence data: A case study. 14th International Conference on Information Technologies and Information Society (ITIS 2023). Conference proceedings (pp. 155–164). Faculty of Information Studies.

[6] Zupančič, P., & Panov, P. (2024). Anomaly detection in time-series employee absence data: A case study. Proceedings of the 47th MIPRO ICT and Electronics Convention (MIPRO) (pp. 1099–1104). IEEE.

Sustainable Practices and Digital Innovation in Serbia and Slovenia: Comparative Analysis and Economic Implications

Milica Stanković[1], Gordana Mrdak[1], Jovana Džoljić[1], Stevan Simić[2]
[1]Academy of Applied Technical and Preschool Studies
Filipa Filipovića 20, 17000 Vranje, Serbia
[2]UFC Holding
milica.stankovic, gordana.mrdak,
jovana.dzoljic@akademijanis.edu.rs
stevansimic@live.com

Abstract: *A modern approach to sustainable development integrates the circular economy with digital innovation to use resources efficiently, reduce the environmental footprint and create new opportunities for green jobs and a competitive economy. The aim of the paper is to reach conclusions about the challenges faced by the respondents and what digital innovations and in what way they can contribute to overcoming the challenges through a comparative analysis of the path of Serbia and Slovenia towards the achievement of sustainable development goals and the results of primary research on the sustainable habits of respondents in these two countries. Primary research shows that despite the differences in sustainable habits in the compared countries, the obstacles to a sustainable life are very similar in both Serbia and Slovenia. After mapping the barriers to sustainable living in Slovenia and Serbia, the paper presents proposed technological solutions in the form of digital innovations, as well as potential benefits that can be realized by implementing them. Based on the summary of potential benefits, it can be concluded that digital technologies can not only improve sustainable practices but can significantly contribute to improving the competitiveness of the economy, reducing the environmental footprint, and creating new economic opportunities through the development of green businesses.*

Key words*: Sustainable Development, Circular Economy, Digital Innovation, Digital Transformation, Sustainable Development Goals (SDGs)*

1 Introduction

A modern approach to sustainable development integrates the circular economy with digital innovation to use resources efficiently, reduce the environmental footprint and create new opportunities for green jobs and a competitive economy. The aim of the paper is to reach conclusions about the challenges faced by the respondents and what digital innovations and in what way they can contribute to overcoming the challenges through a comparative analysis of the path of Serbia and Slovenia towards the achievement of sustainable development goals and the results of primary research on the sustainable habits of respondents in these two countries.

The contribution of this paper is multiple. Primarily, the paper presents a comparative analysis of national progress in achieving sustainable development goals, individual sustainable habits and perceived obstacles in Serbia and Slovenia. Also, the paper shows

the connection between mapped obstacles to sustainable living identified by primary research and digital innovations that can contribute to overcoming them. The first part of the paper provides a theoretical basis and summarizes previous research in the field of sustainable development, circular economy, digital transformation and innovation and creates an initial connection between these concepts necessary for further analysis and research. The second part of the paper provides a comparative overview of the situation in Serbia and Slovenia in terms of progress towards the achievement of 17 sustainable development goals based on the Sustainable Development Report. The third part of the paper presents the results of the primary research that was conducted in February 2025 and includes a total of 85 respondents, of which 60 respondents are from Serbia and 25 respondents are from Slovenia. In the fourth part of the paper, based on the mapped challenges for the implementation of sustainable habits among respondents, digital innovations that can contribute to overcoming challenges are pointed out and their potential benefits are listed. Based on comprehensive primary and secondary research, relevant conclusions are presented.

2 Theoretical Background

A comprehensive approach to economic and ecological sustainability implies a transition from a linear to a circular economy with the creation of a sustainable society that rationally uses natural resources and reduces the ecological footprint [1]. In that process, STEM professionals have a special role, bearing in mind that they are expected to develop digital innovations that will support the principles of sustainability [2]. It is the transition to the circular economy that opens opportunities for new green jobs, especially in the areas of energy efficiency, sustainable technologies and renewable energy sources, which ultimately contributes to strengthening the competitiveness of the economy [3]. The synergy of digital tools and sustainable practices contributes to the optimization of resource consumption and waste reduction and accelerates the transition to sustainable growth models [4]. Digital transformation is one of the key drivers of sustainability that enables more efficient management of resources and significant savings, whereby digital technologies (AI, big data, blockchain, cloud computing, IoT) play a key role in encouraging innovation and sustainability [5]. Especially after the COVID-19 pandemic, digitalization has accelerated the adaptation processes of organizations, enabling more efficient decision-making, optimization of resources, and the creation of new business models [6] [7].

The Coalition for Digital Environmental Sustainability (CODES) emphasizes the simultaneous need to align digital strategies with sustainable development goals, actively mitigate the negative impacts of digital technologies, and implement digital innovations explicitly focused on sustainability outcomes. This framework provides a conceptual basis for public policies and digital strategies, which can be implemented very successfully in countries such as Serbia and Slovenia [8]. The connection of digital technologies and sustainability led to concepts such as digital sustainable entrepreneurship and sustainable digital innovation, which denote a long-term process of introducing digital solutions with the aim of creating economic value, but also reducing environmental and social negative consequences [9] [10]. Digitization and innovation are proving to be powerful drivers of sustainable performance, as they enable organizations to develop new products and services that are both profitable and environmentally responsible [11].

3 Methodology

This research uses a combined approach that includes a comparative analysis of secondary data and a primary survey of respondents in Serbia and Slovenia. The analysis and synthesis of secondary data from the Sustainable Development Report 2025 provided a systematic overview of Serbia's and Slovenia's progress towards the 17 Sustainable Development Goals (SDGs). Secondary data analysis includes comparative visualization and synthesis of key indicators to identify areas of success and challenges in achieving sustainable development, economic performance and digital transformation.

The primary research was conducted in February 2025 on a sample of a total of 85 respondents - 60 from Serbia and 25 from Slovenia. A survey methodology was used with structured questionnaire consisted of four sections: demographic data, sustainable habits practiced in everyday life, barriers to sustainable living and awareness and perception of digital innovations that could support sustainable behavior. Most questions were closed-ended and structured using single or multiple choice formats. The data were analyzed descriptively and through a comparative approach to identify differences and similarities between respondents from the two countries.

A special focus of the methodology is connecting the identified challenges with the potential of digital innovations. Based on the results of primary and secondary research, the key challenges experienced by respondents in the implementation of sustainable practices were mapped. For each challenge, digital innovations are proposed that can contribute to overcoming it, emphasizing the expected benefits.

The combination of analysis and synthesis of secondary data, results of primary research and practical recommendations for digital innovations enables a comprehensive approach to the research question and contributes to the understanding of how different economic, environmental and technological factors affect the sustainable habits of citizens in Serbia and Slovenia.

4 Results and Discussion

In Serbia, awareness of sustainable development is growing, but there is still a gap between environmental awareness and the adoption of sustainable practices in everyday life. To bridge the gap, it is important to emphasize education about sustainable development and the circular economy. Serbia is actively working to achieve the Sustainable Development Goals (SDGs) listed in the Agenda for Sustainable Development until 2030. The country's performance is monitored annually, and the latest data from the Sustainable Development Report for 2025 show that Serbia is in 33rd place out of 167 countries, achieving an SDG index score of 78.2 [12]

While some goals have been largely achieved or steady progress is being made, there are still significant challenges for certain SDGs. The goal of SDG 1 (eradication of poverty) is the only one that has been fully achieved, bearing in mind that Serbia records extremely low rates of extreme poverty, with almost universal access to basic services. Health care indicators indicate a low rate of maternal, neonatal and child mortality, with solid universal health coverage. However, there are still challenges such as high prevalence of obesity, mortality from chronic diseases, traffic accidents and lower life expectancy compared to EU averages. Literacy and primary education coverage are at a high level, although there is a drop in scores for SDG 4 (quality education). In terms of water and sanitation services, Serbia records a high availability of basic services, but a low percentage of wastewater treatment. The energy sector shows full availability of electricity and a solid share of renewable sources, but CO_2 emissions are still a challenge.

In the field of infrastructure and digitization, the country achieves a high level of use of the Internet and mobile technologies. However, investments in research and development (1% of GDP) and the number of registered patents are still below the average of developed countries. In the area of responsible consumption and production, a low level of management of electronic and municipal waste is recorded (Figure 1.) [12]

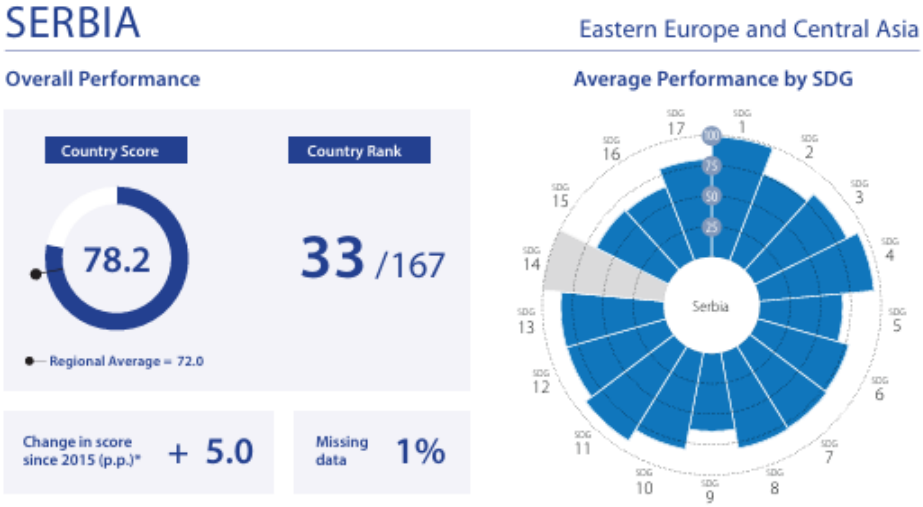

Figure 1.: Serbia overall and average performance and SDG index score [12]

Cities are the key drivers of the Serbian economy, bearing in mind that in the 28 largest cities, about 60% of the total population of Serbia lives and about 74% of employment is realized in them. Nevertheless, demographic trends indicate a general decline in the number of inhabitants in Serbia, with the simultaneous expansion of urban space, which further increases the negative effects on the environment. The main challenges in cities remain air pollution, poor waste management and low energy efficiency, and they further slowdown the green transition. The solution to the mentioned problems can be the framework of 3C: Concentrate, Connect, Capacitate. "Concentrate" implies a focus on regional centers and adapted policies in accordance with the demographic and economic trends of each city. "Connect" emphasizes a better connection between national policies and local actions, as well as coordination between different levels of government and different sectors. "Capacitate" refers to strengthening the technical, institutional and financial capacities of local governments to implement sustainable policies and attract investments. [13]. Nevertheless, it is important to emphasize that agriculture remains one of the key sectors for sustainable living in the Republic of Serbia, and it is important to pay attention to initiatives in villages that support sustainable agriculture [14].

Slovenia achieves a relatively high level of success in implementing the Sustainable Development Goals (SDGs), with an average result above the regional average and a stable overall trend. According to the 2025 report, Slovenia ranks quite high on the global SDG index — it ranks 12th out of 167 countries, with an SDG index score of 81.2. It shows that the country generally ranks well in terms of progress towards the Sustainable Development Goals (SDSN, 2025).

SDG 1 (eradication of poverty) has been fully achieved, while SDG 11 (sustainable cities and communities) shows a stable maintenance trend. The country has almost universal

access to basic services – from electricity and clean water to sanitation and education – with a low poverty rate and high health standards. However, there is still room for improvement in the field of wastewater treatment and the use of renewable energy sources. Slovenia also stands out in terms of universal health insurance, low infant and maternal mortality, as well as long life expectancy, and records a high subjective well-being of its citizens. In education, coverage and quality are at a high level, and in terms of gender equality, a significant balance can be observed in education and the participation of women in the labor market. In addition, the country has a strong position in research and innovation, with high levels of investment in development and widespread digital connectivity, high internet usage and good academic results. However, there is still room for more investment in research and development and the participation of women in STEM fields. Inequalities in society are low, but there is a gender gap in earnings and somewhat greater poverty among the elderly population. Air quality is acceptable, and access to public transport is good. Municipal waste management is effective, but there are challenges in electronic waste management (Figure 2.) [12]

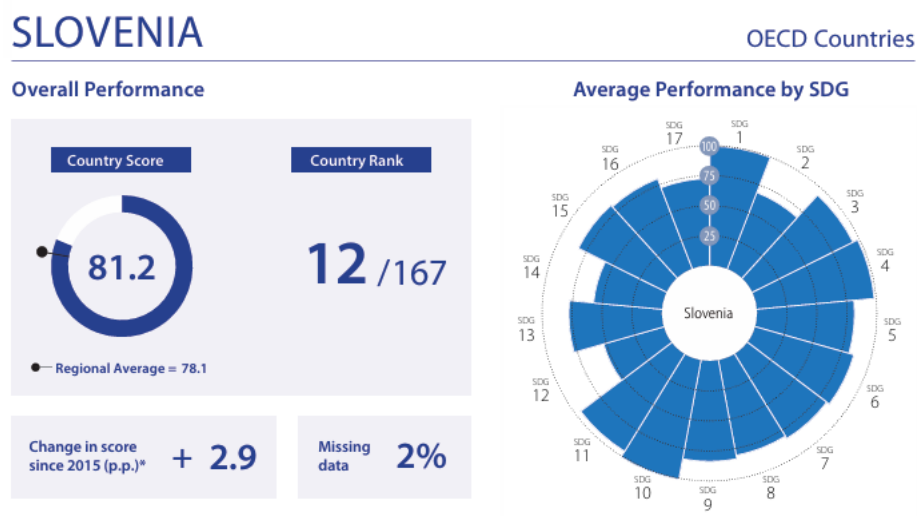

Figure 2.: Slovenia overall and average performance and SDG index score [12]

The authors of the paper conducted primary research during the month of February 2025 in Serbia and Slovenia on a sample of a total of 85 respondents between the ages of 18 and 65. 60 respondents from Serbia and 25 respondents from Slovenia participated in the research. In the sample of respondents from Serbia, 75% are women, while 25% are men, while in Slovenia, 60% of men and 40% of women answered the questionnaire. The research results show that as many as 83.3% of respondents in Serbia believe that sustainable habits and practices are very important or important for the future of the planet, while in Slovenia that percentage is as high as 96%.

When it comes to sustainable practices that they apply in everyday life, in both Serbia and Slovenia, by far the largest percentage of respondents emphasize recycling (66.7% in Serbia, 96% in Slovenia). In addition, in Serbia, a third of respondents emphasize that they consume energy sustainably and the same percentage of respondents opt for organic food. A quarter of respondents buy local products, and only 8.3% opt for reduced use of

plastic, use of public transport, cycling or walking and reduced water consumption (Figure 3.).

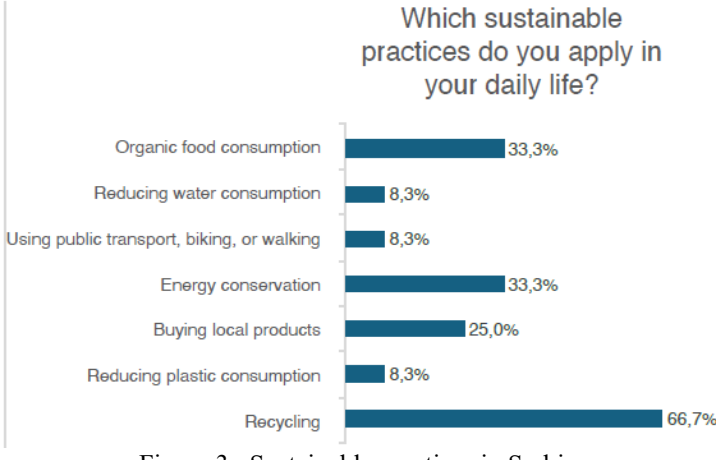

Figure 3.: Sustainable practices in Serbia

It is an interesting fact that respondents from Slovenia have slightly different trends, and as many as 72% of respondents point out that they focus on reducing the use of plastic. Slightly less than half of the respondents opt for energy conservation (48%), buying local products (44%), organic food consumption (44%), using public transport, biking and walking (44%) and reducing water consumption (40%) (Figure 4.). Through a comparative analysis, we can conclude that respondents in Slovenia practice sustainable habits in a balanced way, while in Serbia certain sustainable practices are significantly less common (eg reducing water consumption, reducing plastic consumption, using public transport, biking or walking). So, in Slovenia, sustainable practices are more integrated into everyday life, while in Serbia they are still based only on recycling.

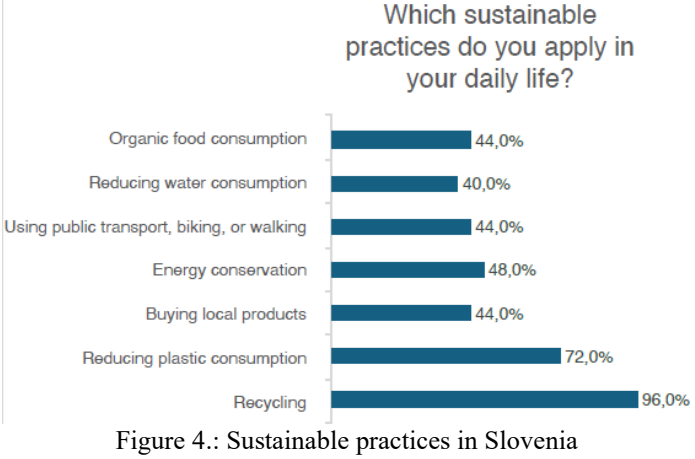

Figure 4.: Sustainable practices in Slovenia

The following are cited as key barriers to sustainable living in Serbia: high costs of sustainable products (41.7%), lack of infrastructure (recycling centers, public transport, etc.) (41.7%) and insufficient community support (41.7%). A third of the respondents

point to a lack of time, while a quarter cites a lack of information as an obstacle to sustainable living (Figure 5.).

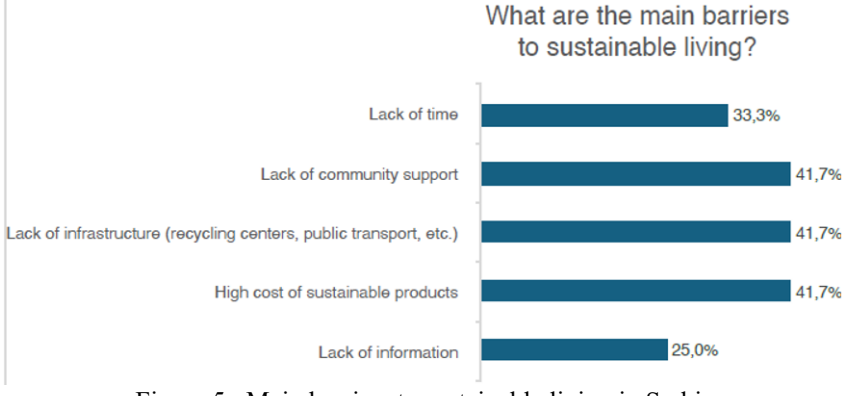

Figure 5.: Main barriers to sustainable living in Serbia

Respondents in Slovenia believe that the biggest obstacle to a sustainable life is the lack of infrastructure (84%), while an equal number of respondents point out the lack of information and the high costs of sustainable products (60%). About a third of the respondents from Slovenia point out the lack of community support as an obstacle, while a fifth of the respondents cite a lack of time (Figure 6.).

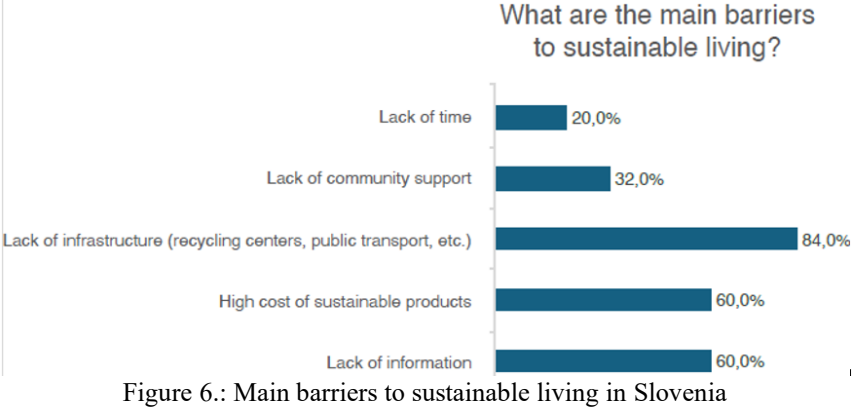

Figure 6.: Main barriers to sustainable living in Slovenia

Recognizing the importance of a sustainable way of life, there is a logical need to overcome the mapped barriers. In this, digital innovations, i.e. solutions based on the use of modern digital technologies, can play an important role. Table 1 shows the mapped barriers for sustainable living in Serbia and Slovenia, as well as digital innovations that can contribute to overcoming each of the mentioned barriers and potential benefits for the economy, ecology and society. Therefore, digital innovations not only provide technical solutions for specific challenges, but also contribute to reducing the ecological footprint, strengthening social connectivity and creating synergy between technological progress and sustainable development goals.

Table 1.: Digital innovation as a solution to identified barriers to sustainable living

Identified barriers	Digital innovations	Potential benefits
High costs of sustainable products **Serbia 41,7%** **Slovenia 60%**	Platforms for exchange and sharing of resources Peer-to-peer applications Digital markets for second-hand products Apps for smart shopping	Reducing costs and waste Increased availability of sustainable products
Lack of infrastructure **Serbia 41,7%** **Slovenia 84%**	Mobile applications with maps of recycling centers and e-vehicle chargers Smart bins AI to optimize public transport	Easier access to existing resources Greater use of public transport Higher recycling rate
Lack of information **Serbia 25%** **Slovenia 60%**	Educational platforms and e-learning applications Eco Footprint Apps Chatbot assistants	Increasing knowledge and motivation Personalized recommendations for sustainable practices
Lack of community support **Serbia 41,7%** **Slovenia 32%**	Digital communities and social networks for sustainability (local groups, digital forums for sharing ideas and resources, digital campaigns, volunteering platforms)	Strengthening social cohesion and joint action Support for local initiatives
Lack of time **Serbia 33,3%** **Slovenia 20%**	Applications for sustainable purchasing planning Applications for optimization of energy, water and waste consumption Digital reminders AI assistants for organization: planners that integrate environmental choices (e.g. shortest green route, recycling reminders)	Automating and facilitating sustainable decisions in everyday life Saving time

5 Conclusion

On the way to a sustainable society, a transition from a linear to a circular economy is necessary with the integration of sustainable practices in the daily life of individuals in parallel with the integration of digital strategies and policies of sustainable development. A key role in this process is played by educators, decision makers, representatives of public institutions, but also society and the community. The results of our research indicate that it is precisely the connection of sustainable practices and digital innovations that is key to the achievement of sustainable development.

Comparative analysis based on secondary data from the Sustainable Development Report 2025 gives us a clear insight that Slovenia is much closer to achieving the SDGs compared to Serbia, but that strong institutional support and education can contribute to both countries improving their SDG index score in the coming period. By comparing sustainable practices in the observed countries, we conclude that in Slovenia the situation is significantly better when it comes to recycling and reduced use of plastic, and that sustainable habits are more comprehensively implemented in the daily life of the

respondents. On the other hand, in Serbia, respondents cited recycling as the dominant sustainable practice, while other sustainable habits are still in their infancy. Primary research shows that despite the differences in sustainable habits in the compared countries, the obstacles to a sustainable life are very similar in both Serbia and Slovenia. The most common challenges faced by respondents in both observed countries are high costs of sustainable products, lack of infrastructure and information, as well as limited community support.

After mapping the barriers to sustainable living in Slovenia and Serbia, the paper presents proposed technological solutions in the form of digital innovations, as well as potential benefits that can be realized by implementing them. Based on the summary of potential benefits, it can be concluded that digital technologies can not only improve sustainable practices but can significantly contribute to improving the competitiveness of the economy, reducing the environmental footprint, and creating new economic opportunities through the development of green businesses.

6 References

[1] Stanković, M., & Džoljić, J. (2022). Green entrepreneurship: a path towards green economy. 37th International Scientific Conference – Knowledge for development, Ohrid, North Macedonia, 18–21 August 2022.

[2] Stanković, M., Džoljić, J., & Dimitrijević, B. (2022). Circular economy and STEM professionals competencies. 36th International Scientific Conference – The teacher of the future, Budva, Montenegro, 03–05 June 2022.

[3] Stanković, M., Anđelković, T., Mrdak, G., & Stojković, S. (2023). Transition to green economy and green jobs. 27th International Eco-Conference, Novi Sad, Serbia, 27–29 September 2023.

[4] Stanković, M., Mrdak, G., & Džoljić, J. (2024). Big Data and the Circular Economy: Synergy for Sustainable Growth. 6th International Conference Path to a Knowledge Society – Managing Risks and Innovation, October 20–21.

[5] Schork, S., Özdemir-Kaluk, D., & Zerey, C. (2025). Understanding innovation and sustainability in digital organizations: A mixed-method approach. Sustainability, 17(2), 415.

[6] Ardito, L. (2023). The influence of firm digitalization on sustainable innovation performance and the moderating role of corporate sustainability practices: An empirical investigation. Business Strategy and the Environment, 32, 5252–5272.

[7] Lu, H. T., Li, X., & Yuen, K. F. (2023). Digital transformation as an enabler of sustainability innovation and performance—Information processing and innovation ambidexterity perspectives. Technological Forecasting and Social Change, 196, 122860.

[8] Coalition for Digital Environmental Sustainability (CODES). (2022). Action Plan for a Sustainable Planet in the Digital Age: Supplement 1 — Accelerating Sustainability Through Digital Transformation: Use Cases and Innovations. Coalition for Digital Environmental Sustainability.

[9] Nasiri, M., Saunila, M., Rantala, T., & Ukko, J. (2022). Sustainable innovation among small businesses: The role of digital orientation, the external environment, and company characteristics. Sustainable Development, 30(4), 703–712.

[10] Yousaf, Z., Radulescu, M., Sinisi, C. I., Serbanescu, L., & Păunescu, L. M. (2021). Towards sustainable digital innovation of SMEs from the developing countries in the context of the digital economy and frugal environment. Sustainability, 13(10), 5715.

[11] Xu, J., Yu, Y., Zhang, M., & Zhang, J. Z. (2023). Impacts of digital transformation

on eco-innovation and sustainable performance: Evidence from Chinese manufacturing companies. Journal of Cleaner Production, 393, 136278.
[12] Sustainable Development Solutions Network (SDSN). (2025). Sustainable Development Report 2025. https://sdgtransformationcenter.org/reports/sustainable-development-report-2025
[13] World Bank. (2022, January 13). Green, Livable, Resilient Cities in Serbia Program. https://www.worldbank.org/en/country/serbia/brief/green-livable-resilient-cities-in-serbia-program
[14] United Nations Serbia. (2024, April 18). United Nations Serbia 2023 Results Report. https://serbia.un.org/en/266266-united-nations-serbia-2023-results-report

From Tool to Co-Author: Legal and Ethical Thresholds of Authorship in the Age of Generative AI

Matevž Mandl, dr. Martina Plantak
DOBA University of Applied Sciences
Prešernova 1, 2000 Maribor, Slovenia
matevz.mandl@doba.si, martina.plantak@doba.si

Abstract: *The rapid progression of generative AI challenges traditional concepts of authorship and copyright. While current frameworks require a human creative input as a threshold for protection, the precise boundaries of "sufficient human contribution" remain unclear and contested. This paper examines how different jurisdictions address AI-assisted works. Building on comparative legal analysis, this paper introduces a normative and operational framework consisting of five dimensions (conception and intention, control and guidance, transformative editing, responsibility, and traceability). Each dimension is mapped onto existing legal standards and ethical requirements, including transparency and recognition principles reflected in existing guidelines and standards. The framework can be of value in a vast number of institutions when evaluating works. We argue that linking legal thresholds with ethical criteria of recognition and accountability is crucial for sustainable creative ecosystems in the age of AI.*

Key Words: *Ethics, Intellectual Property, Artificial Intelligence, Human Authorship, Ethics of Recognition and Responsibility.*

1 Introduction

Generative artificial intelligence (short: GenAI) has blurred the line between tool and co-author, as its contributions are increasingly more intricate and advanced. Texts, images, audiovisual works produced with GenAI assistance raise pressing questions for copyright law: can such outputs/results be protected, and under what conditions?

Most legal systems require human intellectual creativity as a necessary precondition for authorship. However, as GenAI systems are trained on extensive corpora of copyright-protected works and can produce outputs that approximate human expression and originality, the contours of this requirement have become increasingly indeterminate [1]. This paper addresses this normative and practical uncertainty by combining an overview of existing legal practices with ethical inquiry, and by proposing a framework to evaluate the sufficiency of human authorship in AI-assisted creations.

2 Methods

This paper uses a qualitative methodology that combines doctrinal legal analysis with normative ethical analysis. The first step in developing an operational framework involves a close reading od statutory provisions and case law across four different

jurisdictions (EU, USA, UK and China). This identifies how each legal system defines authorship and whether non-human agency can be awarded authorship. The comparison enables highlighting the convergences and divergences in different jurisdictions.

Normative ethical analysis rooted in theories of recognition, responsibility and distributive justice used in editorial policies and international guidelines, such as UNESCO's AI Ethics Recommendation or the OECD AI Principles, are used to frame ethical obligations beyond strict legal threshold. Both emphasize human oversight, accountability, and complete transparency, offering a normative bridge from mere compliance to ethically robust practice [2, 3]. This ensures that the framework incorporates both compliance and value driven considerations.

Finally, insights from analyses are synthesized into a conceptual model, that can be applied in judicial, editorial, and institutional contexts. This methodology ensures that the findings are translated into a usable framework for practice.

4 Legal Frameworks on Human Authorship

Berne Convention for the Protection of Literary and Artistic Works (as amended on September 28, 1979) [4] is the international foundation for modern copyright law. It establishes the extension of protection to "authors of literary and artistic works", which presupposes a human creator. Article 2(1) defines the scope of protection, while Article 3(1) limits authors to "nationals" of contracting countries. This means that the Convention does not recognize non-human entities as authors in all its member states.

EU law is thus harmonized through directives with the Articles of Berne Convention. Although the EU directive never explicitly defines "author" as a natural person, the Court of Justice of the European Union (CJEU) for example made this assumption in Case C-5/08 – Infopaq International A/S v Danske Dagblades Forening (2009) [5], where it clearly stated that copyright can be applied only in relation to a subject-matter which is original in the sense that it is its author's own intellectual creation [5]. This presupposes an individual capable of intellectual and creative choice, therefore a human being. Further developing this approach is the Directive 2001/29/EC on the Information Society (InfoSoc Directive) that consistently refers to authors within human context [6].

The copyright law in US consistently requires human authorship as it presumes that original works of authorship should be fixed in any tangible medium of expression [7]. This interpretation was later further affirmed by the United States Court of Appeals in Thaler v. Perlmutter [8] where court held that copyright can protect only human-made work and generated work by AI is not registrable.

UK copyright law adopts a more flexible approach to authorship as such and takes a unique position by expressly addressing computer-generated works in the Copyright, Designs and Patents Act (1988), where Section 9 provides that in case of computer-generated works, the authorship shall be taken to the "person by whom the arrangements necessary for the creation of the work are undertaken" [9]. This provides a unique situation but still demands a creative choice and significant human involvement, as creative decision-making remains essential. However, even under this provision, UK courts, guidelines and commentators emphasize that originality still depends on a creative choice made by a human author, for example, through selection, arrangement, or

curatorial decisions. In other words, creative decision-making must remain and remains the essential element of authorship, ensuring that the resulting work reflects human intellectual creation rather than autonomous AI computation [10, 11]

Article 11 of Copyright Law of People's Republic of China (2020 Revision, in force 2021) expressly states the authorship can be attributed only to the citizen who has created the work [12], which presumes a human being as an author. Furthermore, when a work is created under the direction and responsibility of a legal entity or organization, authorship may be attributed to that entity or organization, reflecting the principle of institutional authorship in Chinese copyright law.

All mentioned major jurisdictions and international frameworks converge on the principle that authorship is predominantly human. The Berne Convention provides the normative baseline by referring exclusively to authors presupposing a human. This principle permeates EU law through mentioned directives and CJEU rulings. The USA reinforces the same standard through law and practice. The UK is the only example that formally acknowledges computer-generated works but still requires sufficient human creative input. China also confines authorship to humans and demands evidence of human intellectual input.

The similarity is substantive: all systems deny copyright to fully autonomous AI outputs, confirming that creative authorship, and the legal and ethical responsibilities it entails, remains an exclusively human domain.

5 Ethical Consideration

Ethics plays a vital role in this debate because it explains not only *who* can be an author under current law but also *why* human oversight, judgment, and accountability must remain central as generative AI becomes more deeply embedded in creative practices. Ethical frameworks, such as those promoted by the Committee on Publication Ethics (COPE), establish that authors must take responsibility for the integrity of their work, a requirement that AI tools cannot fulfil. Hence, AI systems cannot be listed as authors but must instead be transparently disclosed as aids. [13]

Recognizing human intellectual labour is crucial for sustaining trust in academic, cultural, and creative ecosystems. Failure to do so risks devaluing human expertise and shifting normative authority toward opaque technological systems. As Cath et al. argue, insufficient oversight in AI governance may allow institutional or private actors to shape standards without adequate ethical reflection, thereby prioritizing convenience over responsibility. [14]

Maintaining ethical standards also mitigates risks associated with GenAI, including amplification of biases, misuse of data, and the absence of clear accountability structures. Upholding transparent and human-centered authorship practices therefore serves both moral and practical purposes. [15, 16]

6 Operative Framework

Given the legal and ethical implications, we propose an operative framework that assess whether AI-assisted works meet the threshold of human authorship. It is based on five basic dimensions that jointly capture legal requirements and ethical imperatives. It can be used everywhere where there is a need to evaluate authorship to works involving GenAI.

This framework evaluates whether a work contains sufficient human authorship across five dimensions:

1. Conception and intention – Did the human author define an original expressive concept, beyond generic prompts?

Many editorial committees have extensive authorship criteria. For instance, ICMJE – International Committee of Medical Journal Editors notes that an individual must make substantial contributions to conception or design, participate or critically revise the work, give final approval, and agree to be accountable for all aspects of the work in order to be listed as an author [13].

2. Control and guidance – Did the author exercise iterative, creative decision-making in directing the AI system?

Maintaining human oversight and intentionality throughout algorithmic creation is prerequisite for accountability and moral agency in creative production [16].

3. Transformative editing – Did the author substantially modify or curate the AI output in a creative way?

Human author must have sufficient control, input and contribution to the work, throughout the prompt engineering and creating of the work, based on person's selection or arrangement of elements [17].

4. Responsibility – Is there a clearly identifiable human who assumes responsibility for the work?

Authorship of a work entails accountability for all aspects of a work. AI systems cannot satisfy this requirement by definition, hence cannot be assumed as an author [13].

5. Traceability – Are there verifiable records of prompts, edits, or provenance?

Recent emerging policies ask authors to document prompting techniques, the tool's role in the study or final work, and other choices, so that editors, reviewers can verify who contributed what and how [18].

Each dimension can be scored 0-2. 0 meaning absent, 1 = minimal and 2 = substantial contribution. A cumulative score of $\geq 6/10$ indicates sufficient human contribution for authorship. With scores 5 or below work can be treated as AI-generated and thus unlikely to qualify for copyright protection. A score of 6-8 would suggest that work is borderline but generally should qualify as AI-assisted authorship. Score of 9 or 10 would represent a strong case of human authorship despite AI involvement and therefore definitely qualify for copyright protection.

A score-based application (e.g., minimum 6/10 points) can guide courts, publishers, and cultural institutions in distinguishing AI-assisted works from AI-generated works lacking human authorship.

The five-dimensional operative framework, encompassing conception and intention, control and guidance, transformative editing, responsibility, and traceability, is grounded in legal standards and judicial practices in the EU, US, UK, and China. Each dimension reflects criteria already recognized across these jurisdictions. For example, the requirement of a human-defined conception and intention behind AI-assisted works aligns with the EU's originality test of the "author's own intellectual creation," as established by the CJEU in Infopaq (Case C-5/08) [5], and with US jurisprudence, particularly Thaler v. Perlmutter, which affirmed that only human-made works are registrable for copyright [8]. The emphasis on control and guidance mirrors the UK's approach under Section 9 of the Copyright, Designs and Patents Act 1988, where authorship of computer-generated works is attributed to the person making the necessary arrangements [9], provided human creative choices remain essential [10, 11]. Similarly, the transformative editing dimension corresponds to the widespread requirement for a human intellectual contribution, such as curatorial or editorial input, to meet originality standards [6, 11]. The responsibility criterion reflects legal expectations that an author must be capable of bearing accountability, as illustrated in China's Copyright Law (2020 revision), which limits authorship to natural persons or, under certain conditions, legal entities [12]. Finally, traceability resonates with emerging standards in publishing ethics and institutional practice, which increasingly require transparent documentation of prompts, edits, and the AI's role such as those proposed by COPE and other ethical guidelines [13, 18].

Taken together, the framework operationalizes converging legal trends, namely the requirement for human creative input, the capacity for responsibility, and the need for process transparency, and translates them into a structured, evaluative model. In doing so, it bridges doctrinal standards with institutional application. The model respects current judicial interpretations, avoiding protection for autonomous AI-generated outputs (as rejected in Infopaq [5], Thaler [8], and Chinese law [12]), while introducing normative clarity and usability. By articulating five concrete factors, it offers a unified tool that enhances consistency and accountability in practice without departing from existing legal doctrine.

7 Discussion

In addition to bringing clarity, the proposed framework contributes to a broader understanding of how law and ethics can jointly sustain creative accountability in the age of GenAI. By operationalizing legal principles such as originality and authorship through measurable indicators of human input and creative process, the model provides a bridge between normative theory and institutional practice.

It further encourages transparency and traceability as new forms of evidence supporting both judicial interpretation and ethical compliance in creative industries. Furthermore, this framework invites policymakers and regulatory bodies to consider using standardized protocols as a means of preserving trust and responsibility within human and AI collaboration. In this sense the framework is not merely a diagnostic tool but also a policy-oriented instrument for fostering sustainable, human-oriented innovation. The simple implementation of traceability mechanisms, as the retention and citation of GenAI conversation logs and prompt records, can serve as practical means of evidencing human involvement. Transparent documentation of how prompts were designed, revised and which parts of the finished product were affected can serve as a verifiable proof of

creative agency, complementing tests of originality and reinforcing trust within creative and academic ecosystems [19].

We believe that this framework offers clarity where existing frameworks diverge. For the EU, it operationalizes AOIC; for the US, it provides concrete guidance on human authorship disclosures; for the UK, it suggests a reinterpretation of Section 9 in light of modern systems; for China, it harmonizes emerging case law.

Still, we recognize that the proposed model has important limitations. It possibly simplifies complex human–AI creative processes into a scoring system, which may not capture all aspects of authorship. The evaluation relies on subjective judgments, so different evaluators could reach different conclusions. The model also assumes that records of prompts and edits are available and verifiable, which may not always be the case. Finally, because generative AI is rapidly evolving, the framework reflects current conditions and will need ongoing refinement and testing to remain relevant.

8 Conclusion

The rise of generative AI necessitates a principled yet practical approach to authorship. While jurisdictions vary, all converge on the idea that human intellectual input is indispensable. Proposed framework bridges legal and ethical perspectives, offering a tool for consistent evaluation of AI-assisted works. By embedding recognition, responsibility, and transparency, the framework contributes to building a sustainable, human-centric future of creativity with AI.

The framework still requires empirical validation through case analysis and stakeholder testing.

9 References

[1] Authors Guild. Authors Guild v. Open AI, No. 1:23-cv-8292, https://www.classaction.org/media/authors-guild-et-al-v-openai-inc-et-al.pdf, downloaded: September 22nd 2025.
[2] UNESCO. Recommendation on the Ethics of Artificial Intelligence (2021).
[3] OECD. OECD AI Principles (adopted 2019; updated 2024).
[4] World Intellectual Property Organization. Berne Convention for the Protection of Literary and Artistic Works, Paris Act of July 24, 1971.
[5] Court of Justice of the European Union. Infopaq International A/S v. Danske Dagblades Forening, Case C-5/08, Judgment of 16 July 2009.
[6] European Union. Directive 2001/29/EC on the harmonisation of certain aspects of copyright and related rights in the information society, Official Journal L 167, 22 June 2001.
[7] United States. Copyright Act, Title 17 U.S. Code, Section 102(a)
[8] United States Court od Appeals. Thaler v. Perlmutter, No 23-5233, Judgment of March 18th 2025, United States District Court for the District of Columbia, March 18th 2025.
[9] United Kingdom. Copyright, Designs and Patents Act 1988, Section 9.
[10] UK Intellectual Property Office. Consultation on Artificial Intelligence and Intellectual Property: Copyright and Patents. London, 2022.
[11] UKIPO. AI and IP: A Summary of Responses and Government Response. 2023.

[12] China. Copyright Law of the People's Republic of China (2020 Revision).
[13] Committee on Publication Ethics. COPE Position Statement: Authorship and AI Tools. Available at: https://publicationethics.org/guidance/cope-position/authorship-and-ai-tools.
[14] Cath, C. , Wachter, S. , Mittelstadt, B., Taddeo, M. i Floridi, L. (2018). Artificial Intelligence and the 'Good Society': the US, EU, and UK approach. Science and Engineering Ethics 24, Springer.
[15] Borenstein, J., Howard, A. (2021). Emerging challenges in AI and the need for AI ethics education. AI Ethics 1, pp. 61–65.
[16] Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P. & Vayena, E. (2018). AI4People — An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. Minds & Machines 28, 689–707.
[17] Lee, E. "Prompting Progress: Authorship in the Age of AI." Florida Law Review (2024).
[18] Luo, X., Tham, Y. C., Giuffre, M., et al. Reporting guideline for the use of Generative Artificial Intelligence tools in Medical Research (the GAMER statement). BMJ Evidence-Based Medicine 2025.
[19] Hosseini, M., Resnik, D. B., Holmes, K. The Ethics of Disclosing the Use of AI Tools in Writing Scholarly Manuscripts. Accountability in Research (2023).

Exploring the Adoption of Generative AI in Higher Education - A Cross-National Comparison

Nadia Molek, Lejla Imamović Lerić, Annmarie Gorenc Zoran
Faculty of organization studies in Novo mesto
Ulica Talcev 3, 8000 Novo mesto, Slovenia
nadia.molek@fos-unm.si, Lejla Imamović Lerić
lejla.imamovic.leric@fos-unm.si; a.zoran@fos-unm.si

Nerman Ljevo
Faculty of Management and Business Economics
Azapovići 439, 71250 Kiseljak, Bosnia and Herzegovina,
nerman.ljevo@fmpe.edu.ba

Abstract: *Generative Artificial Intelligence (GenAI) is rapidly transforming higher education. This study examines its adoption in Bosnia and Herzegovina (BiH) and Slovenia (SI) using a cross-sectional, cross-national mixed-methods design, combining quantitative survey data with qualitative content analysis of open-ended responses. Findings show moderate familiarity in both contexts, with higher frequency of use among BiH students. GenAI is mainly employed for research tasks in BiH, while in SI it supports personalized learning and teaching. Perceptions of effectiveness are positive but accompanied by ethical concerns, especially around plagiarism, reliability, and integrity. Support for regulation differs: BiH respondents favour strict controls, whereas SI participants prefer flexible pedagogical approaches. The study highlights the role of cultural traditions, institutional infrastructures, and regulatory contexts in shaping adoption, and calls for universities to strengthen digital literacy, critical awareness, and context-sensitive guidelines for responsible integration of GenAI in higher education.*

Key Words*: Generative AI; higher education; digitalization; ethics; Bosnia and Herzegovina; Slovenia.*

1 Introduction

Digital technologies have long reshaped education as part of broader processes of human development and social reproduction. Following the transformative impact of the Internet on access to knowledge (Alier et al., 2024), Generative Artificial Intelligence (GenAI) is rapidly becoming integral to teaching, learning, and research. Surveys indicate that 50–65% of students and faculty have already experimented with tools like ChatGPT, suggesting GenAI is on track to become fully embedded in higher education (Baytaş & Ruediger, 2024).

Educational uses of GenAI involve three dimensions: the data it draws on, the algorithmic processing of queries, and users' evaluation of outputs. Data shape accuracy and bias, algorithms shape relevance, and user judgment shapes value and legitimacy (Passey et al., 2024). As Molek (2023) notes, while AI is already embedded in higher education, its future trajectory depends on its users.

Research on GenAI in higher education clusters into three strands: (a) empirical studies of faculty and student use (Hwang & Chen, 2023; Xia et al., 2024); (b) analyses of benefits and risks—from productivity and personalization to misinformation, plagiarism, and dependency (Ogunleye et al., 2024; Lim et al., 2023); and (c) attitudinal work balancing enthusiasm for innovation with ethical-governance concerns (Kaplan-Rakowski et al., 2023; Chan & Hu, 2023). Yet comparative, cross-national analyses remain scarce. We still know little about how distinct academic traditions, infrastructures, and regulatory regimes shape GenAI appropriation—especially in Europe's semi-periphery (e.g., the Balkans), where digital transformation intersects with heterogeneous institutional trajectories. Against this backdrop, our study examines knowledge, use, perceptions, and ethics of GenAI among faculty and students in BiH and Slovenia, using a mixed-methods cross-national design. It explores similarities and differences in adoption and addresses five questions: familiarity and use, purposes, perceived effectiveness, ethical concerns, and support for regulation.

2 Literature review

The incorporation of emerging technologies in higher education has been widely studied. Scholars emphasize that universities are not neutral environments for technological adoption, as integration depends on infrastructures, pedagogical traditions, and cultural attitudes toward innovation (Selwyn, 2022; Williamson & Eynon, 2020). This perspective moves the debate beyond purely technical accounts and highlights how practices of use, regulation, and resistance reflect broader academic cultures.

GenAI can be mobilized by different actors within educational contexts. As Hwang and Chen (2023) note, professors employ it for generating examples, summarizing content, or preparing teaching materials; students apply it to tasks such as homework, problem-solving, or creative outputs; peers use it collaboratively in team-based work; and AI itself functions as a partner in data collection and analysis. This blurring of boundaries between teacher, student, and tool introduces hybrid forms of collaboration. Yet, as Noroozi et al. (2024) remind us, meaningful adoption requires ethical and pedagogical framing, considering implications for assessment, learning goals, and academic integrity.

In this line, several studies underscore tensions between human and machine agency. Wong and Looi (2024) stress the need for rational integration, while Okaiyeto et al. (2023) point to the paradox of supporting GenAI despite its dependence on responsible human use. Xia et al. (2024) argue that teacher training, AI literacy, and revised assessment policies are preconditions for meaningful implementation. At the institutional level, McDonald et al. (2024) find that 73% of universities in their sample encourage faculty use of GenAI, with more than half including it in curricula, though often accompanied by ethical warnings.

The literature converges on the double-edged nature of GenAI adoption. On the one hand, GenAI enhances research and teaching practices, facilitating literature reviews, personalized learning, feedback, and efficiency (Rulfi & Spada, 2023; Ogunleye et al., 2024; Guchhait, 2024). On the other, it introduces risks such as inaccuracies (Lim et al., 2023), erosion of trust between professors and students (Luo, 2024), and potential limitations on critical thinking (Abunaseer, 2023).

The legitimacy of Generative AI (GenAI) in higher education is not only grounded in its technical affordances but also in the narratives through which students and professors frame its usefulness. Rather than focusing solely on outputs, adoption is mediated by symbolic expectations of productivity, creativity, and informational access. As Aoun

(2017) argues in his vision of "robot-proof" higher education, the challenge for universities is to prepare students for futures in which automation coexists with human adaptability, critical thinking, and creativity. Similarly, Knox et al. (2020) highlight the role of AI imaginaries—collective beliefs about what AI can and should do—in shaping pedagogical practices and legitimizing its integration.

In this perspective, GenAI is perceived as both a pragmatic and a symbolic ally. It is imagined as a time-saving tool, an assistant for personalized learning, and a catalyst for new forms of creativity. For example, Baidoo-Anu and Anash (2023) describe it as an "endless sea of knowledge" that stimulates curiosity and self-directed learning, while Pavlik and Pavlik (2024) show how it enhances visualization of creative processes, especially in artistic fields. These imaginaries resonate with empirical findings: in Slovenia, respondents emphasized GenAI's role in providing tailored pedagogical support, whereas in Bosnia and Herzegovina they highlighted its capacity to expand access to literature and resources.

At the same time, perceptions of effectiveness are contested. Critical voices caution against over-reliance, warning that GenAI may generate inaccuracies (Lim et al., 2023), undermine trust in student–professor relationships (Luo, 2024), or constrain intellectual autonomy (Abunaseer, 2023). Lee (2024) further questions its pedagogical value, noting that students tend to follow AI-generated suggestions uncritically. These tensions illustrate that GenAI's legitimacy is socially negotiated: while widely embraced as a productive and creative resource, it simultaneously raises doubts about the depth and independence of academic learning.

The integration of GenAI in higher education inevitably raises ethical and governance challenges, particularly around issues of academic integrity, plagiarism, data privacy, and reliability of outputs, showing that its integration has cultural meanings, institutional practices, and governance frameworks. These debates highlight both opportunities—such as productivity and creativity—and challenges related to ethics, critical thinking, and trust, thereby framing the research questions that guide this study.Globally, these debates are structured by emerging frameworks of AI ethics, such as Floridi and Cowls' (2019) principles of beneficence, non-maleficence, autonomy, justice, and explicability, which have influenced both institutional codes and regulatory agendas. In Europe, the discussion is increasingly framed through the EU Artificial Intelligence Act, which sets out a risk-based governance model. While the Act is not yet fully implemented, its projected impact on higher education is already shaping institutional guidelines and expectations. Prior research supports the relevance of ethical reflexivity in education. As Molek (2023) argues in her critical review of digitalization and ethics in organizational and educational contexts, the introduction of AI tools requires not only technical literacy but also ethical literacy, where students and faculty learn to critically evaluate the boundaries of acceptable use.

In this sense, our study investigates how professors and students in Bosnia and Herzegovina and Slovenia—two contexts with different institutional histories, digital infrastructures, and regulatory environments—engage with GenAI in their academic practices. Specifically, the research addresses five guiding questions: (1) What is the level of familiarity and frequency of use of GenAI? (2) What are the primary purposes of its use? (3) How is its effectiveness and productivity perceived? (4) What ethical concerns emerge, and how do they vary across contexts? and (5) To what extent is there support for stricter regulations and institutional guidelines? These questions are operationalized into five working hypotheses (H1–H5), which structure the empirical analysis presented in the following section..

3 Method

This cross-sectional, cross-national exploratory mixed-methods study examined knowledge, use, purposes, perceptions, and ethical orientations toward GenAI among higher education actors in Bosnia and Herzegovina (BiH) and Slovenia (SI). Data were collected in October 2024 through an online questionnaire distributed via institutional and professional channels. The instrument consisted of five sections—demographics, familiarity and use, purposes, perceptions, and ethical concerns—with both closed and open-ended items. Closed items employed 5-point Likert-type scales (ranging from 1 = strongly disagree/never to 5 = strongly agree/very frequently). Versions in Bosnian and Slovenian were semantically aligned for comparability. The instrument was developed for this study and reviewed for face validity by the research team; no formal psychometric validation was performed, consistent with the exploratory scope of the study.

The analytic sample comprised 29 respondents in BiH and 20 in SI. The BiH group was younger and mainly student-based, whereas the SI group was older and more balanced across academic roles, reflecting the exploratory character of cross-national research.

Quantitative data were summarized using descriptive statistics (frequencies, percentages, means). Between-country differences were explored with χ^2 or Fisher's exact tests for categorical variables and Mann–Whitney U tests for ordinal measures, with effect sizes reported alongside p-values. Given small sample sizes, emphasis was placed on descriptive contrasts rather than statistical inference.

Open-ended responses were examined through qualitative content analysis. Two members of the research team independently coded the data, discussed discrepancies, and reached consensus to enhance interpretive reliability. Codes were grouped inductively into broader themes, with attention to both convergence and divergence across cases.

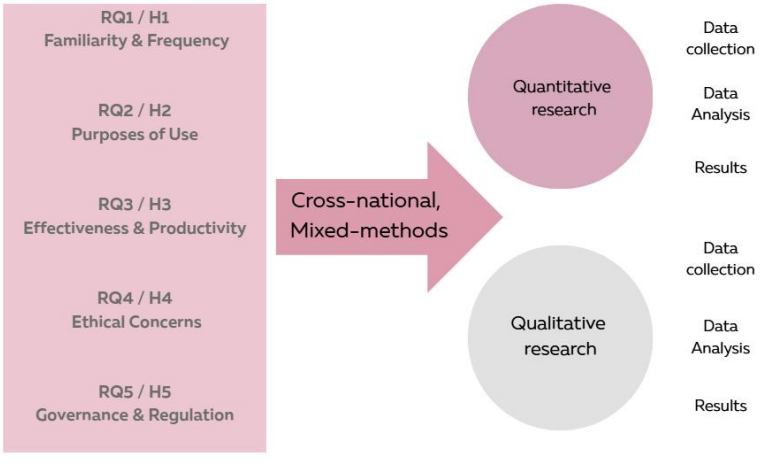

Figure 1: Research model

The study complied with institutional standards and GDPR principles. Participation was voluntary and based on informed consent obtained at the start of the questionnaire. Respondents were informed about the purpose of the study, the voluntary nature of participation, and their right to withdraw at any point. Data were collected anonymously, stored securely on password-protected servers, and retained only for the duration

necessary for analysis. Given the minimal risk and the absence of personally identifiable information, no further institutional ethical approval was deemed necessary.

4 Results

4.1 Demographic Profile of Respondents

The demographic composition of the two national samples reflects differences in age structure, gender distribution, and academic affiliation (Table 1 summarizes the demographic profile of respondents in both national contexts).

In BiH**,** the largest share of respondents were aged 25–34 (37,9%), followed by those aged 35–44 (34,5%). Younger participants aged 18–24 represented 20,7%, while older cohorts were less represented (6,9% aged 45–54, and none above 55). In terms of gender, the sample was predominantly female (58,6%), while men accounted for 37,9%. No respondents identified as non-binary. Regarding institutional roles, the BiH's sample was composed mainly of undergraduate students (48,3%), followed by faculty members (24,1%). Master's students (13,8%) and postdoctoral researchers (10,3%) were less represented, with very few indicating "other" (3,4%).

In Slovenia, the age profile was markedly older. The majority of respondents were aged 35–44 (35%) and 45–54 (40%), with smaller shares in younger cohorts (5% aged 18–24; 15% aged 25–34; and 5% aged 55–65). No participants were above 65. The gender distribution revealed an even stronger predominance of women (75%), with men making up 25%. Regarding academic role, the Slovenian sample was more balanced: undergraduate students (35%) were the largest group, followed by doctoral students (30%), faculty members (20%), and master's students (15%).

4.2 Knowledge and frequency of Use of GenAI (RQ1, H1)

There can be noted contrasts emerged between the two national samples (see Table 1). In Bosnia and Herzegovina (BiH), most respondents reported being somewhat familiar with GenAI (62,1%), while a smaller proportion considered themselves very familiar (27,6%), and 10.3% reported no familiarity at all. In contrast, the Slovenian sample demonstrated a markedly higher level of reported familiarity: 85% of respondents declared themselves very familiar, and the remaining 15% reported being somewhat familiar. Importantly, none of the Slovenian respondents indicated unfamiliarity with GenAI.

Table 1: Familiarity with Generative Artificial Intelligence by national contexts

Variable	**Category**	**Bosnia (%)**	**Slovenia (%)**
Familiarity with GenAI	Not known	10,3	0,0
	Somewhat known	62,1	15,0
	Very known	27,6	85,0

Source: Authors' survey data (2024).

Findings show GenAI is more embedded in Slovenian higher education than in BiH. While institutional embedding of GenAI appeared stronger in Slovenia, self-reported awareness of ethical issues and policy frameworks (such as the EU AI Act) was higher in

Bosnia and Herzegovina. This divergence may reflect differences in sample composition (predominantly students in BiH vs. more balanced roles in SI), as well as contextual factors shaping exposure to GenAI in practice versus awareness at the discursive level.

In Bosnia and Herzegovina, GenAI use was more dispersed, with about one quarter never using it, while in Slovenia usage was more regular, with most respondents reporting daily or weekly engagement. The lower mean score in Slovenia (M = 2.65 vs. BiH M = 3.45) confirms a higher overall intensity of use.

Qualitative insights further contextualize these results. In Slovenia, students emphasized the novelty of encountering GenAI in the classroom, which illustrates a peer-to-peer diffusion dynamic, where familiarity expands informally once the tool is introduced institutionally. Several students also highlighted efficiency gains: *"AI shortened the time needed to search for information because it quickly provides summaries,"* while others warned about over-reliance: *"Sometimes I became too dependent on AI and researched less on my own." In BiH,* participants more frequently stressed the exploratory nature of adoption and unevenness in exposure. Some pointed to GenAI's role in translation or accessing literature, while others reflected on its risks: *"AI sometimes gave me wrong information, which was confusing,"* or *"I noticed that students sometimes rely only on AI and skip the books."*

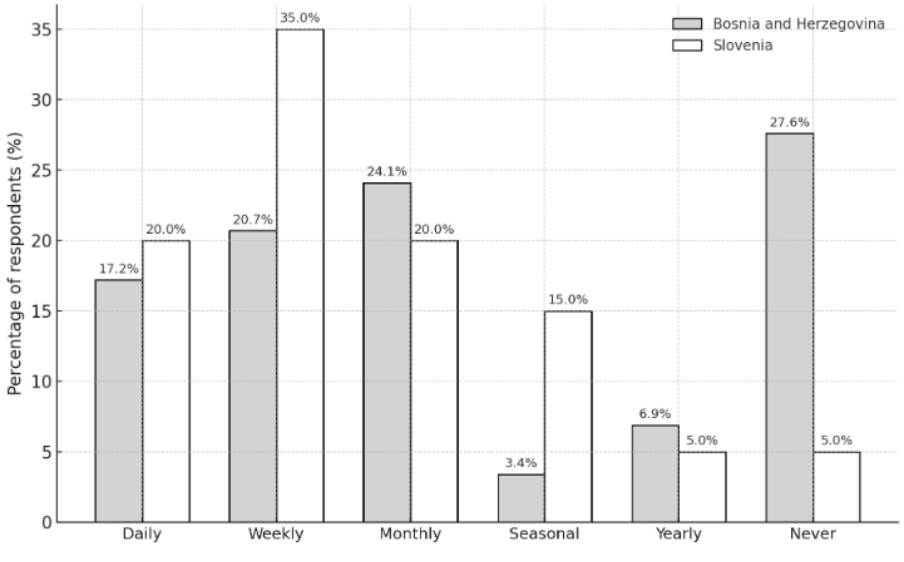

Figure 2: Frequency of Generative AI use among respondents in Bosnia and Herzegovina and Slovenia
Source: Authors' survey data (2024).

These findings point toward a descriptive "usage gap" between the two countries. In Slovenia, GenAI seems more embedded into academic practices, while in Bosnia and Herzegovina, a substantial minority of respondents have yet to engage with such tools. However, while these descriptive contrasts are notable, the difference in frequency of use between the two samples did not reach statistical significance.

4.3 Purposes of Use (RQ2, H2)

Respondents in both countries reported diverse purposes for using Generative AI in higher education, although the relative importance of these purposes varied significantly between the two national contexts (see Figure 3).

In Slovenia, the most frequently cited purpose was assistance in preparing teaching materials (50% of respondents), followed by creative activities (25%), communication support (10%), and solving learning problems (10%). Very few respondents indicated the use of GenAI for writing scientific papers (5%). The mean score (M = 2.75) suggests that Slovenian respondents position their GenAI use primarily toward practical, teaching-related applications, with relatively lower emphasis on research or student learning support. Qualitative comments illustrate this orientation: *"Easier to find definitions of unfamiliar terms," "Faster article scanning/reading, new ideas for problem solving, text polishing, translations."* These narratives confirm that Slovenian users see GenAI as an efficiency-enhancing tool for academic writing, classroom preparation, and linguistic support.

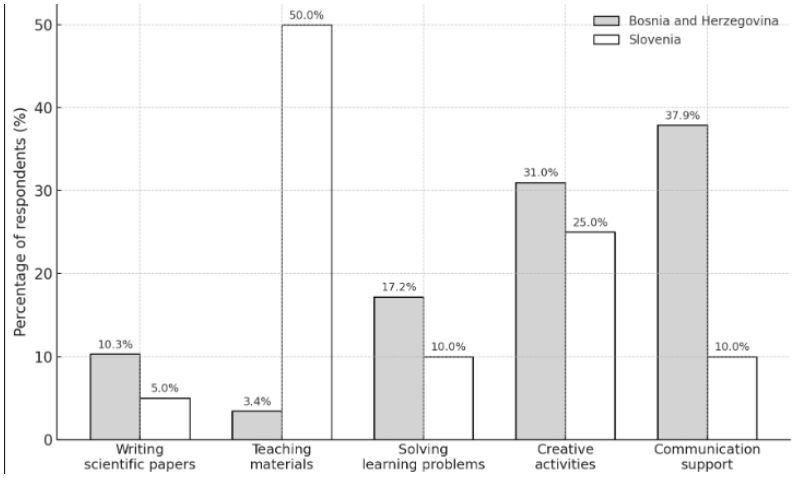

Figure 3: Purposes of GenAI use among respondents in Bosnia and Herzegovina and Slovenia. Source: Authors' survey data (2024).

In BiH, the pattern was markedly different. The most common purposes were communication support (37,9%) and creative activities (31,0%), followed by solving learning problems (17,2%). Only a small share reported using GenAI for writing scientific papers (10,3%) or assistance in preparing teaching materials (3,4%). The higher mean score (M = 4.69) reflects that GenAI is used more for student-centered and interactive purposes—particularly communication, creativity, and problem solving—rather than for faculty-oriented tasks. Respondents described how GenAI served as a conversational partner or brainstorming aid: *"AI helps me rephrase sentences and improve the tone in communication," "It gave me creative ideas for presentations and class projects,"* and *"Sometimes I used it to check or simplify explanations when I didn't understand something in class."* These accounts underscore that in BiH, GenAI is appropriated less through institutionalized teaching routines and more through individualized, exploratory uses that support student learning and expression.

4.4. Perceptions of Effectiveness, Advantages, and Challenges of GenAI (RQ3, H3)

Across both national contexts, respondents generally perceived GenAI as an effective and productivity-enhancing tool in higher education. In Bosnia and Herzegovina, 41,4% rated GenAI as *very effective* and 31,0% as *somewhat effective*, while in Slovenia, half of respondents (50%) considered it *moderately effective* and 35% *very effective*. Mean values (Bosnia M = 2.03; Slovenia M = 1.95, lower values = higher effectiveness) confirm that perceptions of effectiveness were broadly similar, with no statistically significant differences (H3 rejected, p = .938). Respondents in both countries also tended to agree that GenAI increases productivity. In Bosnia, 37,9% completely agreed and 24,1% agreed; in Slovenia, 50% agreed and another 20% completely agreed. Mean scores were nearly identical (Bosnia M = 2.48; Slovenia M = 2.45), reinforcing convergence.

However, some differences emerged in the advantages attributed to GenAI. In BiH, respondents emphasized *increased productivity (34.5%)* and *access to diverse sources and literature (31%)*, while in Slovenia the emphasis was on *productivity (40%)* and *customized learning support (30%)*. This contrast suggests that Slovenian respondents associate GenAI more with personalized learning and efficiency in teaching, whereas Bosnian respondents see it primarily as a gateway to information resources. Qualitative results support this:

- BiH: "*I use it to simplify explanations when I don't understand a concept*"; "*It helps me check the literature I cannot access directly*."
- Slovenia: "*Very helpful in understanding topics, but critical evaluation is needed*" "*Improved and shortened the time wasted searching for correct answers*", "*I now produce better work, because I must think differently*".

Table 2: Perceptions of effectiveness, productivity, advantages, challenges, and future role of GenAI among respondents in Bosnia and Herzegovina and Slovenia

Dimension	Bosnia and Herzegovina	Slovenia
Is GenAI effective?	M = 2.03 (41,4% very effective, 31% somewhat effective)	M = 1.95 (50% moderately effective, 35% very effective)
GenAI increases productivity	M = 2.48 (38% completely agree, 24% agree)	M = 2.45 (50% agree, 20% completely agree)
Main advantages	Increased productivity (34,5%), access to literature (31%)	Increased productivity (40%), customized learning support (30%)
Main challenges	Ethical issues (37,9%), reliance on technology (24%), misinformation (20,7%)	Misinformation (40%), ethical issues (35%), data privacy (15%)
Future role of GenAI	M = 2.07 (45% key tool, 45% limited due to risks)	M = 1.80 (70% additional tool, 25% key tool)
Impact on academic performance	M = 1.97 (52% small improvement, 31% strong improvement)	M = 2.10 (50% small improvement, 25% strong improvement)

Note. Mean values are based on a scale where lower values indicate higher agreement or stronger perceived effectiveness. Percentages reflect the distribution of responses within each national sample. Source: Authors' survey data (2024).

Challenges showed distinct emphases: in BiH, ethical issues (37,9%) were most cited, followed by reliance on technology (24%) and misinformation (20,7%), while in Slovenia misinformation (40%) and ethical issues (35%) predominated. Views on GenAI's future diverged: Bosnian respondents were polarized, with equal shares (44,8%) seeing it as either a key tool or one that should be limited, whereas most Slovenians (70%) considered it a supplementary aid. Perceptions of academic performance were cautious in both contexts: around half in each country reported only small improvements, with a minority noting substantial gains.

4.5. Ethical Concerns and Support for Regulation (RQ4, RQ5, H4, H5)

Respondents in both countries expressed notable concern about the ethical implications of GenAI in educational contexts, though the intensity varied. In Bosnia and Herzegovina, 41.4% reported being *very concerned*, compared to only 15% in Slovenia, where the majority (45%) described themselves as *a little concerned*. Mean values (Bosnia M = 2.45; Slovenia M = 2.35) suggest marginally higher levels of concern in Bosnia, though the difference was not statistically significant (H4 rejected, p = .779).

With regard to plagiarism and cheating, apprehensions were widespread but again stronger in Bosnia (51,7% *very concerned*) than in Slovenia (20%). Slovenian respondents distributed more evenly across *somewhat concerned* (40%) and *neither concerned nor indifferent* (40%). Similarly, Bosnian respondents expressed stronger concerns about privacy: 44,8% *completely agreed* that GenAI reveals personal information, compared to 20% in Slovenia. By contrast, 30% of Slovenian respondents selected *I don't know*, indicating less defined or less engaged positions.

When asked about awareness of ethical issues more generally, BiH reported higher familiarity: 55.2% indicated they were *very familiar*, whereas 80% of Slovenian respondents said they were *not familiar*. A similar pattern was evident with awareness of the EU AI Act. Interestingly, awareness was higher in Bosnia (27,6% well informed, 24.1% somewhat familiar) than in Slovenia (15% well informed, 40% somewhat familiar), despite Bosnia not being an EU member. This indicates that global regulatory debates are circulating transnationally, though they may resonate differently in local contexts.

Differences also emerged in attitudes toward governance. In Bosnia and Herzegovina, 44,8% supported *strict institutional guidelines*, compared to only 20% in Slovenia, where respondents favoured more *flexible regulation* (50%). When asked about broader legal regulation, 58,6% of Bosnian respondents endorsed *stricter legal control*, compared to 35% in Slovenia, where 45% considered current frameworks sufficient. These results align with H5, which hypothesized stronger support for strict regulation in one country. Although the hypothesis is partially confirmed descriptively, statistical testing indicated that differences did not reach significance. Experience with ethical challenges also diverged. In Bosnia, 27,6% reported encountering such issues *often* and another 27,6% *occasionally*, while in Slovenia the majority (55%) reported *never* experiencing them.

Qualitative responses enrich these findings, since Slovenian respondents emphasized the need for "*increasing awareness of ethical approaches to GenAI use*", "*training for teachers and students*", and "*moderate regulation, but not prohibition*". Others stressed individual responsibility: "*It is important to know you must think for yourself and not hand everything over to AI*". By contrast, Bosnian voices expressed stronger anxieties about integrity and fairness, with repeated calls for stricter institutional and governmental oversight, demonstrating higher levels of ethical concern, greater familiarity with ethical debates, and stronger support for strict institutional and legal regulation. Nevertheless,

despite these descriptive contrasts, statistical analysis did not confirm significant differences in overall levels of concern, highlighting the need to interpret the results as indicative of cultural tendencies rather than categorical divides

5 Discussion

Our findings highlight that GenAI adoption in higher education is not simply technical but mediated by academic cultures, infrastructures, and values (Selwyn, 2022; Williamson & Eynon, 2020). GenAI reconfigures tasks, roles, and ethical expectations, operating as what anthropologists might call a *total social fact* (Mauss, 1925/1990), insofar as it cuts across economic, cultural, institutional, and ethical domains simultaneously. Its adoption is shaped by imaginaries of what AI can or should do (Knox et al., 2020) and by agendas for "robot-proof" education that emphasize higher-order human skills (Aoun, 2017).

Our data align with these debates but reveal distinct national patterns. In Slovenia, GenAI is embedded in pedagogical routines, especially in preparing teaching materials and offering "personalized" support. Respondents described it as *"very helpful in understanding topics"* and as a way to *"shorten wasted time searching for answers"*, though with the caveat that critical evaluation of outputs is necessary. In Bosnia and Herzegovina, by contrast, GenAI functions more as a pragmatic resource—supporting access to literature, communication, and creativity. Participants emphasized its value for *"expanding access to sources"* and for helping students with assignments but also raised worries about overreliance and integrity. These contrasts reflect different institutional ecologies: Slovenia emphasizes AI literacy and capacity-building (McDonald et al., 2024; Xia et al., 2024), whereas BiH appropriates GenAI as an informational shortcut compensating for resource gaps.

A paradox emerges in regulatory awareness. BiH respondents report greater concern about ethics and stricter support for regulation, as well as higher familiarity with the EU AI Act, despite the country being outside the EU. Slovenian respondents, by contrast, favor flexibility and training but show lower awareness of formal frameworks. As one Slovenian respondent put it, *"moderate regulation, but not prohibition"*, while another noted, *"it is important to know you must think for yourself and not hand everything over to AI."* This suggests that ethical narratives travel beyond political borders and are re-contextualized locally (Floridi & Cowls, 2019; Molek, 2023).

Across both contexts, GenAI is legitimized as a time-saving and productivity-enhancing tool, though the meaning of "productivity" differs. Personalized support in Slovenia versus expanded access to information in BiH. Qualitative voices underscore this double-edged perception: *"I now produce better work, because I must think differently"* (Slovenia), versus *"it helps, but raises concerns about misuse"* (BiH). These distinctions highlight that GenAI performs different types of work within divergent moral economies of teaching.

We discovered, that governance regimes diverge. In BiH, strict regulation functions as a technology of trust in risk-prone environments, while in Slovenia, competence and training are viewed as safeguards. Comparative scholarship shows both trajectories internationally: legal and institutional oversight (McDonald et al., 2024; Mohmed & Elballat, 2024) alongside curriculum innovation and literacy-based approaches (Xia et al., 2024; Kadaruddin, 2023).

Finally, when revisiting the research questions and hypotheses, only H2 (purposes of use) and H5 (regulatory orientations) were supported. Divergences in purposes of use and governance align with cultural and institutional conditions, while hypotheses regarding

familiarity (H1), productivity (H3), and ethical concern (H4) were not statistically confirmed. The broader implication is that GenAI adoption cannot be reduced to measures of frequency or effectiveness: it is mediated by imaginaries, infrastructures, and governance regimes. In Slovenia, it is integrated into pedagogical innovation; in Bosnia and Herzegovina, it is approached with caution, simultaneously a resource for access and a potential ethical risk. Digitalization in higher education thus appears as a cultural process—uneven, contested, and embedded in local traditions of knowledge and authority.

6 Conclusions

This mixed-method study examined how professors and students in Bosnia and Herzegovina and Slovenia are engaging with Generative Artificial Intelligence in higher education. Despite the modest sample sizes, several patterns stand out. Both groups demonstrated moderate familiarity with GenAI, yet adoption differed. In Slovenia, respondents reported a broader range of purposes, especially pedagogical support and personalized learning. In Bosnia and Herzegovina, GenAI was framed more as a pragmatic resource for expanding access to literature and information. Perceptions of effectiveness and productivity were generally positive but tended to be described in instrumental rather than transformative terms. Ethical concerns were widespread, though demands for strict regulation were stronger in BiH, contrasting with more flexible—yet less informed—approaches in Slovenia. The findings contribute to debates on digitalization in higher education by showing that GenAI adoption is not merely technical but shaped by institutional cultures, infrastructures, and policy environments. Faculty and student narratives highlight ongoing tensions between innovation, productivity, and integrity in academic life. Practically, the results underscore the need for universities to address not only the pedagogical and technical integration of GenAI, but also its ethical and regulatory dimensions. Institutions should develop guidelines, training, and participatory forums to foster critical and responsible engagement with GenAI in teaching, learning, and research.

This exploratory cross-national study offers initial insights into how higher education actors in Bosnia and Herzegovina and Slovenia perceive and engage with GenAI. Findings highlight both institutional embedding and ethical awareness, underscoring the multidimensionality of GenAI adoption in academic contexts. Limitations include the use of small, convenience samples that limit generalizability; the cross-sectional design, which precludes analysis of change over time; and reliance on several single-item measures and self-reports that may bias assessments of familiarity and awareness. Future research should draw on larger, longitudinal, and mixed-methods designs to better capture the evolving dynamics of GenAI in higher education.

8 Acknowledgements

This work was supported by the Faculty of Organisation Studies in Novo mesto and the Faculty of Management and Business Economics from Bosnia and Hercegovina. The authors would like to thank all survey participants from Bosnia and Herzegovina and Slovenia for their valuable contributions.

9 References

[1] Abunaseer, H. *The Use of Generative AI in Education: Applications, and Impact*. *Technology and Curriculum*, Summer 2020, 2023.

[2] Alier, M., et al. *The impact of internet availability on student learning*. *Educational Technology Review*, 12(2):45–60, 2024.

[3] Aoun, J. E. *Robot-proof: Higher education in the age of artificial intelligence*. MIT Press, Cambridge, MA, USA, 2017.

[4] Baidoo-Anu, D., and Anash, L. *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.

[5] Baytas, M. A., and Ruediger, T. *Generative AI in academia: Adoption and futures*. *Journal of Higher Education Policy*, 36(1):55–72, 2024.

[6] Bergkvist, L., and Rossiter, J. R. *The predictive validity of multiple-item versus single-item measures of the same constructs*. *Journal of Marketing Research*, 44(2):175–184, 2007.

[7] Bryman, A. *Social research methods*. Oxford University Press, Oxford, UK, 2016.

[8] Chan, C., and Hu, W. *Students' voices on generative AI: perceptions, benefits, and challenges in higher education*. *International Journal of Educational Technology in Higher Education*, 20:43, 2023. https://doi.org/10.1186/s41239-023-00443-1

[9] Creswell, J. W., and Plano Clark, V. L. *Designing and conducting mixed methods research*. Sage, Thousand Oaks, CA, 2018.

[10] Davidov, E., Schmidt, P., and Billiet, J. *Cross-cultural analysis: Methods and applications*. Routledge, London, UK, 2011.

[11] Dillman, D. A., Smyth, J. D., and Christian, L. M. *Internet, phone, mail, and mixed-mode surveys: The tailored design method*. Wiley, Hoboken, NJ, 2014.

[9] Etikan, I., Musa, S. A., & Alkassim, R. S. Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, 5(1): 1–4, 2016.

[12] European Parliament and Council. *Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. *Official Journal of the European Union*, L 2024/1689, June 13, 2024.

[13] Field, A. *Discovering statistics using IBM SPSS statistics*. Sage, London, UK, 2018.

[14] Floridi, L., and Cowls, J. *A unified framework of five principles for AI in society*. *Harvard Data Science Review*, 1(1):1–15, 2019. https://doi.org/10.1162/99608f92.8cd550d1

[15] Guchhait, T. *The transformative influence of generative AI on modern education*. *International Journal of Research Publication and Reviews*, 5(8):4361–4371, 2024.

[16] Harkness, J. A., Villar, A., and Edwards, B. *Translation, adaptation, and design*. In Harkness, J. A., Braun, M., Edwards, B., Johnson, T. P., Lyberg, L., Mohler, P. P., … and Smith, T. W. (eds.), *Survey methods in multinational, multiregional, and multicultural contexts*, pp. 117–140. Wiley, Hoboken, NJ, 2010.

[17] Hsieh, H. F., and Shannon, S. E. *Three approaches to qualitative content analysis*. *Qualitative Health Research*, 15(9):1277–1288, 2005.

https://doi.org/10.1177/1049732305276687

[18] Hwang, G.-J., and Chen, N.-S. *Exploring the potential of generative artificial intelligence in education: Applications, challenges, and future research directions*. *Educational Technology & Society*, 26(2), 2023. https://doi.org/10.30191/ETS.202304_26(2).0014

[19] Kaplan-Rakowski, R., Grotewold, K., Hartwick, P., and Papin, K. *Generative AI and teachers' perspectives on its implementation in education*. *Journal of Interactive Learning Research*, 34(2):313–338, 2023.

[14] Knox, J., Williamson, B., & Bayne, S. Machine behaviourism: future visions of 'learnification' and 'datafication' across humans and digital technologies. *Learning, Media and Technology*, 45(1): 31–45, 2020. https://doi.org/10.1080/17439884.2019.1623251

[15] Lee, D., Arnold, M., Srivastava, A., Plastow, K., Strelan, P., Ploeckl, F., Lekkas, D., & Palmer, E. The impact of generative AI on higher education learning and teaching: A study of educators' perspectives. *Computers and Education: Artificial Intelligence*, 6, 100221, 2024. https://doi.org/10.1016/j.caeai.2024.100221

[16] Lim, W., Gunasekara, A., Leigh Pallant, J., Pallant, J., & Pechenkina, E. Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education*, 21(2), 100790, 2023. https://doi.org/10.1016/j.ijme.2023.100790

[17] Luo, J. How does GenAI affect trust in teacher-student relationships? Insights from students' assessment experiences. *Teaching in Higher Education*, 2024. https://doi.org/10.1080/13562517.2024.2341005

[18] Mauss, M. (1990). The gift: The form and reason for exchange in archaic societies (W. D. Halls, Trans.). London: Routledge. (Original work published 1925).

[19] McDonald, N., Johri, A., Ali, A., & Hingle, A. Generative Artificial Intelligence in Higher Education: Evidence from an Analysis of Institutional Policies and Guidelines. *Generative Artificial Intelligence in Higher Education*, 2024.

[20] Molek, N. AI and organizational transformation: anthropological insights into higher education. *Izzivi prihodnosti*, 8(3): 148–177, 2023. https://doi.org/10.37886/ip.2023.007

[21] Noroozi, O., Soleimani, S., Farrokhnia, M., & Banihashem, S. Generative AI in education: Pedagogical, theoretical, and methodological perspectives. *International Journal of Technology in Education*, 7(3): 373–385, 2024. https://doi.org/10.46328/ijte.845

[22] Ogunleye, B., Zakariyyah, K., Ajao, O., Olayinka, O., & Sharma, H. A systematic review of generative AI for teaching and learning practice. *Education Sciences*, 14(6): 636, 2024. https://doi.org/10.3390/educsci14060636

[23] Pavlik, J., & Pavlik, O. Art education and generative AI: An exploratory study in constructivist learning and visualization automation for the classroom. *Creative Education*, 15: 601–616, 2024.

[24] Ragin, C. C. *Constructing social research*. Pine Forge Press, Thousand Oaks, CA,

1994.

[25] Selwyn, N. *Education and technology: Key issues and debates* (3rd ed.). Bloomsbury, London, UK, 2022.

[26] Small, M. L. How many cases do I need? *Ethnography*, 10(1): 5–38, 2009.

[27] Tashakkori, A., & Teddlie, C. *Sage handbook of mixed methods in social & behavioral research*. Sage, Thousand Oaks, CA, 2010.

[28] van de Vijver, F. J. R., & Leung, K. Methods and data analysis of comparative research. In Berry, J. W., Poortinga, Y. H., Breugelmans, S. M., Chasiotis, A., & Sam, D. L. (eds.), *Cross-cultural psychology: Research and applications*, pp. 33–60. Cambridge University Press, Cambridge, UK, 1997.

[29] Williamson, B., & Eynon, R. Mapping AI imaginaries in education. *Learning, Media and Technology*, 45(3): 1–16, 2020.

[30] Wong, L., & Looi, C. Advancing the generative AI in education research agenda: Insights from the Asia-Pacific region. *Asia Pacific Journal of Education*, 44(1): 1–7, 2024. https://doi.org/10.1080/02188791.2024.2315704

[31] Xia, Q., Weng, X., & Ouyang, F. A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education*, 21: 40, 2024. https://doi.org/10.1186/s41239-024-00468-z

ABSTRACTS

Tracking Physiological System Integration in Cardiac Rehabilitation Using the Cross-Vector Approach

Pavle Boškoski, Martin Brešar
Jožef Stefan Institute
Jamova cesta 39, 1000 Ljubljana, Slovenia
Faculty of Information Studies
Ljubljanska cesta 31A, 8000 Novo mesto, Slovenia
{pavle.boskoski, martin.bresar}@fis.unm.si

Abstract: *We present the analysis of physiological measurements from coronary patients during rehabilitation following a coronary event. A key feature we analyse is how different physiological systems interact. An increase in interaction strength reflects stronger integration between physiological systems, indicating effective recovery. We analyse signals recorded at the start of the rehabilitation and again after three months, allowing for monitoring changes during this critical period. The dataset includes signals of tissue oxygenation, blood flow, respiration, and electrocardiogram activity. We apply the recently developed cross-vector approach for characterizing interactions from measured time series. This is a nonlinear method based on reconstructed state-spaces. It is broadly applicable to various deterministic and stochastic dynamics. It reliably identifies both the strength and the direction of interactions even in short and noisy data. Applied to patient data, the method reveals changes in interaction strength and direction that may reflect underlying physiological changes. The resulting features hold promise for tracking rehabilitation progress at an early stage and for continuous monitoring during long-term rehabilitation.*

Key Words*: interactions, nonlinear dynamics, coronary rehabilitation*

Evolution of topics in Slovenian science

Borut Lužar, Nika Robida
Faculty of Information Studies
Ljubljanska cesta 31A, 8000 Novo mesto, Slovenia
{borut.luzar, nika.robida}@fis.unm.si

Abstract: *We present an analysis of the development of Slovenian science from 1975 to 2024. Based on the keywords extracted from scientific articles published by Slovenian authors, we created a keyword co-occurrence network (KCN) for each five-year period and, using community detection, detected topics based on communities of keywords. We assigned a disciplinary profile to each community by aggregating the scientific fields of its authors (using Universal Decimal Classification (UDC)). This enabled us to compare topic development across nine primary UDC disciplines. The resulting timeline highlights persistent, emerging, declining, and branching topics, and allows us to explore potential drivers of topic growth, transformation, or disappearance, revealing some notable differences between scientific disciplines.*

Key Words*: Slovenian science, topic evolution, keyword co-occurrence network*

Climate Cards: Battle of the Weather Stations!

Matija Klančar
Faculty of Information Studies
Ljubljanska cesta 31A, 8000 Novo mesto, Slovenia
matija.klancar@gmail.com

Abstract: *This project presents an innovative educational card game designed to familiarize players with climate data from Slovenian meteorological stations. Inspired by classic comparative card games — such as those featuring cars, where players compete based on the highest engine power or acceleration — this game applies a similar format to environmental education. Each card represents a different meteorological station in Slovenia and includes a set of relevant climate indicators, especially temperature records. The gameplay is simple and competitive: players take turns choosing a specific variable to compare, then each reveals the corresponding value from their top card. The player with the highest (or lowest, depending on the category) value wins the round and collects the cards. The goal is to collect as many cards as possible, while also gaining a deeper understanding of local climate characteristics. The cards are based on homogenized meteorological data, allowing players to engage with authentic information in a fun and interactive way. The selection of indicators encourages reflection on both short-term weather phenomena and long-term climate patterns, offering opportunities to discuss climate change, regional differences, and the importance of long-term observations. In addition to entertainment, the game serves as a learning tool suitable for use in schools, workshops, and informal educational settings. It can be adapted for different age groups by adjusting the complexity of indicators or including additional context and explanations. The project also opens doors for further expansion — such as thematic sets (e.g., extreme events, climate change scenarios) or digital versions with interactive visualizations. By combining elements of play, competition, and environmental science, this card game offers a unique approach to raising awareness about climate and fostering curiosity about the natural world through localized, data-driven content.*

Key Words*: data science, gamification, education, meteorology, climatology*

Loneliness, Reflexivity, and AI in Youth

Tea Golob, Matej Makarovič, Romina Gurashi*
Faculty of Information Studies
Ljubljanska cesta 31A, 8000 Novo mesto, Slovenia
*Università degli Studi Internazionali di Roma
Via Cristoforo Colombo, 200, 00147 Roma, Italy
tea.golob@fis.unm.si, matej.makarovic@fis.unm.si
romina.gurashi@uniroma1.it

Abstract: *In the presentation, we address the issue of loneliness in relation to the increasingly all-encompassing use of the AI and the role reflexivity is playing in this regard. Based on previous research we presume that reflexivity significantly affects how one constitutes the world, engages in social relationship, navigates the media contents, and also engages in purposive action related to a more sustainable future. Based on that, we hypothesise that those who are highly reflexive tend to better cope with the potentially negative effects of AI – maintaining healthier social relations and personal empowerment. Research has been conducted on a national representative sample in Slovenia enabling a comparison between the older generations and the younger ones that have experienced direct and immediate socialization in a technologically mediated world, increasingly supported by forms of Artificial Intelligence.*

Key Words*: loneliness, reflexivity, and AI in youth*

Social Infrastructure for Digitalization

Urša Lamut
Rudolfovo - - Science and Technology Centre Novo Mesto
Podbreznik 15, 8000 Novo mesto, Slovenia
ursa.lamut@rudolfovo.eu

Abstract: *This study examines how employees perceive workplace digitalization through institutional, network, and cognitive frames (Beckert, 2010). We treat digitalization as an organizational–cultural transformation, not just a technical upgrade. The objective is to identify conditions that enable or hinder adoption beyond tool rollouts. In 2025, we conducted heterogeneous focus groups in six companies (47 participants). Data was audio-recorded and analysed using qualitative content analysis. Findings show that digitalization is sustainable when tools align with values and employees understand the purpose of change and help co-shape it; otherwise, tensions arise around trust, privacy, relationships, and the preservation of competencies. This requires social—as well as digital—infrastructure: open communication, collaboration, organizational learning, and explicit handling of value tensions. Through a field-theoretic lens, three patterns emerge. Institutionally, adoption accelerates when technology, culture, leadership, vision, and structured training are aligned; lowest-price procurement and late regulator/union involvement slow implementation, and knowledge transfer from training is often patchy. Relationally, success depends on trust, precise goal alignment, and transparent data sharing with partners; administrative burdens, cultural gaps, and research-to-industry lags impede uptake, while education systems refresh digital skills slowly. Cognitively, employees adopt tools they perceive as error-reducing, growth-enabling, and value-consistent; resistance grows with surveillance cues and added administrative load. Two-way communication and mentoring shorten the path from pilots to routine. Contributions: (i) an actionable tri-frame diagnostic to align rules, relationships, and meanings; (ii) evidence that human-centric communication and training improve adoption trajectories; and (iii) guidance for designing the social infrastructure that sustains digital transformation.*

Key Words*: workplace digitalization, social infrastructure, institutions, networks, cognitive frames*

Reliable and Trustworthy Use of AI in Cultural Heritage Digital Transformation

Ines Vodopivec
AI4LAM
ines.vodopivec@nb.no

Abstract: *The AI for Libraries, Archives, and Museums (AI4LAM) community is a global, cross-sectoral, and participatory network committed to advancing artificial intelligence within the cultural heritage sector. It spearheads the development of cutting-edge AI tools and services tailored to heritage institutions - key custodians of the data foundational to digital humanities and AI model training. AI-driven solutions offer transformative potential: revolutionizing data management, enhancing access to digital collections, and streamlining administrative workflows. AI4LAM's mission is to establish a robust framework for organizing, sharing, and evolving knowledge around AI, while promoting the adoption of trustworthy and transparent technologies. Through collaborative exchange, the community fosters innovation across GLAM institutions, academia, and education, addressing the sector's evolving needs in the coming decade. Its dedication to open knowledge aligns with the principles of Open Science, supporting more transparent, accessible, and participatory research practices. Integrating AI into GLAM digital infrastructures equips researchers with powerful tools to analyze and interpret vast datasets. Yet this progress invites a critical reflection: In an AI-enhanced environment, do we still need metadata as we know it - or must we rethink this foundational concept?*
Drawing on successful case studies, this paper proposes a theoretical framework for implementing AI tools and services in cultural heritage workflows. The model is explored through the interconnected lenses of user engagement, data stewardship, and infrastructure management.

Key Words*: artificial intelligence, cultural heritage, digital transformation, AI4LAM*

Cultural Heritage analysis with YOLO based object detection

Lucijano Berus, Vesna Pungerčar
Rudolfovo - Science and Technology
Centre Novo Mesto
Podbreznik 15, 8000 Novo mesto, Slovenia
{lucijano.berus, vesna.pungercar}@rudolfovo.eu

Abstract: *Cultural heritage artefacts that are rich in engraved and embossed ornamentation on vessels, ritual objects, tombstones, and manuscripts. These objects are important for reconstructing social life, ritual practices, and cultural expression across regions and periods. To understand a culture, it is not enough to study objects in isolation; systematic comparison across related artefacts is essential to determine whether and how communities were connected. However, such a comparison requires first a robust, scalable detection of their visual content. We therefore study whether a real-time object detection framework can localise and classify ornamentation. In this study, pretrained You Only Look Once version 8 (YOLOv8) and version 11 (YOLOv11) architectures were employed, ranging from their nano to large model versions, to detect ornaments characteristic of Greek and Hallstatt cultural artefacts. YOLOv8 and YOLOv11 were pretrained on Common Objects in Context (COCO) dataset and were able to detect 80 different object categories. During the testing of YOLO performance different inherent (YOLO specific) hyper-parameter settings were adopted to detect (localise and classify) ornaments. The models demonstrated promising performance in localising and recognising recurring motifs, yet their accuracy remains constrained by the limited availability of ornament-specific training data. To enhance recognition quality, the development of specialised datasets tailored to cultural ornamentation is essential.*

Key Words*: cultural heritage, object detection, ornamentation, deep learning, YOLO*

A Two-Phase AI Framework for Fresco Analysis and Digital Restoration

Branimir Kolarek, Davor Davidović
Ruđer Bošković Institute
Bijenička Cesta 54, 10000, Zagreb, Croatia
branimir.kolarek@irb.hr, davor.davidovic@irb.hr

Abstract: *The conservation of cultural heritage frescoes is hampered by subjective, time-consuming damage documentation and the inherent risks of irreversible physical intervention. Our research objective is to develop a multi-stage, AI-driven framework that can act as a transparent co-pilot for art conservators, enhancing their analytical capabilities without dictating outcomes. The research will be approached in two phases. Phase One involves developing a tool for objective damage analysis. This computer vision system will be designed to automatically identify, classify, and map various forms of damage from high-resolution imagery. The envisioned contribution of this phase is an efficient tool that standardizes condition reporting by generating quantified reports and precise, color-coded damage maps, creating a digital baseline for monitoring changes over time. Phase Two will address the more ambitious challenge of digital restoration by developing a system for generating historically plausible inpainting suggestions. Our proposed methodology involves constructing a multi-modal Knowledge Graph of stylistic and iconographic evidence to heavily constrain a retrieval-augmented generative model. This approach is designed to ensure all proposals are based on historical precedent, not artistic invention. The project aims to culminate in an expert-in-the-loop visualization tool for exploring plausible restorations, complete with a "Certainty Map" tracing each suggestion back to its supporting evidence.*

Key Words: *fresco restoration, computer vision, digital inpainting, cultural heritage, AI-assisted conservation*

Promoting Digital Inclusion through Digital Tools: The Role of eAsistent in the Work of Secondary School Teachers in Slovenia

Branka Klarić
Šolski center Novo mesto
Šegova ulica 112, 8000 Novo mesto, Slovenia
branka.klaric@sc-nm.si

Abstract: *Research explores the importance of teachers' digital literacy and the role of digital platforms in their work. Digitalization is becoming a key part of education, and teachers need to develop digital competencies for the effective use of digital tools. The theoretical part discusses the difference between information and digital literacy, historical development and the state of digital competencies of teachers in Slovenia. Special emphasis is placed on the impact of digital literacy on the employability and professional development of teachers, and on digital inclusion. The use of the eAsistent digital platform in secondary schools is analyzed, as well as its impact on school administration and teacher work. The empirical part examines the impact of using eAsistent on teachers' digital competencies, and the final part contains results and suggestions for further research.*

Key Words*: digital literacy, digital platforms, eAsisitent, teachers*

A Digital Transformation for Student Living in Slovenia: Addressing Key Challenges

Nejra Arnautović, Denis Kantić
Faculty of Information Studies
Ljubljanska cesta 31A, 8000 Novo mesto, Slovenia
arnautovic.nejra@gmail.com, contact@deniskantic.net

Abstract: *Student living in Slovenia is challenged by persistent issues: limited housing availability, rising rental costs, and complex roommate searches. This work presents Domko.si, a comprehensive digital platform that fundamentally reshapes the student housing experience through digital transformation. The platform uniquely integrates two core functionalities: secure Housing Listings posted by verified property owners, and an intelligent Roommate Finder service. The primary methodology involves market analysis, user surveys, and prototype testing to evaluate the platform's effectiveness in reducing barriers to access and fostering a transparent rental market. Critically, Domko.si employs Artificial Intelligence (AI) to validate listing quality, checking content consistency, image clarity, and accuracy against titles. Furthermore, security is maintained through required student certificate submissions for administrative validation upon registration. Preliminary findings indicate that this approach significantly enhances trust and efficiency. The expected contribution is a scalable model for digital housing platforms, offering key insights for European universities and policymakers aiming to modernize the practical and social dimensions of student living. Domko.si demonstrates a meaningful digital intervention in a critical socio-economic sector.*

Key Words*: digital transformation, AI-Powered validation, socio-economic development of Slovenia, student accommodation, trust in digital services*

Inverse Modeling of Flexible Rotational Systems via Physics-Informed Deep Learning

Miloš Ivanović, Milan Matijević, Lazar Krstić*
Faculty of Science, University of Kragujevac
*Faculty of Engineering, University of Kragujevac
Liceja Knezevine Srbije 1A, Kragujevac, Serbia
mivanovic@kg.ac.rs, milan.matijevic@gmail.com,
lazar.krstic@pmf.kg.ac.rs

Abstract: *Inverse modeling of flexible rotational systems, such as motor-load assemblies with elastic couplings, is challenging due to nonlinear dynamics, resonant modes, and non-collocated sensing. This work presents a physics-informed deep learning approach to accurately infer the input momentum required to achieve a desired angular displacement. We employ a Physics-Informed Neural Network (PINN) that encodes the system's governing second-order differential equations directly into the loss function, ensuring physical consistency without explicit numerical integration. To address the critical challenge of PINN architecture selection, we employ a tailored Genetic Algorithm (PINN/GA) that automatically optimizes network depth, layer size, activation functions, and optimizer choice. Starting from simple architectures, the evolutionary strategy progressively adds complexity only when it improves accuracy, balancing performance and simplicity. Being computationally highly expensive, the PINN/GA search runs on an HPC cluster with a Kubernetes deployment, making use of Kubernetes batching capabilities. The method is validated on a torsional system with flexible coupling, where the optimized PINN successfully reconstructs the driving torque from the load trajectory. Results show that the GA-optimized PINN achieves higher accuracy and better generalization compared to random search and Hyperopt-based methods, while maintaining computationally efficient architectures. The approach offers a robust, automated pipeline for inverse modeling in mechatronic systems, with potential applications in precision motion control, robotics, and industrial automation.*

Key Words*:* *inverse modeling, physics-informed deep learning, high-performance computing, optimization, rotational systems*

PINN: machine learning on complex systems described by small or limited quality datasets

Andrej Furlan
Faculty of Information Studies
Ljubljanska cesta 31A, 8000 Novo mesto, Slovenia
andrej.furlan@fis.unm.si

Abstract: *The "classical" machine learning (ML) suffers from weakness, that it is very greedy and sensitive to input data. The performances of ML models are completely dependent on the quantity and quality of the data used to train the models, and, the output is strictly dependent on the quality of input data being analysed. There are however, many fields, where collection of data is impractical or expensive, resulting in small datasets, sometimes also of questionable quality, thus often limiting severely application of ML tools in such areas. In order to address such data scarce complex systems, Physical Informed Neural Networks (PINN) methods have been developed recently. Those machine learning methods are relying on physical principles underlying the systems being analysed, in order to find patters in the data, and make predictions. The effect is that, now, classical neural networks are, in addition by data, constrained also by differential equations describing the system, giving thus to the data also a physical sense. The such defined neural networks also obtain a boost for parts of the vector space where data are missing. In this work, we demonstrate, using python libraries for deep neural networks, the extension of the standard data loss function by the additional part – physics loss – which introduces physical constraints to neural networks used to model the systems. We focus also on the complex systems, of which description can potentially be extended to areas of application such as climate science and medicine.*

Key Words*: PINN, neural networks, physical principles, complex systems, limited datasets*

Smart AI-Based System for Turning Tool Condition Monitoring

Nika Brili
Rudolfovo - Science and Technology
Centre Novo Mesto
Podbreznik 15, 8000 Novo mesto, Slovenia
nika.brili@rudolfovo.eu

Abstract: *The turning process is a widely used cutting operation in industry. Any optimization of this process can significantly improve product quality, streamline costs, or reduce unwanted events. With automatic monitoring of turning tools, we can reduce costs, increase efficiency, and decrease the number of undesirable events that occur during machining (scrap, tool breakage, etc.). In single-piece or small-batch production, tool wear is monitored by the machine operator; however, such wear assessment is left to subjective judgment and requires intervention in the process. The presented solution eliminates this problem with automated monitoring of the cutting tool's condition. An IR camera was used for process monitoring, which also captures the thermographic state. The camera was properly protected and mounted right next to the turning tool, enabling close-up observation of the machining. During the experiment, constant cutting parameters were set for turning the workpiece (low-alloy steel designated 1.7225, i.e., 42CrMo4) without the use of coolant. Using turning inserts with varying levels of wear, a database of more than 6,000 images was created during the turning process. With a convolutional neural network (CNN), a model was developed to predict wear and damage to the cutting tool. Based on the captured thermographic image during turning, the model automatically determines the cutting tool's condition (no wear, minor wear, severe wear). The achieved classification accuracy was 99.55%, confirming the suitability of the proposed method. Such a system enables immediate action in the event of tool wear or breakage, regardless of the operator's knowledge and training.*

Key Words: *deep learning, tool condition monitoring, turning, tool wear*

Perceived vs. Actual Spatial Data Quality: Challenges, Consequences, and an Innovative Management Framework

Tomaž Podobnikar
Faculty of Information Studies
Ljubljanska cesta 31A, 8000 Novo mesto, Slovenia;
Ministry of Natural Resources and Spatial Planning
Dunajska cesta 48, 1000 Ljubljana, Slovenia
tomaz.podobnikar@fis.unm.si

Abstract: *Quality is often regarded as an abstract and elusive concept, becoming tangible only when deficiencies undermine real-world decisions. In spatial planning, for example, inaccuracies in planned land use data within the building permit process may lead to biased outcomes or, in critical cases, impede decision-making altogether. This study examines use cases of spatial data quality issues and their potential consequences, emphasizing both technical dimensions and the human factor, with particular attention to the gap between perceived and actual data quality. The research introduces the concept of this discrepancy as a key challenge in governance, with implications for decision-making processes and stakeholder trust. Building on these insights, an innovative approach to spatial data quality management is proposed through an automated Data Quality Management (DQM) framework. The framework is operationalized via the OPIAvalid toolkit, designed to enhance the effectiveness and reliability of national Spatial Information Systems (PIS). The expected contribution lies in providing a systematic methodology for managing spatial data quality, thereby strengthening evidence-based decision-making and fostering greater trust in spatial data infrastructures.*

Key Words*: perceived and actual data quality, data quality management (DQM), quality assurance/quality control (QA/QC), spatial data quality, automation*

Following Quantum Innovation Flows: The Feedback Loop Between Strategic Timing and Patent Activity (2014–2023)

Tamara Besednjak Valič, Karin Dobravc Škof
Rudolfovo – Znanstveno in tehnološko središče Novo mesto
Podbreznik 15, 8000 Novo mesto, Slovenia
{tamara.valic, Karin.dobravc.skof}@rudolfovo.eu

Abstract: *Quantum technologies are central to the global innovation race. While national strategies are designed to secure technological sovereignty, the relationship between strategic timing and actual innovation output is complex. However, the fundamental question remains: does policy actively drive the innovation cycle or merely follow it? This study addresses this relationship by focusing on the temporal alignment between the release of national quantum strategies and the resulting patent application volume across countries (2014–2023).*

Utilizing PATSTAT data, with a focus on the patent application date, we establish that the global application peak occurred in 2022. This finding reveals a significant temporal paradox: while early movers like the US (National Strategic Overview for Quantum Information Systems and Related Documents, 2018) and the Netherlands (National Agenda for Quantum Technology, 2019) acted proactively, the majority of nations (including Germany, France, and Japan) released their strategies in 2023—after the innovation peak had already been reached. We further analyse the China situation (leading patent volume without a publicly available strategy) and the Netherlands paradox (early strategy despite low domestic patent count).

The study's primary quantitative measure is the lag time between a strategy's publication date and the subsequent peak in a nation's domestic quantum patent applications. By analysing this temporal gap, the research provides empirical evidence to validate the effectiveness of strategic foresight versus reactive policymaking.

Key Words*: quantum technologies, national strategies, innovation flows, patent activity*

Methodological Framework for Studying Industrial Path Development: Social Fields Analysis

Kseniia Gromova
Faculty of Information Studies
Ljubljanska cesta 31A, 8000 Novo mesto, Slovenia
kseniia.gromova@fis.unm.si

Abstract: *The contribution presents the methodological framework for an ongoing doctoral study that explores why some industrial localities thrive while others lag behind in their development. Applying the Social-Fields-Approach (SOFIA), the research regards industrial localities as social fields influenced by three main social forces: institutions, social networks, and cognitive frames. It examines how these forces (and their combined impact) enable or hinder new path creation within the selected localities during the latter half of the 20th century. A comparative multiple case study will be conducted on three distinct industrial localities – Novo mesto (Slovenia), Pernik (Bulgaria), Aalborg (Denmark) – selected via purposive sampling as localities with varying levels of innovation performance. Data collection will involve two rounds of semi-structured interviews with key stakeholders from business, academia, and policy-making areas in each industrial locality, as well as with experts in local regional development. The first round aims to identify the main periods and turning points within the developmental trajectory of each industrial locality since the 1950s; the second round will consider the impact of the three social forces (institutions, cognitive frames, social networks) within each period and turning point while shaping the path-creation process as well as the success of a certain locality. The research contributes to the existing theory and practice in regional studies by offering new insights into the reasons behind the uneven geography of industrial development through the new application of the SOFIA conceptual framework.*

Key Words*: Social-Fields-Approach (SOFIA); industrial locality; path-creation; comparative case study; regional development*