



ITIS 2023 BOOK OF PROCEEDINGS

"Future of digital society in
the age of AI and
ChatGPT"

14TH INTERNATIONAL CONFERENCE
ON INFORMATION TECHNOLOGIES
AND INFORMATION SOCIETY

NOVEMBER 2023, LJUBLJANA, SLOVENIA

14th International Conference on Information Technologies and Information Society

ITIS 2023

Conference Proceedings

November 9-10, 2023

Ljubljana, Slovenia

Organized by



Faculty of Information Studies in Novo mesto

Supported by



CONFERENCE COMMITTEES

Organizing Committee

Nuša Erman, *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](#)

Petra Roginič

[Faculty of Information Studies in Novo mesto, Slovenia](#)

Katja Peterlin

[Faculty of Information Studies in Novo mesto, Slovenia](#)

Program Committee

Zoran Levnajič, *chair*

[Faculty of Information Studies in Novo mesto, Slovenia](#)

Filipo Sharevski

[DePaul University, USA](#)

Sanda Martinčić – Ipšić

[University of Rijeka, Croatia](#)

Petra Kralj Novak

[Central European University, Austria](#)

Vesna Andova

[“SS. Cyril and Methodius” University in Skopje, North Macedonia](#)

Tomáš Hlavsa

[Czech University of Life Sciences Prague, Czech Republic](#)

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](#)

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

Published in: 2024

Edited by: Nuša Erman

Kataložni zapis o publikaciji (CIP) pripravili v Narodni in univerzitetni knjižnici v Ljubljani

COBISS.SI-ID 182711555

ISBN 978-961-96549-0-3 (PDF)

EDITOR'S 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

Thursday, 9th November		
08.30 - 09.00	<i>Registration, meet & greet</i>	
09.00 - 09.20	Welcome address by Matej Makarovič, Dean of Faculty of information studies in Novo mesto Welcome address by Miroslav Kranjc, Ministry for Digital Transformation of Slovenia	
09.20 - 10.20	Keynote lecture by Denise Potosky, Penn State University, USA "I am not a robot" – Human Communication in the AI Era	
10.20 - 11.00	Nika Brili ChatGPT and Science: From Introduction to Practical Examples Biljana Mileva Boshkoska and Katarina Rojko Identification of the strengths and weaknesses of outdoor STEM learning using LLM	First session: ChatGPT and LLM (Under the project AI4VET4AI) (chairs Biljana Mileva Boshkoska, Šrđan Škrbić)
11.00 - 11.30	<i>Coffee break</i>	
11.30 - 12.50	Fatima Aziz and Martin Žnidaršič Potential of lexicon-based bug severity classification Nuša Erman and Katarina Rojko Addressing challenges in higher education study programs: The case of the RRP pilot project "Advanced Con Mustafa Bešić and Selena Kurtić AI-Based Analysis and Studying Cheating Behaviours Ana Hafner and Adela Černigoj ChatGPT as Psychic: Excellent for Ceremonial Speeches	Second session: General applications of AI (chairs Srđan Škrbić, Andrej Furlan)
12.50 - 14.30	<i>Lunch</i>	
14.30 - 15.30	Keynote lecture by Slavko Žitnik, University of Ljubljana, Slovenia How did we come to chat with AI?	
15.30 - 16.30	Nuša Erman and Hlavsa Tomaš Comparative analysis of Slovenian and Czech digitalization index: an AI approach Valerij Grasic and Biljana Mileva Boshkoska Future-based smart city with AI-powered flood support Žan Pogač and Tomaž Aljaž Enhancing Urban Traffic Flow with AI: A Case Study of YOLO-Based Vehicle Counting	Third session: Circular, green & environmental (Under the project SRC EDIH) (chairs Nuša Erman, Zoran Levnajić)
16.30 - 17.00	<i>Coffee break</i>	
17.00 - 18.20	Jelena Joksimović, Zoran Levnajić and Bernard Ženko Detecting Corruption in Slovenian Public Spending from Temporal Data Biljana Jolevska Tuneska Mathematica and art: designs with RegionPlot Albert Zorko and Zoran Levnajić Using AI and physiological data for diagnosing depression Marijana Ribičić The interconnection of Artificial Intelligence and Property Technology: literature review	Fourth session: Society 5.0 and Industry 4.0 (chairs Blaž Rodič, Srđan Škrbić)
18.20 - 18.40	Janez Povh Rudolfovo - a new driver of applied research	
19.00 - 19.30	<i>Stand up - Marina Orsag</i>	
19.30 -	<i>Dinner</i>	
Friday, 10th November		
08.30 - 09.00	<i>Meet & greet</i>	
09.00 - 10.00	Keynote by Juraj Petrović, University of Zagreb, Croatia The future of education in the era of ChatGPT	
10.00 - 11.20	Milica Stankovic, Gordana Mrdak and Tiana Andjelkovic The Role of Artificial Intelligence in Circular Economy Erika D. Uršič, Urška Fric and Alenka Pandiloska Jurak Test Environment for the Implementation of the Circular Economy Mare Srbinovska, Vesna Andova, Aleksandra Krkoleva Mateska and Maja Celeska Krsteska Breathing Easy in North Macedonia: The Effect of Green Infrastructure and Movement Restrictions on The Air Mohamed Abdel Maksoud The Ecological Impact of Server-Side Rendering	Fourth session: Society 5.0 and Industry 4.0 (chairs Blaž Rodič, Srđan Škrbić)
11.20 - 11.50	<i>Coffee break</i>	
11.50 - 13.50	Blaž Rodič and Matej Barbo Application of Simulation Modelling for Decision Support in Traffic Safety Matjaž Drev and Bostjan Delak Automating privacy compliance Dino Arnaut and Damir Bečirović The Necessity of Digital Transformation of Developing Countries: The Case of Bosnia and Herzegovina Peter Zupančič and Pance Panov Clustering of Employee Absence Data: A Case Study Bojan Pažek and Slavko Arh A Multi-Factorial ANOVA Framework for STL File Quality Optimization via Point Cloud Analysis Lavdim Menxhiqi and Galia Marinova Knowledge base assisting PCB Design tool selection and combination in Online CADCOM platform	Fourth session: Society 5.0 and Industry 4.0 (chairs Blaž Rodič, Srđan Škrbić)
13.50 - 14.20	Victor Cepoi (JM Module TIC2030) and Tea Golob Jean Monnet Module "Technology and Innovation Communities 2030": Reflexive responsibility and society 5.0	

KEYNOTE SPEAKERS

Lectures by keynote speakers

"I am not a robot" – Human Communication in the AI Era

by [Denise Potosky](#), *Professor of Management and Organization, Penn State University, USA*

How did we come to chat with AI?

by [Slavko Žitnik](#), *Laboratory for Data Technologies, Faculty for computer and information science, University of Ljubljana, Slovenia*

The future of education in the era of ChatGPT

by [Juraj Petrović](#), *Department of Electronic Systems and Information Processing, Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia*

CONFERENCE PAPERS

ChatGPT and Science: From Introduction to Practical Examples

Nika Brili

Rudolfovo – Science and Technology Centre Novo mesto
Podbreznik 15, 8000 Novo mesto, Slovenia
nika.brili@rudolfovo.eu

Abstract: *This paper explores the impact of ChatGPT on academic research and scientific discourse. The technical aspects of ChatGPT are presented, including its generative pre-trained Transformer (GPT) architecture and tokenization process, which supports its text generation and processing abilities. Despite its advanced capabilities, ChatGPT is constrained by a token limit and a knowledge cutoff, limiting its access to the most recent scientific, technological, and societal developments.*

The paper concludes by examining the role of ChatGPT in writing scientific papers, highlighting its diverse applications, from literature analysis to hypothesis generation and educational support. The varying guidelines and rules set by publishers and academic bodies regarding the use of ChatGPT in scientific work are also addressed.

Key Words: *ChatGPT, Large Language Models, Research, Artificial Intelligence*

1 Introduction

ChatGPT is a natural language processing model developed by OpenAI. Its main advantage lies in its remarkable capability to analyse extensive data and generate comprehensible textual responses in a conversational format. The potential of ChatGPT has already been recognized in academic research [1], demonstrating its utility in streamlining literature reviews, formulating hypotheses, and even assisting in writing research papers. Furthermore, its ability to engage in dialogue simulates scholarly discourse, offering researchers a novel tool for exploring complex scientific questions.

Disclaimer: ChatGPT was used in the process of scientific research.

2 Large Language Models

The name ChatGPT comes from the acronym:

- ▲ G – Generative (can "generate" or create text)
- ▲ P – Pre-trained (already trained on vast amounts of text)
- ▲ T – Transformer (machine learning model, good at handling sequences of information, like language; transformers are brains of the model) – see Figure 1

The most common function of Large Language Models (LLM) is text creation. However, these AI systems can create visual content, such as images. Furthermore, these models extend their versatility to generate music by interpreting patterns and structures characteristic of musical composition [3]. ChatGPT can communicate in multiple languages in the single conversation. It is not limited to major global languages, it can interact in less common dialects.

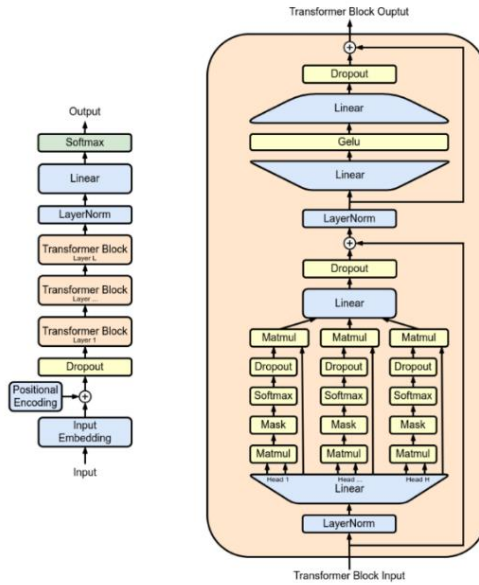


Figure 1: Original GPT network architecture [2]

2.1 How does ChatGPT process the text?

The operation of AI language models like ChatGPT involves translating text into tokens (Figure 2), which can be considered numerical representations of pieces of language, much like numbers. These tokens act as the building blocks or "letters" for the model. ChatGPT processes input by interpreting these tokens. The 'alphabet,' or the range of tokens recognized by GPT-3, is expansive, encompassing over 50,000 unique tokens. This vast library allows for a wide variety of language patterns and nuances to be understood and generated by the model.

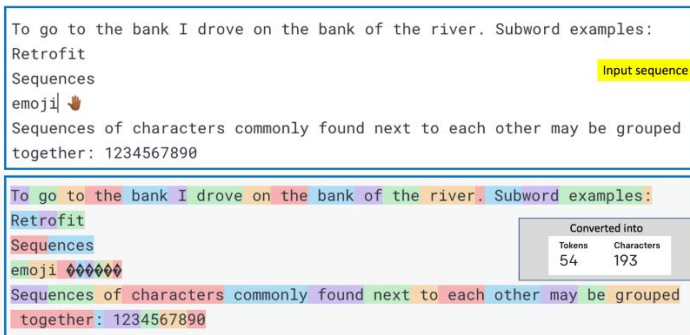


Figure 2: Tokenization of input sequence [4]

However, there is a limitation to the number of tokens that GPT models can handle in a single processing instance. When the token limit is reached, the model does not consider any excess tokens during that interaction. This constraint necessitates efficient use of language when interacting with the model to ensure that the most essential information is conveyed within the token limit and thereby receives attention from the AI.

2.2 Training and fine-tuning

ChatGPT was trained on billions of words from diverse sources, like books, web pages, and news articles (for example, Wikipedia represents only 3% of its knowledge database) [4]. Its database goes up to September 2021 for GPT-3 and January 2022 for GPT-4. Later on, it was trained and fine-tuned with the assistance of human trainers, whose task was to rank different ChatGPT outputs to the same prompt from best to worst. However, it is not totally clear in detail how ChatGPT was trained [5].

2.3 Generating the text

The model operates on the principle of predicting the next word in a sequence. When you input a question or sentences, the model tries to predict which word is most likely to follow to form a coherent response.

Every word the model generates is based on probabilities determined during its training. Predictions are not based on the meaning of sentences, but on how often individual words statistically appear next to other words. The model does not always select the token with the highest probability score for the output, thereby the generated answers are never exactly the same.

3 Limitations of ChatGPT

ChatGPT's knowledge is confined to information available up until a specific cutoff date; hence, it lacks the latest updates in science, technology, and societal developments. The information it provides is based on pre-existing texts, which may contain inaccuracies or lack precision, as the model is not capable of real-time fact-checking or exercising critical thinking [6].

The model operates with a significant dependence on human-generated labels and data categorization, which guide its understanding and responses. Whether ChatGPT can deliver a qualified answer to a question depends on the scope of the inquiry and the quality of the data it was trained on. While users cannot directly verify the answers in real-time through the model, they can cross-reference the information with reliable sources.

There exists a high risk of propagating misinformation if the model's responses are taken at face value without additional verification, especially given its limitations and reliance on possibly flawed data. As for the intentions of the provider OpenAI it aims to inform users about these limitations, encouraging critical engagement with the model's responses and an understanding that it should not be seen as an infallible source of information.

3.1 Data privacy

Concerns regarding data protection pose significant challenges for ChatGPT. The tool operates as a "black box," meaning the intricate workings of how it processes and stores data are not entirely transparent. Since the servers hosting ChatGPT are located in the United States, adherence to the European Union's stringent data protection standards is not assured, raising concerns about the security of data that is inevitably transmitted to U.S. servers.

Additionally, ChatGPT utilizes data from user interactions to further refine and train the AI, which could involve sensitive information. The implications for privacy and consent under such a framework are complex and have sparked current debates and legal actions.

For instance, after worries about data and youth protection surfaced, Italy imposed a ban on ChatGPT in March 2023, which remained in effect until OpenAI introduced improved data protection measures in April 2023 [7]. This case highlights the ongoing discussions about the necessity of similar protective measures in Germany and at the EU level, where regulators are contemplating how best to balance technological innovation with the fundamental right to data privacy.

Since the model can learn from previous interactions, researchers face a potential risk. Sharing preliminary or unpublished research data with the model could lead to unintended disclosure.

4 ChatGPT & Writing Journal Papers

ChatGPT can be utilized in the scientific environment in different ways:

- ⤴ Analysis and summarization of literature: ChatGPT can swiftly read and summarize key points from scientific papers, aiding researchers in literature review.
- ⤴ Generating hypotheses: based on input data or research queries, ChatGPT can suggest potential hypotheses or research directions.
- ⤴ Interactive education: ChatGPT can serve as a tool for educating students or new members of the research team, providing real-time answers to questions.
- ⤴ Assistance in writing: the model can help in crafting clear and coherent paragraphs, suggest improvements, or even generate draft texts based on provided information.

Different publishers have already accepted internal rules about the usage of ChatGPT for scientific work and writing scientific articles. For example, Springer accepts no LLM tool as a credited author on a research paper. If an author uses such models, one has to mention this in the "Methods" or "Statements and Declarations" sections of the manuscript [8]. The APA style of citing suggests that you put the full text of long responses from ChatGPT in an appendix of your paper since the results of a ChatGPT chat are not retrievable by others [9]. Scribbr created guidelines for different types of ChatGPT usages [10]. The summary is in Figure 3.

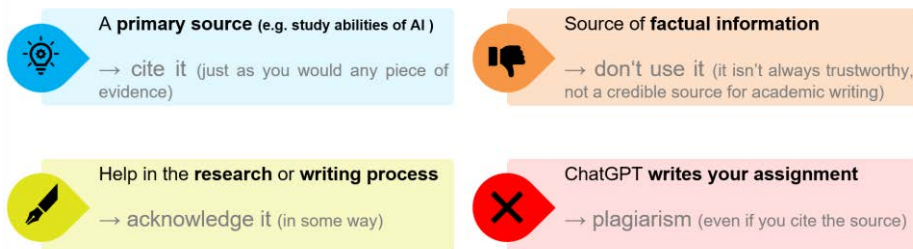


Figure 3: Is it allowed to use ChatGPT for scientific work?

5 Discussion and conclusion

In the field of science, Large Language Models have proven to be exceptionally useful tools. ChatGPT has been at the forefront of this transformation, fundamentally changing the usability of computational assistance in research. As the academic sphere continues to adapt to these rapid changes, it becomes imperative for researchers to stay informed

about the guidelines and policies set forth by individual universities and publishers. Operating according to these rules ensures that the integration of ChatGPT into scientific work remains ethical and productive.

Funding

This research was funded by the Slovenian Research Agency (Javna agencija za raziskovalno dejavnost Republike Slovenije—ARIS), research project grant number ARRS-RPROJ-JR-PRIJAVA/2023/1233.

6 References

- [1] M. Eppler et al. Awareness and Use of ChatGPT and Large Language Models: A Prospective Cross-sectional Global Survey in Urology, *European Urology*, Nov. 2023, doi: 10.1016/j.eururo.2023.10.014.
- [2] Generative pre-trained transformer, Wikipedia. https://en.wikipedia.org/w/index.php?title=Generative_pre-trained_transformer&oldid=1185644518, accessed: Oct. 10, 2023.
- [3] S. Rice; S. R. Crouse; S. R. Winter; C. Rice. The advantages and limitations of using ChatGPT to enhance technological research. *Technology in Society*, vol. 76, p. 102426, Mar. 2024, doi: 10.1016/j.techsoc.2023.102426.
- [4] Demystifying the Architecture of ChatGPT: A Deep Dive | LinkedIn. <https://www.linkedin.com/pulse/demystifying-architecture-chatgpt-deep-dive-vijayarajan-a/>, accessed: Oct. 08, 2023.
- [5] A. Pfau; C. Polio; Y. Xu. Exploring the potential of ChatGPT in assessing L2 writing accuracy for research purposes. *Research Methods in Applied Linguistics*, vol. 2, no. 3, p. 100083, Dec. 2023, doi: 10.1016/j.rmal.2023.100083.
- [6] M. Žmihorski. The hallucinating chatbot “ChatGPT” poorly estimates real bird commonness. *Biological Conservation*, vol. 288, p. 110371, Dec. 2023, doi: 10.1016/j.biocon.2023.110371.
- [7] A. Satariano. ChatGPT Is Banned in Italy Over Privacy Concerns. *The New York Times*, Mar. 31, 2023. <https://www.nytimes.com/2023/03/31/technology/chatgpt-italy-ban.html>, accessed: Okt. 12, 2023.
- [8] Artificial Intelligence (AI) | Springer — International Publisher. <https://www.springer.com/gp/editorial-policies/artificial-intelligence--ai-/25428500>, accessed: Oct. 16, 2023.
- [9] How to cite ChatGPT, <https://apastyle.apa.org>. <https://apastyle.apa.org/blog/how-to-cite-chatgpt>, accessed: Oct. 17, 2023.
- [10] J. Caulfield. ChatGPT Citations | Formats & Examples. Scribbr. <https://www.scribbr.com/ai-tools/chatgpt-citations/>, accessed: Oct. 02, 2023.

Identification of the strengths and weaknesses of outdoor STEM learning using LLM

Biljana Mileva Boshkoska, Katarina Rojko
Faculty of Information Studies
Ljubljanska cesta 31A, 8000 Novo mesto, Slovenia
{biljana.mileva, katarina.rojko}@fis.unm.si

Abstract: *The Learn GREEN - Outdoor STEM project, aimed at promoting outdoor STEM education through authentic tasks, physical activities, usage of mobile devices, and teamwork, has shown great potential in enhancing students' learning experiences. This paper provides qualitative analysis (using BERTopic) of students' views after testing the learning outdoor trails developed within the project on the major strengths and weaknesses of this type of learning. The analysis reveals that BERTopic may be used for qualitative analysis of even small datasets. Specifically, it reveals two topics for both strengths and weaknesses which most significantly influenced learners. Regarding strengths the nature and collaboration-related motives were on the front on one hand and culture-related strengths on the other hand, while regarding weaknesses general everyday imperfections and time-related requirements.*

Key Words: *outdoor learning, topic modelling, BERTopic*

1 Introduction

Teaching and learning outside the classroom can be a powerful tool to enrich students' educational experiences. The *Learn GREEN - Outdoor STEM (GREEN&STEM)* project seeks to harness this potential by incorporating real-life experiences and hands-on learning into STEM education. The project focuses on providing students with engaging activities that promote teamwork, problem-solving, and an understanding of the environment's importance in STEM subjects. STEM fields were chosen as, e.g., according to Li et al. (2019), STEM education is positioned to provide diverse opportunities to facilitate students' learning through design and develop their design thinking, while design thinking as a model of thinking is important for every student to develop and have in the twenty-first century.

This research aims to evaluate students' views on the major strengths and weaknesses of this type of learning, including green trail paths along with mobile devices. We used large language models (LLM), in particular BERTopic, for the analysis. Thus, our research question is which are the main topics that BERTopic defines using students' answers on the major strengths and weaknesses of experienced outdoor learning within the project?

2 Literature overview

2.1 STEM, green trails

There are multiple interpretations of what STEM (science, technology, engineering, and mathematics) education is, and these interpretations usually involve the integrated appearance of the four disciplines that make up the acronym (Martín-Páez et al., 2019). Based on various definitions, Bybee (2013) employs STEM education as (i) it involves an application that relates at least two of the science, technology, mathematics and engineering fields, (ii) these fields are brought together in a context based on real-life problems, and (iii) it helps to teach students the subject-matters or enriches their learning.

Identifying the goals and the content are noted as two critical steps in the design of STEM education programs. Building the programs on students' early interests and experiences and engaging them in the practices of STEM education are noted as crucial factors in developing and sustaining their motivation and engagement with STEM education (National Research Council, 2011 in Baran et al., 2016). For this reason, the Learn GREEN - Outdoor STEM project was focused on students' early interests, i.e., on younger students. In addition, the project also aimed to increase engagement by developing outdoor STEM trails, due to the identified (e.g., James and Williams, 2017; Martín-Páez et al., 2019) benefits and effectiveness of outdoor learning.

2.2 Large language models

Large language models (LLM) are designed to process and generate human-like text, making them exceptionally versatile in understanding and generating human language. They are typically built on deep learning architectures and are trained on massive datasets, which allow them to learn patterns, grammar, and context from the text. BERT (Bidirectional Encoder Representations from Transformers) is a specific instance of LLM. It is developed by Google and is designed for tasks like text classification, natural language inference and text summarization. BERTopic uses BERT to learn representations of documents and then uses these representations to cluster documents into topics (Grootendorst, 2022). It is designed to produce interpretable topics, with clear and distinct descriptions. In this paper we use BERTopic for analysis of qualitative answers from students obtained during the project evaluation phase.

2 Methodology

To assess the impact of the Learn GREEN - Outdoor STEM project, data was collected from participating students in various educational settings. The project's activities were implemented in the period from February 2022 to June 2023, and both quantitative and qualitative data were gathered. After each project activity related to learning and training of students, project surveys were used to measure students' satisfaction with the implemented projects' activities. To process the students' answers obtained in the questionnaires, we used BERTopic, which leverages BERT's embeddings and c-TF-IDF to create clusters and easily interpretable topics. TF-IDF is a statistical measure that evaluates how relevant a word is to a document in a collection of documents. It takes into account how frequently a term appears in a document (TF) relative to how often it appears across all documents (IDF). c-TF-IDF is a slightly adjusted version which

calculates the TF-IDF for only a specific set (or "class") of documents. When documents are transformed using BERT's embeddings and then clustered based on those embeddings, each cluster of documents forms a "class." c-TF-IDF will then highlight words that are frequent in this class of documents but not so frequent in the entire corpus. We have also used n-grams. Ngrams are a sequence of n words. For example, a unigram is a single word, a bigram is a sequence of two words, and a trigram is a sequence of three words. For instance: a 1-gram is a single word (e.g., "stem"); a 2-gram is a sequence of two words (e.g., "lerning stem"). In terms of usage in BERTopic, when calculating c-TF-IDF, the model can be set up to consider n-grams in addition to unigrams. This helps to capture more context and to obtain more interpretable topics.

3 Results

We obtained 95 questionnaires from students who validated the results of the Green&STEM project. We required qualitative answers to the following question: *“What are the major strengths and weaknesses of this type of learning?”*.

We wanted to be able to gain an understanding of the answers of the students, without compromising the analysis with subjective thinking. Using BERTopic, we modelled the answers and obtained the two topics presented in Figure 1. A closer examination of the bag-of-words comprising both topics reveals that the left one regards the “negative” emotions and the right plots the “positive” ones.

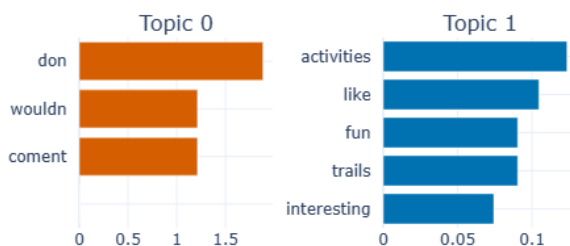


Figure 1 Topic modelling of the questionnaires' answers on *“What are the major strengths and weaknesses of this type of learning?”*

For example, selecting a representative from Topic 0 is *“No, I wouldn't”* refereeing to having no other comments, and a representative from Topic 1 is *“Interesting place to search shapes and math in practice, exploring river and nature, fountains etc,...”*.

Given the small number of documents, altogether 95, BERTopic did not provide more than 2 topics. To alleviate this issue, we divided the answers into strengths and weaknesses and analysed them separately.

The analysis of the answers regarding the *strengths* of the project activity resulted in the identification of two topics presented in Figure 2. A closer examination of the obtained words in Topic 0 reveals that a group of students focused on nature and collaboration and the representative comments are *“Learning in nature and in the city are major strengths. Also working in groups.”*, or *“The major strengths of this type of learning is teamwork and friends relationship”*. At the same time, Topic 1 resulted from the group

of students focused on culture developed during such activities, resulting from representative answers such as “*Doing activities with people from other countries*”, and “*I liked getting to know cultures and people from other countries*”.

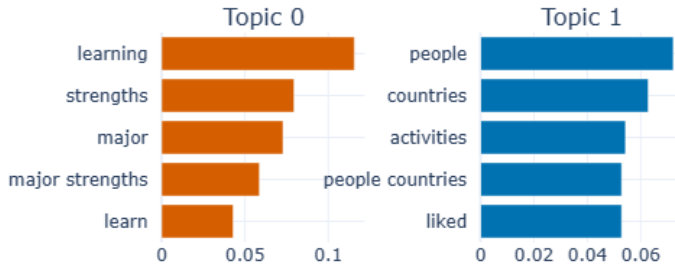


Figure 2 Topic word scores for the "Strengths" of the project activities

Finally, Figure 3 shows the obtained topics when analysing the “weaknesses” stated by the students.

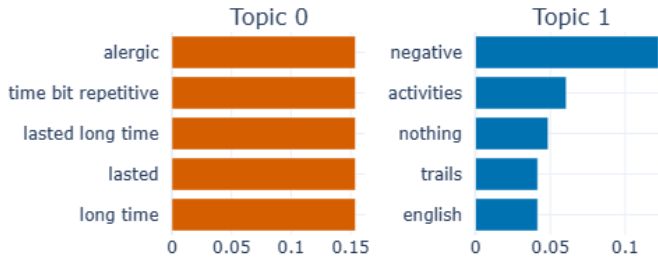


Figure 3 Topic word scores for the "Weaknesses" of the project activities

Topic 0 refers to students being sensitive to general everyday imperfections such as '*if someone is allergic he could have problems*', and '*lasted a long time, it was all a bit repetitive*'. In Topic 1, the weaknesses focused on the time-related requirements to perform activities. The representative answers in this topic are '*having more free time to spend with them*' ('them' most likely referring to the outdoor activities :), '*doing activities in English all the time*', and '*more free time*'.

4 Conclusion

The Learn GREEN - Outdoor STEM project, focused on STEM education, due to its importance in the 21st century (Li et al., 2019), has demonstrated its potential to enhance STEM education by providing students with practical, hands-on learning experiences. The GREEN&STEM toolkit, mobile application, and Outdoor STEM educator guide have played crucial roles in promoting students' interest in STEM subjects and fostering their academic and personal growth. To obtain feedback from the involved students in the project we used questionnaires after each of the LTTA (learning, training and teaching activity).

While analysis of questionnaires using traditional methods is still sound, BERTopic allowed us to have an in-depth understanding of the qualitative answers provided by the

students. Using BERTopic, we identified the main topics that prevailed in the students' answers regarding major strengths and weaknesses of the experienced outdoor learning within the project, indicating nature and collaboration-related motives on one hand and culture-related strengths on the other hand. Weaknesses resulted from general everyday imperfections and from the time-related requirements to perform activities. The analysis answers our research question, and therefore we would recommend such analysis also for future project activities where qualitative text analysis is required. Nonetheless, we are aware of the limitation of our presented qualitative analysis, mainly resulting from a small number of surveys analysed.

In conclusion, we claim that the Learn GREEN - Outdoor STEM project is a valuable initiative that holds immense potential for revolutionising traditional classroom-based learning approaches. Yet further expansion, sustained funding, and ongoing teacher training are essential to ensuring the project's continued success and its contribution to advancing STEM education in schools and beyond.

5 Acknowledgements

This work has been prepared in frames of the ERASMUS+ project **GREEN – Outdoor STEM** financed by the European Union under the **Call 2022 KA220-HED: Cooperation partnerships in tertiary education**.

6 References

- [1] Baran, E; Bilici, S. C; Mesutoglu, C; Ocak, C. Moving STEM Beyond Schools: Students' Perceptions about an Out-of-School STEM Education Program. *International journal of education in mathematics science and technology*, 4(1): 9-19. 2016.
- [2] Bybee, R.W. *The Case for STEM Education: Challenges and Opportunities*. National Science Teachers Association (NSTA) Press. 2013.
- [3] Grootendorst, M. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv*, 2022.
- [4] Green&Stem Learn Gren – outdoor stem. <https://stem.fis.unm.si>. Last accessed: September 23rd 2023.
- [5] James, J. K.; Williams, T. School-Based Experiential Outdoor Education: A Neglected Necessity. *Journal of Experiential Education*, 40(1): 58-71, 2017.
- [6] Li, Y; Schoenfeld, A. H; diSessa, A. A; Graesser, A.C; Benson, L.C; English, L. D; Duschl, R. A. Design and Design Thinking in STEM Education. *Journal for STEM Education Research*, 2: 93-104, 2019.
- [7] Mann J; Gray T; Truong S; Sahlberg P; Bentsen P; Passy R; Ho S; Ward K; Cowper R. A Systematic Review Protocol to Identify the Key Benefits and Efficacy of

Nature-Based Learning in Outdoor Educational Settings. *International Journal of Environmental Research and Public Health*, 18(3): 1199, 2012.

[8] Martín-Páez, T; Aguilera, D; Perales-Palacios, F. J; Vílchez-González, J. M. What are we talking about when we talk about STEM education? A review of literature. *Science Education*, 103: 799-822, 2019.

Potential of lexicon-based bug severity classification

Fatima Aziz

Jožef Stefan International Postgraduate School
Jamova cesta 39, 1000 Ljubljana, Slovenia
fatima.aziz6655@gmail.com

Martin Žnidaršič

Jožef Stefan Institute, Jožef Stefan International Postgraduate School
Jamova cesta 39, 1000 Ljubljana, Slovenia
martin.znidarsic@ijs.si

Abstract. *The severity of a reported bug is a critical factor in deciding how soon it needs to be fixed. Bug severity classification is commonly automated and approached with machine learning methods, some of which employ lexicons. Interestingly, there does not seem to be a dedicated lexicon available for this task and in this paper we report on a study that is aimed at assessing potential usefulness of such a resource. For this purpose, we propose a lexicon development approach and assess the performance of pure lexicon classification in comparison with standard machine learning approaches on the problem of bug severity classification. Our results indicate that it would be sensible to develop a specialized lexicon for this problem domain.*

Keywords. Lexicon-Based Classification, NLP, Machine Learning, Software Bug Report, Software Bug Severity, Software Maintenance

1 Introduction

Software bug is a fault that prevents a software system to operate as intended. Bug reports are usually submitted to bug-tracking systems and can range from minor issues that have little impact on the overall functionality, to critical issues that can cause the software to crash. To speed-up the process of bug triaging, these reports need to be accurately classified and assigned to appropriate developers. Since software testing and bug triaging is time-consuming, proper management of bugs is an important factor in minimizing the cost of the software development process.

A general bug report contains various attributes such as Bug ID, Status, Submission Date, Summary, Severity, Description, and Priority [11]. Based on the information in the summary or description in the bug reports, the bug reporters usually classify the bug according to its severity [5]. The process of assigning the bug severity level is time-consuming and error-prone [4]. According to its severity, a bug is prioritized and assigned to the related developer [8]. One way of assigning severity to a bug is through automated classification of the text description of a bug report. Several studies have proposed solutions for automated bug severity prediction and are mostly focused on machine learning

approaches [14], [7], deep learning methods [10],[2] and the use of word-embeddings [1]. Interestingly, few studies have also employed lexicons or word lists [3, 13, 2] either as sources of features for machine learning or as baseline approaches for comparison. However, in almost all the cases these are not lexicons that would be tailor-made for the problem of bug severity classification, but lexicons that were designed for sentiment analysis. Classification approaches based on lexicons are usually very resource friendly and can be employed in situations when the labeled data is scarce. They are usually outperformed by machine learning approaches, but remain commonly used in some domains, for example in finance [9] and serve as baseline approaches to compare against.

In the field of bug severity classification, one of rare works that provides insight into performance of such approaches is the work of Barrah et al. [2], but the lexicon used there is a general sentiment one, like in many similar studies on bug severity. There seems to be a simple reason for this, namely, according to our literature and research review, there is no lexicon available yet for bug severity classification. An attempt in this direction was the SentiStrength-SE [6], which is an adaptation of a general sentiment analysis lexicon to software development although, focusing only on sentiment. The only specific lexicon of this kind that we could find was reported to be developed for an experimental comparison [12]. However, the lexicon that was used is not provided, there are no insights about its standalone performance, as it was used as feature filter for machine learning approaches, and the study leaves a lot to be desired in terms of methodology and experimental setup.

Our goal in this problem domain is to inspect whether lexicon based approaches can be used as relevant baseline solutions for the development of a robust general lexicon of this kind. Motivation for this is twofold, on the one hand, lexicons serve as baseline solutions in situations when either data or computational resources are low, on the other hand, good general lexicons can serve as sources of seed words for other approaches, including for approaches that make use of the state-of-the-art machine learning and word-embedding based techniques. This paper is focused on the initial stages of the planned research — on assessment of potential usefulness of lexicon approaches in the context of the problem of bug severity prediction. For this purpose, we develop tailor made lexicons for two datasets of diverse characteristics and compare their performance to performance of classic machine learning methods. The rest of this paper is structured as follows: Section 2 introduces the data and data preprocessing used. Section 3 explains our proposed lexicon-based approach to predict software bug severity. Section 4 outlines our experimental setting, results and discussion of empirical assessment. Section 5 concludes this paper and suggests future work.

2 Data

The dataset used in this study was taken from Bugzilla¹ and Eclipse² online platforms. The bug report provides information on Bug Id, Product, Component, Assignee, Status, Resolution, Summary, Changed On, Priority, Severity, and Type. The targeted column in this work was the Severity column whereas, the Summary column was used as a feature having an unstructured textual description for an explanation of the bug. The severity levels in the target column are generally set as a blocker, critical, major, normal, trivial, and minor or in some cases, the severity level is assigned as S1, S2, S3, and S4 as explained in Table 1. The Type column contains information on the topics such as defects,

¹<https://bugzilla.mozilla.org/>

²<https://bugs.eclipse.org/>

enhancement, or tasks. A defect here refers to a fault in the system whereas, enhancement represents a new feature in the system. Task on the contrary defines any refactoring, removing, replacement, enabling, or disabling of functionality. A general error report holds also various other information such as Bug ID, Status, Submission Date, etc., but the most critical attribute of the bug report is the severity level of a bug, on the basis of which one decides how rapidly it should be resolved.

Table 1: Bug severity levels.

Sn	Severity	Description
1	Blocker(S1)	(Catastrophic) Blocks development/testing, may impact more than 25% of users, causes data loss, likely dot release driver, and no workaround available
2	Critical (S2)	(Serious) Major functionality/product severely impaired or a high impact issue and a satisfactory workaround does not exist
3	Major (S2)	(Serious) Major functionality/product severely impaired or a high impact issue and a satisfactory workaround does not exist
4	Normal(S3)	(Normal) Blocks non-critical functionality and a workaround exists
5	Minor (S4)	(Small/Trivial) minor significance, cosmetic issues, low or no impact to users
6	Trivial (S4)	(Small/Trivial) minor significance, cosmetic issues, low or no impact to users

Following the extraction of the bug reports, the data from both software projects were combined. The dataset used in this study had 51154 data items. Data that were tagged as N/A were also excluded from the dataset, since the default severity level in Bugzilla is "-" which means the bug with severe levels "-" are not assigned with any severity levels. In order to make it a binary classification problem, the severity level was categorized into Severe and Non-Severe, where blocker, critical, and major levels of severity were categorized in a Severe category while, normal, minor, and trivial levels fell into the Non-Severe category. After excluding the irrelevant data items, the dataset is comprised of 33924 non-severe bugs and 8214 severe bugs.

The Summary column was further pre-processed to transform it into structured data. The Summary attribute was therefore pre-processed using NLP techniques including tokenization, removal of stop words, and stemming. Tokenization is a process where the text or sentences in the summary attribute from the bug report are broken down into tokens. In the next step, the stop words were removed which are also known as common words such as "the", "a", "an", and "in" etc. Lastly, the stemming technique was applied, which is a normalization technique where the tokenized words are converted into shortened base words in order to consider word variations equally.

3 Lexicon-based approach

Our lexicon for experimentation was not designed manually, but was created from the words in the Summary attribute in the software bug reports. For this purpose, first the training dataset was turned into a corpus of sentences filtered for punctuation and stop words. Next, the sentences of the corpus were split into individual words. A method was then created that picks an individual word from the list and traverses through the Summary column of the training dataset to count the number of times each individual word appears in the Severe and Non-Severe categories. This resulted into two wordlists consisting of counts for appearance of individual words in the Severe and Non-Severe categories. In the next step, the ratio of the appearance of individual words in Severe and Non-Severe

categories was calculated as shown in equation 1.

$$R_s = C_s / (C_s + C_{ns}) \quad (1)$$

where R_s stands for the ratio of appearance in the severe category, C_s denotes the count of appearances in the severe category and C_{ns} the count of appearances in the non-severe category. Similarly, the ratio for words that appeared in the non-severe category was calculated as defined in equation 2

$$R_{ns} = C_{ns} / (C_s + C_{ns}) \quad (2)$$

where R_{ns} represents the ratio of appearance in the non-severe category.

Lexicons for Severe and Non-Severe bugs are then defined based on a given threshold, for Severe in line with equation 3 and for Non-severe in line with equation 4

$$R_s \geq T_s \quad (3)$$

$$R_{ns} \geq T_{ns} \quad (4)$$

where R_s and R_{ns} are the severity and non-severity ratios of individual word from the created wordlist whereas, T_s and T_{ns} refer to the threshold, which was defined between 0.1 and 1.0 in our experiments and assessed on a validation part of the learning data (considering F1 score). The thresholds that yielded the maximum score were selected to produce an output lexicon, which was then used on test data in the experiments.

The classification approach for the lexicon-based classifier that we used was very basic and straightforward — a new bug report was classified into the category that corresponded to the lexicon with which the intersection of the words was the largest. The ratios were therefore not taken into consideration as weights, but used together with the thresholds only to limit the amount of words considered.

4 Empirical Evaluation

4.1 Experimental Setting

For the purposes of this study, we empirically evaluated our lexicon learning and classification approach and three classic machine learning methods in four experiment scenarios (dataset subsets). In each experiment the dataset was always randomly split into training (20%), validation (20%), and testing (60 %) datasets. The lexicon-based classifier was evaluated with tests at different severe and non severe thresholds. The same dataset-split was then used to train the machine learning (ML) classifiers. ML classifiers having a contrasting set of features were chosen for this purpose: Logistic Regression, Naïve Bayes multinomial, and SVM where, logistic regression offers low computational costs whereas, the assumption of the distribution of data is not required for Naïve Bayes. Similarly, SVM classifier can handle linear and nonlinear datasets and is also commonly used for text classification. Each experiment was repeated ten times, with a diverse random split each time.

A total of four experiments were performed using different combinations of datasets as shown in Table 2. In experiment 1, we evaluated the performance of the classifiers on Eclipse subset only since Eclipse’s bug report is written by the developers and thus is technical in nature. Experiment 2 used only the Mozilla Firefox related subset while a

Table 2: Overview of the experiments.

	Dataset	Normal
Experiment 1	Eclipse	Included
Experiment 2	Firefox	Included
Experiment 3	Eclipse&Firefox	Included
Experiment 4	Eclipse&Firefox	Excluded

combined Eclipse and Firefox datasets were utilized in Experiment 3. For this purpose, both datasets were merged in the preprocessing step so that both the learning and the prediction datasets have instances from both of the datasets. In Experiment 4, the combined data of Eclipse and Firefox was used similar to that done in Experiment 3 although, the bugs with the Status equal to Normal were excluded from the final combined dataset. This results in a big difference in the distribution of data between Severe and Non-Severe categories.

The learning step includes machine learned model and lexicon construction, so a separate lexicon was also built for each dataset. This way, we had software product specific lexicons in Experiments 1 and 2, and generic (albeit based on only two products) ones in case of Experiments 3 and 4. The performance of the classifiers was measured primarily by the F1 score (of the Severity category), as the data distribution was very uneven in most experiments, but we gathered the confusion matrix values as well as CPU processing time for both the learning and the classification parts of the processes.

4.2 Results and Discussion

Results of our experimental assessment are presented in Tables 3—6. According to the F1 score for the Severe category—as the most relevant performance measure in our setting—the lexicon approach looks useful, as it is better in three experiment variations and very comparable to ML approaches in the fourth. However, there are at least two cautionary notes to consider in view of these results. First, as explained in Sections 3 and 4 above, we employed a learning approach to learn and tune the lexicon to the training data, so it is not a classic lexicon-based classification approach, but currently machine learned to some extent. Second, the SVM approach is very sensitive to some parameters (we used defaults) and a much more informative assessment of its performance can be gained if the parameters are tuned with the validation set, like in the case of the lexicon approach. The SVM tuning was left out, as it took prohibitive amounts of time. However, we have the F1 score results of SVM with parameter fitting for the Experiment1 and it is 0.3314, considerably better than with defaults, but still comparable to other approaches.

Running time of the learning and classification parts of the approaches are also reported in the result tables (in seconds). As mentioned, the SVM, even without parameter tuning, took much more time than other methods. The learning part of our lexicon approach is also time consuming, but as learning is not necessary given a fixed static lexicon, which is our end goal, the more important is the classification time. Interestingly, even during classification, some classic machine learning approaches are faster than our (though not time optimized) use of the lexicon. This is a relevant result to consider, as besides usefulness in situations with low (or no) training data, lexicons are commonly argued also as fast and not computationally intensive.

Table 3: Results of Experiment1 as averages of 10 iterations.

	F1	F1mean	TP	FP	TN	FN	tl	tc
lexicon	0.4423	0.5356	870.4	1760.6	2413.2	434.8	344.2	2.7
SVM	0.2625	0.56037	232	229.8	3944	1073.2	1696.2	189.5
multiNB	0.3528	0.6033	356.4	358.7	3815.1	948.8	1.9	0.1
logreg	0.2954	0.5773	269.6	250.2	3923.6	1035.6	10.7	0.3

Table 4: Results of Experiment2 as averages of 10 iterations.

	F1	F1mean	TP	FP	TN	FN	tl	tc
lexicon	0.3764	0.5148	205.7	554.1	2061.9	127.3	190.8	1.4
SVM	0.2831	0.6104	266.3	71.2	66.7	2544.8	109.3	27.9
multiNB	0.3231	0.6269	88.6	125.9	2490.1	244.4	0.6	0.06
logreg	0.2958	0.6186	67.1	52.8	2563.2	265.9	2.9	0.05

Table 5: Results of Experiment3 as averages of 10 iterations.

	F1	F1mean	TP	FP	TN	FN	tl	tc
lexicon	0.4126	0.5283	1253.8	3190.8	3602.1	381.3	635.6	4.4
SVM	0.2261	0.5581	237.4	224.9	6568	1397.7	5791.2	406.9
multiNB	0.3688	0.6227	508.1	611.3	6181.6	1127	4.7	0.2
logreg	0.2791	0.5849	309.8	273.2	6519.7	1325.3	22.3	0.3

Table 6: Results of Experiment4 as averages of 10 iterations.

	F1	F1mean	TP	FP	TN	FN	tl	tc
lexicon	0.8857	0.7545	1573.2	347.1	277.7	59	129.3	1.0
SVM	0.8831	0.7637	1500.8	265.5	359.3	131.4	80.9	21.0
multiNB	0.8914	0.7791	1517.7	255.1	369.7	114.5	0.4	0.05
logreg	0.8909	0.7691	1536.1	279.8	345	96.19	1.5	0.03

5 Conclusion

Bug severity classification is a relevant real-world problem, which is commonly approached via text classification. Interestingly, this problem domain does not yet have available lexicons, which represent common baseline solutions in similar domains. In the work presented in this paper, we assessed the potential of such an approach in domain of bug severity classification. According to the results of our study, it makes sense to develop a specialized lexicon for this domain. Encouraged by these results, we will now focus on gathering additional data sources and development of a fixed static lexicon.

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).

References

- [1] Rashmi Agrawal and Rinkaj Goyal. “Developing bug severity prediction models using word2vec”. In: *International Journal of Cognitive Computing in Engineering* 2 (2021), pp. 104–115.
- [2] Aladdin Baarah et al. “Sentiment-based machine learning and lexicon-based approaches for predicting the severity of bug reports”. In: *Journal of Theoretical and Applied Information Technology* 99.6 (2021).
- [3] Fabio Calefato et al. “Sentiment Polarity Detection for Software Development”. In: *Empirical Software Engineering* 23 (2017), pp. 1352–1382.
- [4] Anh-Hien Dao and Cheng-Zen Yang. “Severity prediction for bug reports using multi-aspect features: a deep learning approach”. In: *Mathematics* 9.14 (2021), p. 1644.
- [5] Luiz Alberto Ferreira Gomes, Ricardo da Silva Torres, and Mario Lúcio Côrtes. “Bug report severity level prediction in open source software: A survey and research opportunities”. In: *Information and software technology* 115 (2019), pp. 58–78.
- [6] Md Rakibul Islam and Minhaz F. Zibran. “SentiStrength-SE: Exploiting domain specificity for improved sentiment analysis in software engineering text”. In: *Journal of Systems and Software* 145 (2018), pp. 125–146. ISSN: 0164-1212. DOI: <https://doi.org/10.1016/j.jss.2018.08.030>.
- [7] Ashima Kukkar et al. “A novel deep-learning-based bug severity classification technique using convolutional neural networks and random forest with boosting”. In: *Sensors* 19.13 (2019), p. 2964.
- [8] Ahmed Lamkanfi et al. “Predicting the severity of a reported bug”. In: *2010 7th IEEE working conference on mining software repositories (MSR 2010)*. IEEE, 2010, pp. 1–10.
- [9] Tim Loughran and Bill McDonald. “Textual analysis in accounting and finance: A survey”. In: *Journal of Accounting Research* 54.4 (2016), pp. 1187–1230.

- [10] Waheed Yousuf Ramay et al. “Deep neural network-based severity prediction of bug reports”. In: *IEEE Access* 7 (2019), pp. 46846–46857.
- [11] Korosh K Sabor et al. “Predicting bug report fields using stack traces and categorical attributes”. In: *Proceedings of the 29th Annual International Conference on Computer Science and Software Engineering*. 2019, pp. 224–233.
- [12] Gitika Sharma, Sumit Sharma, and Shruti Gujral. “A novel way of assessing software bug severity using dictionary of critical terms”. In: *Procedia Computer Science* 70 (2015), pp. 632–639.
- [13] Geunseok Yang et al. “Analyzing emotion words to predict severity of software bugs: A case study of open source projects”. In: *Proceedings of the Symposium on Applied Computing*. 2017, pp. 1280–1287.
- [14] Tao Zhang et al. “Towards more accurate severity prediction and fixer recommendation of software bugs”. In: *Journal of Systems and Software* 117 (2016), pp. 166–184.

Addressing challenges in higher education study programs: The case of the RRP pilot project "Advanced Computer Skills"

Nuša Erman, Katarina Rojko
Faculty of Information Studies
Ljubljanska cesta 31A, 8000 Novo mesto, Slovenia
{nusa.erman, katarina.rojko}@fis.unm.si

Abstract: *The appeal to higher education institutions to be more flexible in responding to labour market needs is also one of the highlights of the EU's Resilience Recovery Plan. Under this plan, various projects are funded, including pilot projects that offer shorter study programs. However, such free-of-charge shorter study programs also face similar challenges as full-time study programs. Among these challenges, we have observed that motivation, lack of prior knowledge, and lack of time management skills are at the forefront among learners, while drop-outs and questions about the introduction of compulsory attendance and hybrid delivery method are at the top of the list among education providers. We, therefore, decided to use ChatGPT to provide us with suggestions on how to address these challenges and found that the recommendations were mostly appropriate for us, many of which we have even already implemented to address these challenges, while we have also obtained some valuable new suggestions.*

Key Words: *higher education, challenges in education, ChatGPT.*

1 Introduction

The ever-changing needs of the labour market, driven in particular by the rapid development of information technologies, also require employees to continuously upgrade their skills and knowledge. This has highlighted the need to adapt higher education to make it more flexible and responsive to labour market needs, which is also the focus of the Recovery and Resilience Plan (RRP) pilot project "Advanced Computer Skills" at the Faculty of Information Studies in Novo mesto¹. The educational process is designed as a half-year higher education program, in which learners take courses one after the other within two study fields, starting from the basics and gradually progressing to more advanced content.

During the implementation of the project so far and on the basis of evaluation results, certain challenges appeared on 2 levels - on the side of the learners and on the side of the faculty, which we want to address appropriately in the continuation of the project.

Thus, our research questions are divided into 2 subsets, the first three (Q1 - Q3) relevant to learners and the next three (Q4 - Q6) to education providers:

Q1: How to increase motivation to learn and keep it at an adequate level throughout a half year of higher education courses?

Q2: Is a lack of prior knowledge crucial for success in and completion of higher education courses?

Q3: How to manage your time so that you have enough time to complete a six-month of higher education courses alongside work, family and friends?

Q4: How to prevent dropouts from free half-year higher education programs?

Q5: Does compulsory attendance make sense in free half-year higher education programs and why yes or no?

Q5.1: What proportion of compulsory attendance is recommended in free half-year higher education programs?

Q6: Is hybrid delivery of lectures and tutorials appropriate for free half-year higher education courses?

Linked to research questions, the six topics are presented in the brief literature overview regarding the observed challenges in education, followed by the presentation of methodology, research results, discussion and conclusion.

2 Literature overview

¹ Pilot project "Advanced Computer Skills" at the Faculty of Information Studies in Novo mesto is financed by the Slovenian Ministry of Higher Education, Science and Innovation, and the European Union - NextGenerationEU. Its implementation period is between 1st July 2022 and 31st December 2025.

The motivation to learn is one of the crucial factors for the students to be engaged in the learning process, which leads to increased levels of both, success, and achievement (Pantzios et al., 2022; Ferrer et al., 2022). It is generally divided into intrinsic or internal motivation, and extrinsic or external motivation (Ryan and Deci, 2020). It can be assumed that one of the important tasks of educators is to increase learners' motivation, i.e., to try to influence learners' intrinsic motivation through the means of extrinsic motivation.

Since learners enrolling in any form of higher education (HE) come from different backgrounds, their level of prior knowledge varies. Theories, such as expertise development theory (e.g., Van Lehn, 1996), suggest that in the educational process learners with sufficient prior knowledge combined with practical experience would need less guidance as compared to the ones with lower levels of prior knowledge.

However, an important factor influencing the learners' success and achievement is also their ability to manage time, which is even more important for learners with different life roles. Studies (e.g., Astudillo et al., 2019) suggest that most working students have problems focusing on their studies, whereas a similar situation is also found among learners whose role includes parenting (e.g., Sicam et al., 2021).

The absence of the above-mentioned motivation, along with different backgrounds and different life roles can quickly lead to dropouts. Although the word 'dropout' in higher education (HE) refers to various notions as leaving the course, program or institution, there is absolute consensus that it creates a loss in the economic and social prosperity of both individuals and institutions (Sandeep Kumar et al., 2020). Therefore, educational institutions are also looking for ways to reduce dropout rates, and a possible approach includes the introduction of compulsory attendance. Namely, attendance is usually considered to reflect students' level of engagement with their course and to be critical to student success (Moores, Birdi and Higson, 2019).

To make compulsory attendance less financially and time-consuming for students, overcoming distance through mediated technological practices founded in pedagogical principles, led to the development of a wide range of educational possibilities, including hybrid education, which has evolved alongside social, educational, and technological changes (Jiménez-Cortés and Aires, 2012).

3 Methodology

During the first cycle of the half-year higher educational programs within the RRP pilot project, we successfully implemented 3 surveys: the survey among the learners for each individual course, the survey among the courses' teachers, and the final survey among the learners, after the end of their half-year program.

Other qualitative and quantitative data on the implementation of the educational program was also collected on an ongoing basis, and the compilation of all these data allowed us to identify the main opportunities for improving the next two in continuation cycles of the half-year higher educational programs within the project.

Among others, we decided to examine the observed top challenges with the help of artificial intelligence, by asking ChatGPT to answer our research questions, each related to one challenge. The questions were asked at the end of August 2023, and as project leaders, we then commented on ChatGPT's answers, as summarized in the following section.

4 Results

This section includes a summarized answer given by ChatGPT to each of our six research questions, followed by our comments based on the experiences gained within the project so far.

4.1 Learning motivation

When asked Q1, ChatGPT outlined some strategies on how to increase and maintain motivation throughout a half year of higher education courses:

- *Set clear goals.*
- *Break tasks.*
- *Create a study schedule.*
- *Find meaning in your coursework.*

- Stay *organized*.
- *Reward* yourself for accomplishments.
- Stay curious and *ask questions*.
- Join *study groups* for collaboration.
- *Visualize* success.
- Take *breaks* and practice *self-care*.
- Maintain a *positive mindset*.
- *Seek support* when needed.
- Mix up your *learning methods*.
- Reflect on your *progress*.
- Embrace *failure* as a learning *opportunity*.

Comment by project leaders:

Experience has shown that some learners engaged in the first cycle of the project's half-year educational programs were unlikely to use certain strategies to increase motivation. Although they probably had a *clear goal* when they enrolled in the education, their goals were *not in line* with the purpose of the project's half-year educational programs. Consequently, they *lost motivation* soon after starting the courses and a few learners even developed negative *mindsets*, instead of positive ones.

4.2 Prior knowledge

When asked Q2, ChatGPT answered that a lack of prior knowledge is not crucial for success in higher education courses. While some foundational knowledge can be helpful, higher education is designed to teach what is required. What's more important are qualities like a *willingness to learn*, *critical thinking*, *time management*, and a *strong work ethic*.

Comment by project leaders:

When setting *entry requirements* for the project's half-year educational programs, we followed the current Slovenian legislation, namely the Higher Education Act. As the project's half-year educational programs have been developed in analogy with the first cycle higher education programs of our faculty, we have taken into account the provision of Article 38 of the Act, which states that anyone who has gained a high school (in the duration of 4 years²) diploma may enroll in a first cycle professional higher education program. Moreover, we designed the curricula and the sequence of courses in both project's half-year educational programs so that the courses follow one after the other, starting from the basics and *gradually progressing* to more advanced content, to overcome possible problems due to the lack of prior knowledge.

4.3 Time management

When asked Q3, ChatGPT suggested that to manage your time effectively while balancing higher education, work, family, and friends:

- Set *clear goals*.
- Create a *schedule*.
- Use a *calendar*.
- *Prioritize tasks*.
- Practice *time blocking*.
- Set *realistic expectations* and *avoid overloading*.
- Learn to *say no*.
- *Group* similar tasks.
- Utilize *productivity tools* and *technology*.
- *Communicate* your schedule and boundaries with others.
- *Delegate* or *outsource tasks*.
- *Prioritize self-care*.
- Stay *flexible* and *adapt* to unexpected events.

² Simplified. For the exact provision, see the Higher Education Act (Slovene: ZVis), Article 38.

- Regularly *review* and *adjust* your schedule and commitments.

Comment by project leaders:

As mainly adult learners enrolled in the first cycle of half-year educational programs, we would expect them to have a certain amount of *life experience*, including the *ability to manage time* but it soon became clear that a lot of learners were not aware of the intensity of the education, and therefore didn't set *clear* and *realistic expectations*, and didn't *review* and *adapt* their schedule and commitments.

4.4 Dropouts

When asked Q4, ChatGPT suggested that to prevent dropouts from free half-year higher education programs, institutions can take the following key steps to support students academically, emotionally, and practically:

- *Early Identification and Intervention.*
- *Clear Expectations and Orientation.*
- *Mentoring and Peer Support.*
- *Personalized Academic Support.*
- *Financial and Practical Assistance.*
- *Regular Check-ins.*
- *Flexible Learning Options.*
- *Engaging Curriculum.*
- *Supportive Environment.*
- *Continuous Feedback and Improvement.*
- *Celebrate Milestones.*

Comment by project leaders:

Despite we offered *Flexible Learning Options* and set *Clear Expectations and Orientation*, it quickly became obvious that the latter was not understandable to students. Thus, we tried with *Early Identification and Intervention*, but only at a limited level, and this should be improved.

It is possible that the dropout rate would be decreased by *Mentoring and Peer Support* and *Personalized Academic Support*. We could also provide even more *Supportive Environment* and *Celebrate Milestones* with learners, but the proposal for *Financial and Practical Assistance* is not applicable. We also already addressed *Continuous Feedback and Improvement* with surveys and other qualitative and quantitative data collection. According to the project plan, teachers also provided *Engaging Curriculums*, which are to be updated as required each year but could devote more attention to *Regular Check-ins*.

4.5 Compulsory attendance

When asked Q5, ChatGPT didn't take a position but only provided reasons for and against, which are presented in continuation.

Reasons for:

- *It can maximize learning opportunities.*
- *It maintains a structured learning environment.*
- *Interaction and engagement.*
- *It ensures participation.*
- *Promotes accountability and responsibility.*

Reasons against:

- *Higher education students are often adults.*
- *Students have diverse learning styles.*
- *It can lead to wasted time.*

- *Non-traditional students.*
- The *focus* should be on *learning outcomes*.

Comment by project leaders:

Since the learners in the projects' half-year educational programs are supposed to be mainly *adults* and *non-traditional students* with *diverse learning styles*, we didn't want to *waste their time*. We therefore didn't require compulsory attendance, and we focused on *learning outcomes*.

When asking ChatGPT what the recommended proportion of compulsory attendance is (Q5.1), it answered that a common approach might be around 70-80% attendance while allowing for flexibility based on the specific needs of the course and students which depend on several factors:

- *Course structure and content.*
- *Skills and practical training.*
- *Lecture-based courses.*
- *Student demographics.*

Comment by project leaders:

Regarding the factors, we were focused especially on *student demographics*, and therefore we didn't require compulsory attendance.

4.6 Hybrid delivery model

According to ChatGPT the implementation of a hybrid delivery model for free half-year higher education courses (Q6) can be appropriate due to:

- *Flexibility.*
- *Engagement.*

But ChatGPT also listed the following conditions:

- *Technology.*
- *Communication.*
- *Course Design.*
- *Assessment.*
- *Support Services.*
- *Faculty Training.*
- *Continuous Improvement.*

Comment by project leaders:

Since we already offered a hybrid delivery model, based on the direction of the faculty, we were aware of the required conditions regarding *technology*, *communication*, *assessment*, *support services*, and *faculty training*. We provided all the required. Also, we collected feedback to provide *continuous improvement*. But what we see that could be more clearly focused and addressed is the recommendation for *course design*.

5 Discussion

Regarding the research questions relevant to learners, we summarize our findings as follows:

Q1: How to increase motivation to learn and keep it at an adequate level throughout a half year of higher education courses?

ChatGPT provided us with 15 strategies for the learners to stay motivated and be successful in their education, e.g., *setting clear goals*, *creating a study schedule*, *staying organized* etc. However, as a higher education institution, we have little influence on the extent to which learners use these strategies. Nonetheless, it would

probably be easier for learners if we would at least offer them guidance on how to keep their motivation as high as possible.

Q2: Is a lack of prior knowledge crucial for success in and completion of higher education courses?

Following the arguments provided by ChatGPT, the answer is *no*. Although concerns about the prior knowledge needed to participate in the half-year educational programs were frequently raised in learners' feedback, we need to be aware that the half-year educational programs are at the level of first-cycle higher education. This has been followed in the design of the half-year educational programs, in setting the *entry requirements* and, last but not least, in the emphasis on upgrading knowledge, starting from the courses with basic content and *gradually progressing* to more advanced content.

Q3: How to manage your time so that you have enough time to complete a six-month of higher education courses alongside work, family and friends?

According to ChatGPT, to manage time efficiently, one should follow 14 suggestions, mainly concerning the organization of the daily *schedule*. In order to facilitate the balance between work and education, all courses offered were held in the afternoons. Furthermore, it is also necessary for the learners to balance family and friends with their education, compulsory attendance was not required in any (except in one out of 12) of the courses.

With regard to the research questions relevant to providers of education at the faculty, we present the following summary of our findings:

Q4: How to prevent dropouts from free half-year higher education programs?

Out of the 11 key steps to support students academically, emotionally, and practically that ChatGPT recommended we already took 6 of them. As regards to other proposals that are relevant for our case, we believe that also based on the literature review (e.g., Joseph et al., 2021) we should focus most on *Regular Check-ins* and provide rewards for learners' attendance and cooperation.

Q5: Does compulsory attendance make sense in free half-year higher education programs and why yes or no?

ChatGPT listed reasons for and against but didn't take the position. Nonetheless, despite the reasons against compulsory attendance prevailed in our preparation of the first cycle of the project's half-year program, especially due to educating *adults* and *non-traditional students* with *diverse learning styles*, a critical low attendance rate, along with a significant dropout rate led us to consider about the reasons for compulsory attendance for the forthcoming free half-year education programs.

Q5.1: What proportion of compulsory attendance is recommended in free half-year higher education programs?

According to ChatGPT, the recommended proportion of compulsory attendance is around 70-80% but some flexibility based on the specific needs of the course and students should be regarded.

Q6: Is hybrid delivery of lectures and tutorials appropriate for free half-year higher education courses?

ChatGPT answered that the implementation of a hybrid delivery model can be appropriate due to: *Flexibility* and *Engagement*. The same two reasons were also our initial reasons, based on experiences from the hybrid delivery model from the regular study programs at the faculty.

6 Conclusion

Besides these six challenges, we addressed in the paper, we are aware of several others as well. Among them, one of the most frequently mentioned challenges was the fast pace of course delivery, a comment to which we will also pay more attention in future implementations by adjusting the timetable even though in the project we are limited in terms of content and time. For this reason, we will also communicate the fast pace and difficulty level of courses even more clearly to all interested in participating in our future half-year study programs.

Nonetheless, we are aware of the limitations of our research presented in this paper. As regards the methodological approach, asking ChatGPT for advice on how to address challenges can be questionable because, as claimed by

Spohn (2023) “you cannot trust AI, answers have to be validated”. For this reason, we commented on each recommendation based on our experiences and considered findings from the literature review. Last but not least, the limitation is also that we set the top six challenges which were researched in this paper based only on the first cycle of half-year educational programs at our faculty. However, since those challenges are also among the top in general, this research should provide readers with some useful insight and knowledge.

7 Acknowledgements

The research presented in this paper is a part of the evaluation activity of the RRP pilot project »Applied Computer Skills« financed by the Slovenian Ministry of Higher Education, Science and Innovation, and the European Union - NextGenerationEU.

8 References

- [1] Astudillo, M. L.; Martos, R.; Reese, T. M.; Umpad, K. J.; Dela Fuente, A. The Effects of Time Management among Working Students of Selected Grade 12 General Academic Strand of Senior High School in Bestlink College of the Philippines. *Ascendens Asia Singapore – Bestlink College of the Philippines Journal of Multidisciplinary Research*, 1:1, 2019.
- [2] Chansaengsee, S. Time management for work-life and study-life balance. *Humanities and social sciences*, 10, 20-34, 2017, 2017.
- [3] Ferrer, J.; Ringer, A.; Saville, K.; Parris, M. S.; Kashi, K. Students’ motivation and engagement in higher education: the importance of attitude to online learning. *Higher Education*, 83, 317-338, 2022.
- [4] Jiménez-Cortés, R.; Aires, L. Feminist trends in distance and hybrid higher education: a scoping review. *International Journal of Educational Technology in Higher Education*, 18: 60, 2021.
- [5] Joseph, M. A.; Natarajan, J., Buckingham, J., Al Noumani, M. Using digital badges to enhance nursing students’ attendance and motivation, *Nurse Education in Practice*, 52: 103033, 2021.
- [6] Moores, E.; Birdi, G. K.; Higson, H. E. Determinants of university students’ attendance, *Educational Research*, 61:4, 371-387, 2019.
- [7] Pantzos, P.; Gumaelius, L.; Buskley, J.; Pears, A. Engineering students’ perceptions of the role of work industry-related activities on their motivation for studying and learning in higher education. *European Journal of Engineering Education*, 48:1, 91-109. 2022.
- [8] Ryan, R. M.; Deci, E. L. Intrinsic and Extrinsic Motivation from a Self-Determination Theory Perspective: Definitions, Theory, Practices, and Future Directions. *Contemporary Educational Psychology*, 61: 101860, 2020.
- [9] Sandeep Kumar, G.; Jiju, A.; Fabian, L.; Jacqueline, D. Lean Six Sigma for reducing student dropouts in higher education – an exploratory study, *Total Quality Management & Business Excellence*, 31:1-2, 178-193, 2020.
- [10] Sicam, E. B.; Umawid, M. D.; Colot, J. D.; Dagdag, J. D.; Handrianto, C. Phenomenology of Parenting while Schooling among Filipino College Student Mothers in the Province. *Jurnal Pendidikan Luar Sekolah*, 9:2, 80-94, 2021.
- [11] Spohn, H. Use of ML - AI in the military environment. *Infosek conference*, 4.9.-6.9.2023, Nova Gorica, Slovenia.
- [12] Van Lehn, K. Cognitive skill acquisition. *Annual Review of Psychology*, 47, 513-539, 1996.

AI-Based Analysis and Studying Cheating Behaviours

Mustafa Basic, Selena Kurtic

Faculty of Information Studies, Novo Mesto, Slovenia

International Business-Information Academy, Tuzla, Bosnia and Herzegovina

{ basic.mustafa@yahoo.com, selena.kurtic@yahoo.com }

Abstract: *In the era of educational technology, academic integrity has become a paramount concern. This paper presents a comprehensive study that combines AI-based analysis with an investigation into cheating behaviors among students. Our research delves into the use of artificial intelligence as a tool for assessing student projects and assignments. We employ advanced AI algorithms to evaluate the authenticity and originality of submitted work.*

Furthermore, our study incorporates a questionnaire where students were asked about their utilization of external assistance or AI-related tools in completing their projects. The collected data helped us to study the cheating behaviours and the motivations behind seeking external support.

This study serves as a significant contribution to the ongoing discourse on maintaining academic honesty while harnessing the potential of AI in education.

Key Words: *AI-based analysis, Student projects, AI tools, Academic dishonesty, External assistance, Student questionnaire, Student Projects, Plagiarism detection, Ethics in education, AI in education*

1 Introduction

The phenomenon of cheating in education and assessment has been a persistent challenge that has drawn significant attention from educators, researchers, and institutions worldwide. The whole process of exams, evaluations and student projects does not only impact on the quality of learning, they also effect on the credibility of educational institutions. The educational landscape is being reshaped by technological advancements, which means that cheating behaviors are also evolving, necessitating innovative methods for detection and analysis.

This paper delves into the dynamic intersection of artificial intelligence (AI) and the intricate world of cheating behaviors. Cheating, in its multifaceted forms, poses a complex problem, encompassing activities ranging from plagiarism in written assignments to the use of unauthorized resources during examinations. Traditional methods for detecting cheating have often proven limited in their scope, relying heavily on manual inspection and subjective judgment. As such, there is a growing imperative to harness the power of AI for more comprehensive and data-driven insights into these behaviors.

The objectives of this paper are twofold: firstly, we studied the responses from a questionnaire, that was filled up by the students. Secondly, we used two AI tools to study the code from the projects that the same students have submitted.

The goal of this research is to bridge the gap between traditional methods and cutting-edge technology. This paper contributes to the ongoing discourse on the role of the AI in educational institutions, underscoring its potential to identify dishonesty, educational practices and ultimately elevate the quality of learning and assessment.

Through a synthesis of AI methodologies and a critical examination of cheating behaviors, this paper endeavors to equip educators, institutions, and researchers with a deeper understanding of the challenges posed by cheating in the digital age and the transformative possibilities that AI brings to this critical arena of education.

2 Data Collection

We collected the code submissions from 43 students (39 male and 4 female). It included students who were enrolled in the subject for the first and second time. Both group of students, had to submit the code at the same time, with the project assignment from this year. The students were given a questionnaire at the final test, which was held in 4 terms, to review their honest while the creation of the projects and writing code.

Beside the questionnaire, we used the help of two AI tools (contentdetector.ai and copyleaks.com) to evaluate the code of the submissions, in order to check for plagiarism.

ContentDetector.AI is an accurate and free AI Detector and AI Content Detector that can be used to detect any AI-generated content. It provides a probability score based on the likelihood that the text content was generated by AI tools or chatbots.

CopyLeaks.com checks for plagiarism using advanced AI to detect the slightest variations within the text, including hidden and manipulated characters, paraphrasing, and AI-generated content.

3 Data Analysis

The research is structured to analyze completed project tasks and a questionnaire that was given to the students at the final exam. In this research we used two samples of the code that we tested and performed analysis.

For the project task, students had to submit a written code in the C++ programming language. The task description involves crating a terminal application through which they will navigate using menu commands numbered from 1 to n. The application is intended to have features such as individual application display (window). Students can make arbitrary additions, deletions and other actions. They worked on project task in groups and divided the parts of the code individually by themselves.

At the end of the semester, the students were given a test that included a questionnaire of what type of help did they use while developing their projects.

3.1. Analyzing Questionnaire

The questionnaire was formed in such a way as to contain a range of answers what kind of help do students use while working on their projects. They had the option to select one of the following options: “YouTube”, “ChatGPT”, “Google”, “Friend’s help”, "Other" and "No help at all".

As we can see on the following chart (Figure 1.), most of the students used “YouTube” as a source for studying and developing and what attracted our attention was that 17 students selected “other” as source for extra help and only 1 student selected “Friends help”.

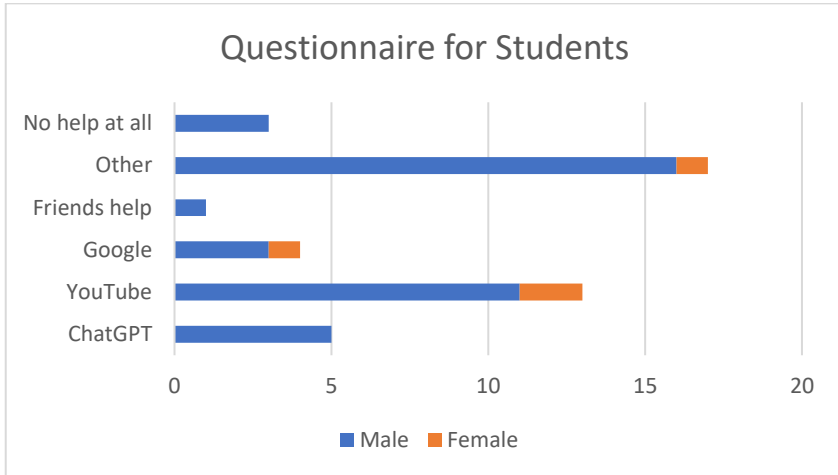


Figure 1: Questionnaire for Students

By putting themselves on “the safe place” and selecting “other” on the questionnaire, they were dishonest with what kind of help they used for their projects.

If we would just take a look at the results from the questionnaire, we might think that the students did all correctly and by the given rules.

3.2. Testing AI tools

For the second part of this research, we used two samples of code and tested it via two AI tools (*contentdetector.ai* and *copyleaks.com*). The code was used from the same group of students, were we tested each one of them individually through both AI tools.

With the given results, form comparing the code that students attached to their projects, we came to the conclusion that the honest of the students was not accurate.

As we can see on Figure 2., the percentage sum of the comparability of the codes showed that they cheated, and even used the help of artificial intelligence to complete the tasks.

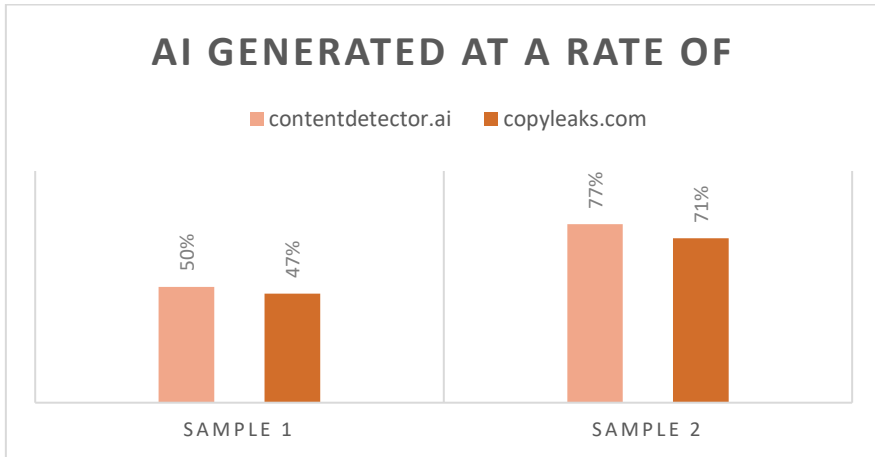


Figure 1: AI generated at a rate of two tools

After inputting the first sample with the *contentdetector.ai* tool, we obtained the results with the rate of 50% and with the second sample the rate of 77%. The second tool for verification we used is *copyleaks.com*, where we obtained the results of 47% for the first sample and 71% for the second sample.

The results give us the insight of the high number of plagiarisms. While proceeding with the tests on both of the tools, even when we used small code sequences, the outcome provided a huge number in percentages. The threshold was 50%, which means that all numbers that are below 50% are human work with minor help from ChatGPT or other sources and presents an individual submitted code. The problem arises with those who are above 50%, meaning that it is not human work and it has been either copy/pasted between students or they used a lot of help from Google and open source codes, including help from ChatGPT. This is where we set the red alert and worry about cheating as plagiarism is illegal and should not be the habit for feature developers.

The research results show a correlation between the data obtained from the questionnaire and verification using AI tools. It is important to note that currently, we do not have tools that can definitively confirm whether a document, image, or code is 100% generated using AI. There are some possibilities that companies could add a hidden signature to their AI generators, which could be identified after a user copies or downloads the content.

4 Conclusion

The findings of this study shed light on a concerning trend in educational settings, where the boundaries between authentic student work and external assistance, particularly from AI tools, have become increasingly blurred. Through a comprehensive analysis of student responses and AI-generated results, it becomes evident that a significant proportion of students have resorted to dishonest practices in their coding projects.

The high incidence of AI tool usage among students, as indicated by the data, underscores the urgent need for a reevaluation of academic integrity policies and practices. It is essential for educational institutions to not only detect and deter cheating behaviors effectively but also to address the root causes driving students to seek external assistance.

Furthermore, the integration of AI into education, while promising, poses ethical and pedagogical challenges that demand careful consideration. Educators and policymakers must collaborate to strike a balance between leveraging AI for educational enhancement and preserving the fundamental principles of learning and assessment.

For feature analysis, we would like to compare the behaviors between different demographics or educational institutions. Where we would also expand the questionnaire with additional questions, in order to get demographic variables (e.g., gender, age), types of cheating behaviors or educational settings. Also, test the projects of different subjects, which means testing the code of different code languages in order to see if there is a code language that the student prefer more and/or cheat more.

5 References

- [1] Foster Provost and Tom Fawcett, "Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking", O'Reilly Media; 1st edition, 2013.
- [2] Kevin P. Murphy, "Machine Learning: A Probabilistic Perspective", The MIT Press Cambridge, Massachusetts London, England, 2012.
- [3] Robert S. Witte and John S. Witte, "Statistics", Wiley: 11th Edition, 2017.
- [4] Alejandro Peña-Ayala (Editor), "Educational Data Mining: Applications and Trends", 2014.
- [5] Philipp K. Janert, "Data Analysis with Open Source Tools: A Hands-On Guide for Programmers and Data Scientists", O'Reilly Media; 1st edition, 2010.
- [6] Ranjit Kumar, "Research Methodology: A Step-by-Step Guide for Beginners.", SAGE Publications Ltd; Third Edition, 2010.
- [7] David S. Moore, George P. McCabe, and Bruce A. Craig, "Introduction to the Practice of Statistics", W.H. Freeman; Ninth edition, 2016.
- [8] Mark D. Shermis and Jill Burstein, "Automated Essay Scoring: A Cross-Disciplinary Perspective", Routledge, 2013.
- [9] M. S. Hale, D. Jensen, and M. Tompkins, "Detecting Cheating in Online Games", SIGKDD Explorations Newsletter, 2008.
- [10] Andrew H. Sung, Yang Cao, and Qiong Wu, "The Application of Machine Learning Algorithms in the Detection of Cheating Behaviors in Online Learning Environments", International Journal of Information and Education Technology, 2016.
- [11] S. Srivastava, S. Vashishtha, and P. Prason, "Artificial Intelligence and Its Role in Education", International Journal of Computer Science and Information Security, 2012.
- [12] Kent M. Kaiser and Joanna C. Dunlap, "Assessment of Student Learning in the Online Classroom", Online Journal of Distance Learning Administration, 2012.

ChatGPT as Psychic: Excellent for Ceremonial Speeches

Ana Hafner

Rudolfovo – Science and Technology Center Novo mesto
Podbreznik 15, 8000 Novo mesto, Slovenia
ana.hafner@rudolfovo.eu

Adela Černigoj

Victoria University of Wellington
8 Kelburn Parade, Wellington, New Zealand
adela.cernigoj@vuw.ac.nz

Abstract: *This paper presents the results of a case study, an unintentional experiment in which ChatGPT wrote a birthday speech. By analysing the text, we aimed to answer the question: why was the speech so effective and how ChatGPT adheres to the recommendations for successful speech design? We can confirm that ChatGPT has excellent abilities for designing different speeches, especially relating to a specific person since it successfully applies cold reading techniques.*

Key Words: *ChatGPT, ceremonial speeches, speech analysis, case study, experiment, cold reading, probability calculus, probability theory*

1 Introduction

ChatGPT has caused much excitement and fear since its release in November 2022. It has drawn a great deal of attention from the natural language processing community since it can generate high-quality responses to human input and self-correct previous mistakes based on subsequent conversations [1]. It can answer several questions, write English essays on various topics [2], and is able to generate various other forms of writing, including poems, lyrics for a song, and even an academic essay: Yes, reports have emerged of scientific articles in journals that were written by using ChatGPT [3]. But it deserved the most attention so far when it comes to student assessment, from euphoria to doomsday predictions [4].

Research so far has produced mixed results. While Vázquez-Cano et al. [5] clearly showed that ChatGPT text obtained the best grades among texts of other real 15-year-old students (evaluated by 30 Spanish language teachers), study by Bašić et al. [6] gave different results: they found no evidence that using GPT improves essay quality since the control group which did not use ChatGPT outperformed the experimental group in most parameters. This discussion will probably continue in the next year along with a proposal of methods and tools to effectively detect texts written by artificial intelligence, as traditional plagiarism detection tools are no longer effective enough (Khalil and Er gave a very good suggestion: ask ChatGPT if it wrote the text [7]).

In this paper we focus on another topic which is not researched yet. If ChatGPT is successful in writing student essays and even scientific articles, there is almost no doubt that it can perform very well in designing all kinds of speeches as well. Some anecdotal

evidence exists about the successfulness of ChatGPT in preparing wedding vows [8] and eulogies [9]. Our main research question however, is not how successful is ChatGPT at writing valedictory, acceptance, commencement, funeral, and other speeches – but why is it so successful? Though our research is based on one case only, we can confirm that ChatGPT has tremendous potential in the field of speech preparation.

2 Methods

Our research is based on a single case study which was an unplanned (unintentional) experiment. The first author was supposed to write a birthday speech for a correspondent author of this paper. The date of the birthday celebration was the 19th of August, 2023. Due to the lack of time, she asked ChatGPT to do it. Since the experiment was unplanned, there was no systematic recording of audience responses, but since that date is not so far away, both authors still remember the reactions well. What is more, the text is still available on ChatGPT, exactly as it was read at the event, only that it was translated into Slovenian. Thus, the text also enables linguistic and psychological analysis i.e., what crucial aspects in the text make it so effective.

3 Results

At the birthday celebration, the first author announced to the audience that this speech was written by ChatGPT. However, after the presentation, many guests said that they simply could not believe that artificial intelligence was able to write such a text.

Now let us try to analyse what it is within this speech that makes it so effective. On the internet, we can find several pieces of advice on how to write a good speech. We will take eight steps of developing a good speech by a career development website [10] – this is the first score in Google if you enter a question: How to write a good speech?



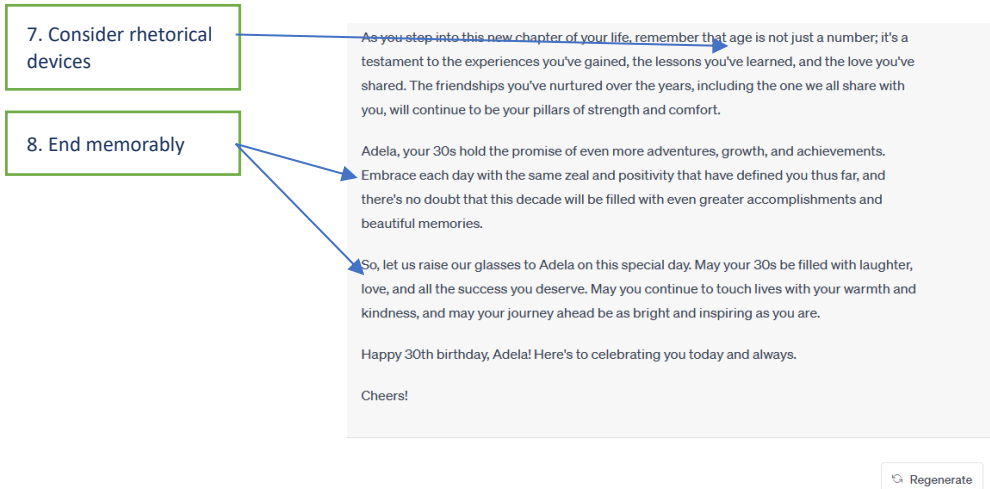


Figure 1: Birthday celebration speech with eight steps of developing a good speech

Due to a lack of space, we will not discuss all eight points. We will leave out the most self-evident ones, such as the first point. We will go to the second point: How did ChatGPT consider the audience? Since this was a birthday celebration, ChatGPT correctly assumed that mainly friends and family would be there, so there certainly is a collective love and admiration, because the audience knows Adela very well.

We continue with the fourth point. “Starting with a strong, clear purpose can show your audience where you intend to lead them,” they say on our reference website [10]. ChatGPT’s start is remarkable: “Today, as we mark this significant milestone in Adela’s life, we are not just celebrating the passing of another year, but we are rejoicing in the incredible journey that has brought her to this point.”

The fifth point suggests using concrete details such as brief stories, interesting examples, or factual data that can help engage the audience [10]. These concrete details were: “...you’ve achieved so much in these three decades – from your academic accomplishments to your professional successes. Your dedication and hard work are an inspiration to us all...”

How about the personal element (sixth point)? There are probably several, but this one is really effective: “...you have always been the heartbeat of our group – the one who brings joy to every gathering...” Who would not be touched by that?

The seventh point suggests a rhetorical device which is a method of using words to make them especially memorable. They suggest considering some of the most memorable lines from famous speeches [10]. Because ChatGPT takes everything from previous speeches it already consists of rhetorical devices. But some are probably used especially frequently, like “age is not just a number” or “it’s a time of reflection on the past and anticipation for the future”.

Finally, they suggest that at the end of the speech we have to return to the strong purpose we began with and create a vision of the future [10]. There is probably almost no other choice than to end the birthday speech with an invitation to celebration: “...let us raise our

glasses /.../ May you continue to touch lives with your warmth and kindness, and may your journey ahead be as bright and inspiring as you are.”

4 Discussion and Conclusion

While the ChatGPT is very effective according to all the tips on how to write a good speech, we can also see the speech is completely “empty”. In other words, the originality of the speech is close to zero: speech is mostly made of phrases we have heard before. If we replace “Adela” with “Andrej” all could fit as well with the exception of “embodiment of grace” since grace is rarely attributed to men in our cultural context.

There are some statements, however, that might look surprising at first sight. We have to know that ChatGPT had no information about Adela except that she is 30 years old. How did ChatGPT know Adela has no children yet? How did ChatGPT know Adela had some academic accomplishments? ChatGPT is example of large language models which are a kind of artificial intelligence technology used to generate natural language [11] and are, first and foremost, artificial neural networks [12]. Basically, when ChatGPT is writing a response is just asking over and over again “given the text so far, what should the next word be?” – and each time adding a word or, more precisely, a “token” [13]. To put it simply, it calculates the probability of appearance of (next) words. Since it was trained on vast amounts of textual data [14], we can assume that there have been quite a few birthday speeches for 30-year-olds that emphasize the importance of family and children. The question we therefore meaningfully ask is, does ChatGPT provide answers based on any data it collects about the person who is asking? What information about the users does OpenAI, a developer of ChatGPT, collect? In their frequently asked questions, they admit to collecting IP address which can be used to estimate the device’s country, state, and city [15].

Therefore, an interesting question is, does ChatGPT also form responses based on statistical predictions from the data collected from the user. That could very much explain everything: if ChatGPT recognized that the user comes from Slovenia, it can know that the average age of women who gave birth for the first time in 2020 was 29.6 years in Slovenia [16]. The situation is similar to the UK where 53% of women born in 1991 were childless by their 30th birthday last year [17]. More than half of Slovenian women, aged between 30 and 39 (in 2019) have completed higher education. ChatGPT is on the safe side, mentioning academic achievements as well as avoiding commenting on children.

Regardless of whether ChatGPT solely computes the probability of word occurrences (using pre-trained data) or its responses incorporate the information it has about the user, one can conclude that ChatGPT effectively employs cold reading techniques. These techniques, most frequently practiced by psychics, fortune-tellers, astrologists, tarot card readers, and other similar practitioners, can be defined as “a procedure by which a ‘reader’ is able to persuade a client whom he has never before met that he knows all about the client’s personality and problems” [18, p.81]. Cold readers can vary from a simple reliance on using statements that are true of most people to a broader definition that includes pre-session information-gathering about a client [19]. They especially exploit the Barnum effect (also called the Forer effect) which relies on the need of people to make connections between what is said and some aspect of their own lives within statements

that seem personal, yet they apply to many people [20]. Most accomplished cold readers have as an information foundation, knowledge of probability calculus and the common denominators of the human condition [21], and so does ChatGPT.

We can conclude that ChatGPT has a bright future for writing various speeches. Since not all people have the same ability to express themselves verbally, especially when they have to deal with heavy, emotionally demanding topics (such as death and writing a eulogy), tools such as ChatGPT might come very handy in these unavoidable situations.

5 References

- [1] C. Qin, A. Zhang, Z. Zhang, J. Chen, M. Yasunaga, and D. Yang, "Is ChatGPT a general-purpose natural language processing task solver?" *arXiv preprint arXiv:2302.06476*, 2023.
- [2] T. N. Fitria, "Artificial intelligence (AI) technology in OpenAI ChatGPT application: A review of ChatGPT in writing English essay", *ELT Forum: Journal of English Language Teaching*, vol. 12, no. 1, pp. 44-58, Mar. 2023.
- [3] H. Zheng and H. Zhan, "ChatGPT in scientific writing: a cautionary tale", *The American Journal of Medicine*, 2023.
- [4] J. Rudolph, S. Tan, and S. Tan, "ChatGPT: Bullshit spewer or the end of traditional assessments in higher education?" *Journal of Applied Learning and Teaching*, vol. 6, no. 1, 2023.
- [5] E. Vázquez-Cano, J. M. Ramírez-Hurtado, J. M. Sáez-López, and E. López-Meneses, "ChatGPT: The Brightest Student in the Class," *Thinking Skills and Creativity*, p. 101380, 2023.
- [6] Ž. Bašić, A. Banovac, I. Kružić, and I. Jerković, "Better by You, better than Me? ChatGPT-3 as writing assistance in students' essays", 2023 [Online]. Available: https://www.researchgate.net/publication/368393093_Better_by_you_better_than_me_chatgpt3_as_writing_assistance_in_students_essays, Sep. 4, 2023.
- [7] M. Khalil and E. Er, "Will ChatGPT get you caught? Rethinking of plagiarism detection", *arXiv preprint arXiv:2302.04335*, 2023.
- [8] M. K. Samantha, "Who says romance is dead? couples are using ChatGPT to write their wedding vows", *CNN Wire Service*, Apr. 2023. Retrieved from <https://go.openathens.net/redirector/wgtn.ac.nz?url=https://www.proquest.com/wire-feeds/who-says-romance-is-dead-couples-are-using/docview/2799424388/se-2>, Sep. 12, 2023.
- [9] J. Rose, "I Had an AI Chatbot Write My Eulogy. It Was Very Weird", Feb. 17, 2023. Retrieved from: <https://www.vice.com/en/article/4axjj9/ai-chatbot-wrote-my-eulogy-chatgpt-death>, Sep. 12, 2023.
- [10] Indeed, "The 8 Key Steps to Successful Speech Writing (With Tips)". Retrieved from: <https://ca.indeed.com/career-advice/career-development/speech-writing>, Sep. 5, 2023.
- [11] A. Gokul, "LLMs and AI: Understanding Its Reach and Impact", 2023. Retrieved form: <https://www.preprints.org/manuscript/202305.0195/v1>, Oct. 29, 2023.
- [12] D. Valdenegro, "A LLM digest for social scientist", 2023. Retrieved form: <https://osf.io/preprints/socarxiv/m74vs/>, Oct. 29, 2023.

- [13] S. Wolfram, "What Is ChatGPT Doing ... and Why Does It Work?" 2023. Retrieved from: <https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/> Oct. 29, 2023.
- [14] G. Orrù, A. Piarulli, C. Conversano, and A. Gemignani, "Human-like problem-solving abilities in large language models using ChatGPT," *Frontiers in Artificial Intelligence*, vol. 6, p. 1199350, 2023.
- [15] OpenAI, "Frequently asked questions about the ChatGPT Android app," 2023. Retrieved from: <https://help.openai.com/en/articles/8142208-chatgpt-android-app-faq>, Oct. 29, 2023.
- [16] SURS, "V 2020 manj kot 19.000 živorojenih otrok", 2020. Retrieved from: <https://www.stat.si/StatWeb/news/Index/9636>, Sep. 9, 2023.
- [17] J. Davies, "Half of women are now childless at thirty for the first time ever: Official statistics show most common age for giving birth has risen to 31 - compared to 22 for baby boomers", MailOnline, 2022. Retrieved from: <https://www.dailymail.co.uk/health/article-10447507/Half-women-childless-thirty-time-ever.html>, Sep. 9, 2023.
- [18] R. Hyman and K. Frazier, "'Cold reading': How to convince strangers that you know all about them", *Paranormal borderlands of science*, pp. 79-96, 1981.
- [19] C. A. Roe and E. Roxburgh, "An overview of cold reading strategies", *The Spiritualist Movement: Speaking with the dead in America and around the world*, vol. 2, pp. 177-203, 2013.
- [20] D. L. Dutton, "The cold reading technique", *Experientia*, vol. 44, pp. 326-332, 1988.
- [21] R. Novella, "Cold Reading: The Psychic's True Power", 1997. Retrieved from: <https://theness.com/index.php/cold-reading-the-psychics-true-power/>, Sep. 14, 2023.

Comparative analysis of Slovenian and Czech digitalization index: an AI approach

Nuša Erman[†], Tomáš Hlavsa[‡]

[†]Faculty of Information Studies
Ljubljanska cesta 31A, 8000 Novo mesto, Slovenia

[‡]Czech University of Life Sciences Prague, Faculty of Economics and Management
Kamýcká 129, 165 00 Praha – Suchdol, Czech

nusa.erman@fis.unm.si, hlavsa@pef.czu.cz

Abstract: *The increased adoption of digital technologies has triggered a variety of changes in business and social environments, which consequently led to the so-called digital transformation. The European Commission has developed an online tool, known as the Digital Economy and Society Index (DESI), which allows data to be collected to monitor important digitalization indicators and make comparisons between EU Member States. The paper presents the evolution of the level of digital economy and society in Slovenia and Czechia and compares their digital performance. The results show that both countries are active in the field of digitisation and their performance is improving year by year. But they both still have room for improvement; Slovenia in the Human capital dimension, and Czechia in the Connectivity and Integration of digital technology dimension.*

Key Words: *digital transformation, digital economy in society index, level of digitalization, digital performance.*

1 Introduction

The increased adoption of digital technologies has triggered a variety of changes in business and social environments, which consequently led to the so-called digital transformation. Accordingly, countries and regions worldwide have become increasingly aware of the importance of shifting focus to digital by harmonizing the digital environment at both national and international levels. In this context, in 2015 the European Union launched the EU's Digital Single Market Strategy [1], which represents a common European framework focusing on three main pillars: 1) allowing better access to online goods and services for consumers and businesses, 2) enabling the flourishing of digital networks and services by the creation of the right conditions, and 3) gaining to the maximization of the European Digital Economy's growth potential.

Increasing awareness of the digitalization of the global economy and the drive to maintain a position as a world leader in the digital economy, the European Union also needed a tool which would enable the assessment of the development of EU countries towards a digital economy and society. To this end, the European Commission has developed an online tool, known as the Digital Economy and Society Index (DESI), which allows data

to be collected to monitor important digitalization indicators and make comparisons between EU Member States through annually released reports [2].

DESI reports initially offered insights into a five-dimension structure of the Member States' digital progress, but in 2021 the structure changed and began to follow the digital targets of Europe's Digital Decade policy programme [3]. The structure of the report now includes four dimensions: 1) the Human capital dimension, which measures advanced digital skills, as well as internet use skills; 2) the Connectivity dimension, which focuses on the fixed and mobile broadbands, their take-up, coverage, and prices; 3) the Integration of digital technology dimension, which covers the areas of business digitalization and e-commerce; and 4) the Digital public services dimension, which focuses on the e-government services.

Focusing on the presented dimensions, the aim of this paper is twofold: 1) to explore how the status of the digital economy and society evolved over time, focusing on the comparison between two EU Member States, i.e., Slovenia and the Czechia; and 2) to provide further insights into the data relating to the measurement of scores for four dimensions which add up to an overall DESI score.

The rest of the paper is organized as follows: after defining the methods used for the research, we present the analysis results which are then followed by a discussion section. We conclude the paper by highlighting the main findings and outlining further research work.

2 Methods

To compare the Slovenian and Czech digitalization index, we used data gathered in the framework of monitoring the progress of EU Member States towards a digital economy and society (DESI), in the 2014-2022 period (see refs. [4] - [19]). We focused on the overall DESI score data indicating the overall level of the country's digitalization of economy and society as well as the DESI scores of the above-presented dimensions.

Since the objective of the research presented in this paper is to offer preliminary insights into the similarities and/or differences between the observed countries, we first present the data using a bar chart to compare the countries' overall DESI scores in the observed time period, followed by the stacked bar chart to compare the countries' DESI scores by dimensions. In both cases, DESI scores were also used to calculate annual growths, and average annual growth rates (AAGRs). To consider the impact of compounding that accrues over the years as well, we also calculated compound annual growth rates (CAGRs).

3 Results

In order to make a comparison between Slovenia and Czechia, in Figure 1 we present annual values of the DESI overall score for the two countries, to which we have also added the values for the EU.

If we first compare the overall level of digitalization in Slovenia and Czechia with the EU's overall level of digitalization, we can notice that in the last two years, Slovenia has achieved a somewhat higher level of digitization as compared to the digitalization level of the EU. On the other hand, we can also observe that Czechia as compared to the EU scores in a lower level of digitization throughout the whole time period observed.

As also shown in Figure 1, the overall level of digitalization of economy and society shows that Czechia has outperformed Slovenia in the first two years, but in 2016 we can acknowledge a turning point as Slovenia's overall level of digitalization caught up with

that of the Czechia, and as of 2017, Slovenia's overall level of digitalization of economy and society is higher than the Czechia's. Slovenia's stronger progress in the level of digitalization compared to Czechia can also be seen in terms of AAGR and CAGR values. In the observed period, the DESI overall score in Slovenia increased by 4.2% on average, exhibiting a CAGR of 4.0%, while in Czechia it grew by 2.3% on average, exhibiting a CAGR of 2.2%.

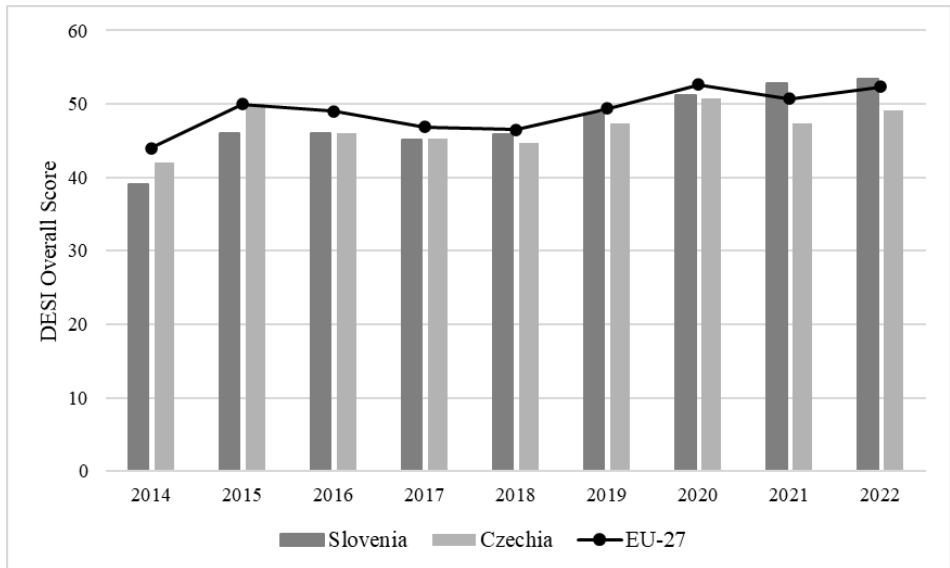


Figure 1: DESI overall scores for Slovenia, Czechia and EU in 2014-2022 (refs [4] – [19])

In Figure 2 we present the annual values of the DESI score for four dimensions, i.e., the Connectivity dimension, the Human capital dimension, the Integration of digital technology dimension, and the Digital public services dimensions, for both Slovenia and Czechia.

Regarding DESI scores on the Connectivity dimension, Czechia has led Slovenia until 2017, but in 2018 we can acknowledge a turning point as Slovenia's DESI connectivity score is higher than the Czechia's. Progress in connectivity in Slovenia is also reflected in average annual growth which indicates that in the observed period DESI connectivity score in Slovenia increased by 6.2% on average, exhibiting a CAGR of 5.5%, whereas the average annual growth of the DESI connectivity score in Czechia reaches a value of 2.3%, exhibiting a CAGR of 1.2%.

The situation is different for the Human resources dimension. The data show that DESI human resources scores are higher in Czechia in the observed time period, whereas Slovenia's performance in 2017, 2018, and 2019 is only marginally better. However, we note that both Slovenia and Czechia have experienced a decline in the average value of the DESI human resources score. Namely, in the observed period, the DESI human resources score in Slovenia dropped by 1.1% on average, exhibiting a CAGR of -1.5%, while in Czechia it dropped by 0.6% on average, exhibiting a CAGR of -1.4%.

Focusing on the DESI score on the Integration of digital technology dimension, we can observe that Czechia scores higher as compared to Slovenia, but in the last two years, Slovenia has taken the lead. In the observed period, the DESI integration of digital technology score in Slovenia increased by 4.8% on average, exhibiting a CAGR of 4.5%, but in Czechia it grew only by 0.1% on average, exhibiting even a negative value of CAGR (-0.8%).

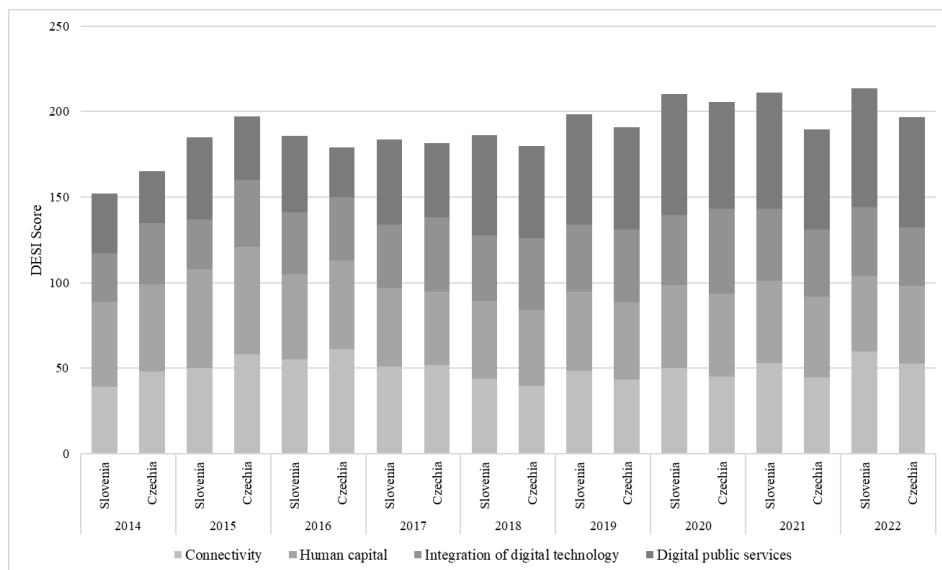


Figure 2: DESI scores for four dimensions (connectivity, human capital, integration of digital technology, and digital public services) for Slovenia and Czechia in 2014-2022 (refs [4] – [19])

Finally, the DESI scores on the Digital public services dimension show significant progress in both Slovenia and Czechia. Although the DESI digital public services scores for Slovenia are higher than for Czechia over the whole observed time period, the progress is even more notable for Czechia. In the observed period, the DESI digital public services score in Slovenia increased by 9.7% on average, exhibiting a CAGR of 9.0%, while in Czechia it grew by 11.9% on average, exhibiting a CAGR of 10.0%.

4 Discussion

The data analyzed in our research showed that both Slovenia and Czechia are on track to meet the targets set by Europe's Digital Decade policy programme [3]. Both countries are evolving in terms of the overall level of digitalization of economy and society, as well as in terms of individual dimensions. According to the DESI 2022 report, Slovenia ranks 11th [18] and Czechia ranks 19th among the 27 EU Member States regarding the overall digital economy and society index. Both countries face opportunities and challenges in terms of progress in digitization, which they are managing relatively successfully over time.

Referring to the two-fold aim of this paper mentioned, let us first focus on exploring how the status of the digital economy and society evolved over time, focusing on the

comparison between two EU Member States, i.e., Slovenia and Czechia. The analysis showed that in the overall level of digitalization of economy and society, Czechia outperformed Slovenia in the first two years, but in 2016 Slovenia caught up with Czechia, and in 2017 took the lead.

As regards the second fold of the aim, we also provided further insights into the data relating to the measurement of scores for four dimensions which add up to an overall DESI score, i.e., the Human capital dimension, the Connectivity dimension, the Digital public services dimension, and the Integration of digital technology dimension, which also covers the use of Artificial Intelligence (AI). In terms of the levels of digitalization in the areas of connectivity and integration of digital technology, both the higher average annual growth rate and the higher compound annual growth rate suggest that Slovenia is outperforming Czechia. Regarding the level of digitalization in the area of digital public services, both countries exhibit strong performance, with Czechia showing a higher average annual growth rate and a higher compound annual growth rate compared to Slovenia. Concerning digitization in the area of human resources, we note that both countries still have room for improvement, since they both encounter drops in average annual growth rate as well as in compound annual growth rate.

The main findings of our research are in line with the state of the digitalization of economy in society. Namely, according to [18] and [20], Slovenia's weak point is recognized especially in the low share of digitally skilled individuals, i.e., those with at least basic digital skills, and those with above basic digital skills, which consequently influences the relatively poor performance of Slovenia in the Human capital dimension. On the other hand, Czechia is characterized by strong performance in the Human capital dimension, while according to [19] and [21] its performance lags in the Connectivity dimension and in the Integration of technology dimension.

5 Conclusion

Although the research showed progress in the level of digitization of economy and society for both countries over the observed time period, both Slovenia and Czechia have to strive towards achieving the targets of Europe's Digital Decade policy programme [3] by the year 2030.

The research presented in this paper can be seen as an initial phase in the study of Slovenian and Czechian levels of digitalization. In the future, we plan to further develop our research in the direction of examining individual indicators used for measuring each of the four dimensions which add up to an overall DESI score. We also intend to focus on the field of artificial intelligence itself, which is not yet well covered by the DESI online tool. Data on the percentage of enterprises using AI is only available for 2022, so it will take some time before we can compare the development of countries in the use of AI. Finally, we also plan to use a combination of several data sources, such as the OECD's AI policy observatory [23] or EUROSTAT's Digital Economy and Society data [22].

9 References

- [1] European Commission. A Digital Single Market Strategy for Europe, [SWD(2015) 100 final], https://ec.europa.eu/commission/presscorner/api/files/attachment/8210/DSM_communication.pdf, downloaded: September 27th 2023.
- [2] European Commission. What is the Digital Economy and Society Index?,

https://ec.europa.eu/commission/presscorner/detail/en/MEMO_16_385, downloaded: October 10th 2023.

[3] European Commission. Europe's Digital Decade: digital targets for 2030, https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/europes-digital-decade-digital-targets-2030_en, downloaded: October 10th 2023.

[4] European Commission. Digital Economy and Society Index (DESI), 2015 Country Report, Czechia, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=8778, downloaded: September 25th 2023.

[5] European Commission. Digital Economy and Society Index (DESI), 2015 Country Report, Slovenia, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=8795, downloaded: September 25th 2023.

[6] European Commission. Digital Economy and Society Index (DESI), 2016 Country Report, Czechia, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=14119, downloaded: September 25th 2023.

[7] European Commission. Digital Economy and Society Index (DESI), 2016 Country Report, Slovenia, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=14137, downloaded: September 25th 2023.

[8] European Commission. Digital Economy and Society Index (DESI), 2017 Country Report, Czechia, <https://ec.europa.eu/newsroom/dae/redirection/document/42999>, downloaded: September 25th 2023.

[9] European Commission. Digital Economy and Society Index (DESI), 2017 Country Report, Slovenia, <https://ec.europa.eu/newsroom/dae/redirection/document/43041>, downloaded: September 25th 2023.

[10] European Commission. Digital Economy and Society Index (DESI), 2018 Country Report, Czechia, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=52216, downloaded: September 25th 2023.

[11] European Commission. Digital Economy and Society Index (DESI), 2018 Country Report, Slovenia, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=52237, downloaded: September 25th 2023.

[12] European Commission. Digital Economy and Society Index (DESI), 2019 Country Report, Czechia, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=59889, downloaded: September 25th 2023.

[13] European Commission. Digital Economy and Society Index (DESI), 2019 Country Report, Slovenia, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=59912, downloaded: September 25th 2023.

[14] European Commission. Digital Economy and Society Index (DESI), 2020 Country Report, Czechia, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=66910, downloaded: September 25th 2023.

[15] European Commission. Digital Economy and Society Index (DESI), 2020 Country Report, Slovenia, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=66929, downloaded: September 25th 2023.

[16] European Commission. Digital Economy and Society Index (DESI), 2021 Country Report, Czechia, <https://ec.europa.eu/newsroom/dae/redirection/document/80485>, downloaded: September 25th 2023.

[17] European Commission. Digital Economy and Society Index (DESI), 2021 Country Report, Slovenia, <https://ec.europa.eu/newsroom/dae/redirection/document/80491>, downloaded: September 25th 2023.

[18] European Commission. Digital Economy and Society Index (DESI), 2022 Country

Report, Czechia, <https://ec.europa.eu/newsroom/dae/redirection/document/88698>, downloaded: September 25th 2023.

[19] European Commission. Digital Economy and Society Index (DESI), 2022 Country Report, Slovenia, <https://ec.europa.eu/newsroom/dae/redirection/document/88715>, downloaded: September 25th 2023.

[20] European Commission. Digital Decade Country Report 2023, Czechia, <https://ec.europa.eu/newsroom/dae/redirection/document/98618>, downloaded: September 30th 2023.

[21] European Commission. Digital Decade Country Report 2023, Slovenia, <https://ec.europa.eu/newsroom/dae/redirection/document/98634>, downloaded: September 30th 2023.

[22] OECD. AI Policy Observatory, <https://oecd.ai/en/>.

[23] EUROSTAT. Digital Economy and Society, <https://ec.europa.eu/eurostat/web/digital-economy-and-society>.

Future-based smart city with AI-powered flood support

Valerij Grasic

Telekom Slovenia

Cigaletova 17, 1000 Ljubljana, Slovenia

grasic.se@gmail.com

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: *Smart city research has been underway for a long time, but we are still in the early stages of development. Our primary interest lies in understanding the role of artificial intelligence (AI) in shaping the blueprint of future-centric smart cities. In recent years, there have been many natural disasters that have been related to floods. Our key inquiry is: How can AI assist emergency services in gearing up for potential flood events? Centering our research on flood alarm forecasting, we are evaluating the effectiveness of the Random Forest and Naïve Bayes classification algorithms, specifically in the context of the smart city of Ljubljana. Our comparative study spans a designated timeline, breaking down both holistic and quarterly data. Insights from this analysis aim to establish a robust framework for situational awareness and enhance flood readiness in the smart cities of the future.*

Key Words: *Smart City, Open Data, Emergency, Artificial Intelligence (AI), IoT, Floods*

1 Introduction

In recent years, there have been many natural disasters that have been related to floods. Figure 1 shows the consequences of floods in Ljubljana in 2010. We are interested in how artificial intelligence can influence the framework of a future-based smart city. Our goal is to recognise in advance the extent of incoming emergency calls due to natural hazardous events, like forthcoming flooding, which will be so large that it will be necessary to trigger an alarm within the smart city.

Random Forest and Naïve Bayes classification algorithms are envisaged for this purpose. A total of 20 attributes are used, including open IoT (Internet of Things) data for 2013 - 2016. This data is directly related to the city of Ljubljana and includes data on the river flows and weather data. In this paper we compare different classification models.

2 Problem definition

Although smart city research has been underway for a long time, it is estimated that we are still at the very beginning of smart city development [2], [3]. Flood government is an important part of smart cities. An essential question is how to increase situational

awareness within Smart Cities' control rooms since any additional information is welcome [4]. Approaches for such cases include solutions like early warning systems and social networks [4]. A valuable factor is the modeling of the flood area, also in connection with satellites [5]. Besides this, AI (Artificial Intelligence) can also be useful for sounding the alarm due to emergency calls [6]. While the existing systems provide an average and a maximum number of emergency calls, the issue is how to provide a more dynamic, artificial intelligence-based forecasting of the alarm-triggered number of emergency calls [6], [7].



Figure 1: Consequences of the floods in Ljubljana in 2010 [1]

3 Methodology description

3.1 Framework for incoming calls

At the beginning of November 2014, floods covered central Slovenia, including Ljubljana. Figure 1 shows the trend in the number of incoming calls towards the emergency systems for the first seven days of November.

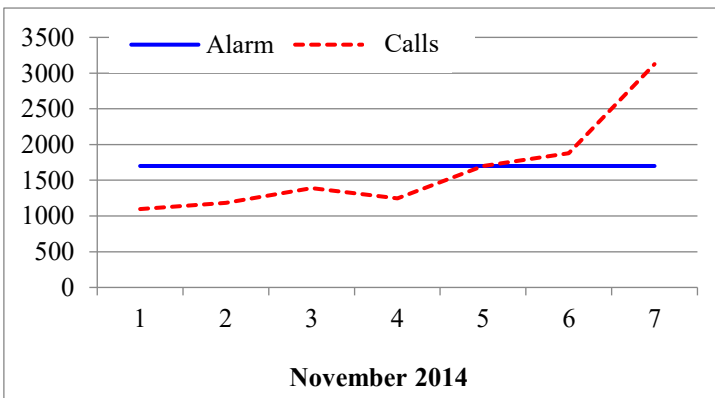


Figure 1: Movement in the number of incoming calls in November 2014

A significant increase in the number of incoming calls can be seen. The idea is to sound an alarm for the sixth and seventh day due to a significantly increased number of incoming calls. In general, that should be valid. For example, 10 to 15% of all incoming calls are due to the significantly increased number. Based on this, the number of incoming calls triggering the class alarm is 1,700 calls per day. Such classes are called regular classes and alarm classes.

3.2 Datasets

The motivation for preparing the data was to consider factors that have caused the most significant damage in recent years, such as floods, and that the data are distributed throughout the country. Twenty different datasets are used, including dataset for river flows (with 16 measuring points) and weather data (consisting of 10 measuring stations, with 16 attributes at each station). The catalogue of open IoT data used is available on GitHub [8].

3.3 Classification methods

Two classification methods are used: Naïve Bayes and Random Forest. The Naive Bayes classifier is one of the most widely used classification methods in industry today [9], [10]. It comes from the Bayes rule, which describes the probabilities of an event, and assumes independence of attributes. Random Forest is an ensemble of decision trees [11], [12]. Prediction is achieved by voting and applying the majority rule of the individual trees. It is suitable for handling missing data, data of different types, and irrelevant data.

3.4 Evaluation metrics

Various metrics are used to evaluate different classification models: accuracy (Acc), F1 measure and ROC curve [13]. Accuracy is used as the primary metric. The tool used for comparison is Weka [14]. The entire process of model evaluation is as follows:

- Perform data normalization and data balancing (by SMOTE).
- Using a cost sensitivity approach, various combinations of C1 and C2 are checked, looking for the value where the accuracy is best.
- A classification is performed, and the associated model is built using cross-validation.
- Classification with different methods is obtained for the whole period (2013–2016). The results and the obtained metrics are checked and compared with each other.
- In the hope of obtaining even better classification results, the data of the given period are additionally divided by quarters. Models are created, and model testing is performed for each quarter separately (Q1–Q4).

4 Evaluation of results

In the following, the forecasting of the number of incoming calls is evaluated. The evaluation is made for data for the entire period and quarters. At the end, the discussion is given.

4.1 Data evaluation

Figure 2 shows the basic statistical parameters for the data set for a given period (2013–2016), as seen from the Weka software tool. In total, we have 1,186 instances of data. The Figure 2 shows the maximum, minimum and mean values of incoming calls. In the given

observed time, Slovenia's smallest number of calls per day was 890, the largest 9,538, and the average number was 1,478.90.

Name: call_in		Type: Numeric
Missing: 0 (0%)		Distinct: 627
		Unique: 302 (25%)
Statistic	Value	
Minimum	890	
Maximum	9538	
Mean	1478.897	
StdDev	437.254	

Figure 2: Basic statistical parameters for the data set for a given period

Figure 3 shows the distribution of incoming calls divided into two classes for a given period. On the left, it refers to the number of calls for the regular class, and on the right, to the number of calls for the alarm class.

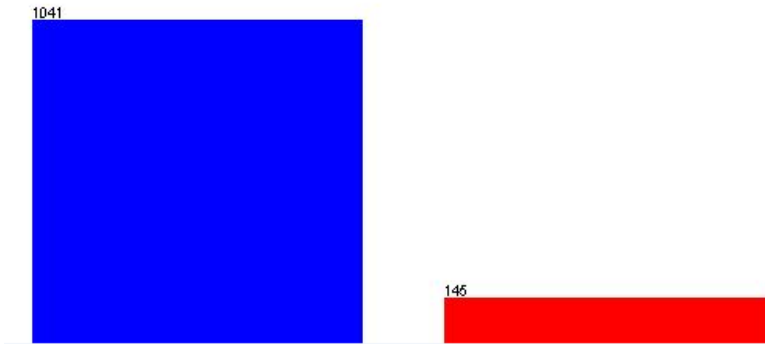


Figure 3: Incoming calls divided into two classes (left: regular, right: alarm) for the data set for a given period

4.2 Evaluation for the whole period

An evaluation has been made for the entire period. Table 1 shows the accuracy, F1 score and the ROC value of the obtained classification models. The best results are given in bold.

Table 1: Classification models for Ljubljana and two classes for the entire period

Method	Acc	F1	ROC
Naïve Bayes	0.747	0.746	0.786
Random Forest	0.914	0.914	0.969

For the two classes for Ljubljana, for a given period, the best classification results were obtained with the Random Forest method, with an accuracy of 91.4%.

4.3 Evaluation for quarters

The timeframe was divided into four time quarters, each one referencing to three consecutive months in the year. We build classification models for each quarter separately and the results are given in Table 2. The best results are in given in bold.

The best classification results were obtained with the Random Forest method in the first quarter, with an accuracy of 95.5%.

Table 2: Classification models for Ljubljana and two classes, by individual quarters

Quarters	Method	Acc	F1	ROC
Q1	Naïve Bayes	0.772	0.772	0.838
	Random Forest	0.955	0.955	0.986
Q2	Naïve Bayes	0.770	0.770	0.834
	Random Forest	0.920	0.920	0.996
Q3	Naïve Bayes	0.743	0.742	0.776
	Random Forest	0.917	0.917	0.970
Q4	Naïve Bayes	0.722	0.722	0.809
	Random Forest	0.933	0.933	0.978

4.4 Discussion

We demonstrated that it is possible to create classification models with high accuracy as presented in Figure 4 for each of the methods. The best results for accuracy were achieved in the case of the Random Forest method. In such a case, the best values for accuracy were 95.5% for the first quarter of a given period (2013-2016).

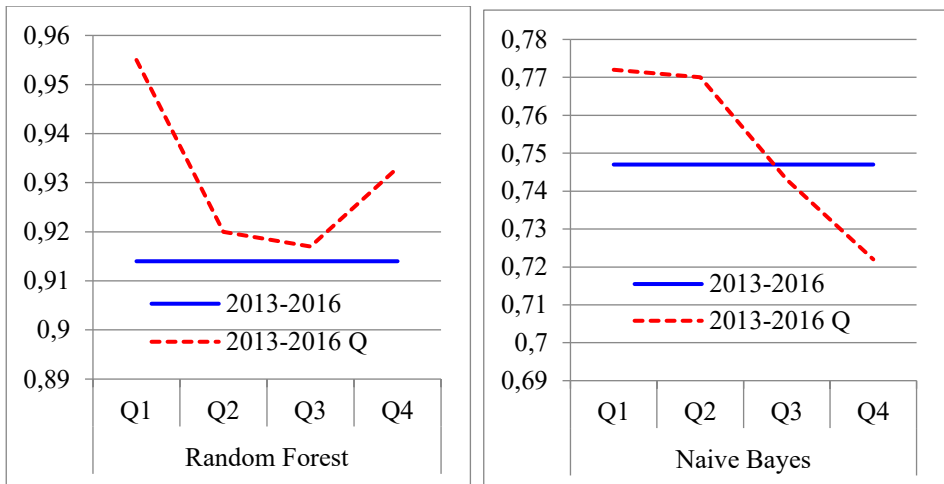


Figure 4: Accuracy by individual methods

The conclusion is that Random Forest is the best method for all the cases, both for the whole period and for quarters. In the case of Naïve Bayes, although the method is very effective in practice, it failed to perform well, when used with the default settings.

5 Conclusion

In recent years, there have been many natural disasters that have been related to floods. There exist different approaches to artificial intelligence flood management. Such cases are warning systems, use of social networks, predictive models of the surface of the flooded area and models of the damage caused. Satellites can also be included.

This article looked at the framework of the future-based smart city in case of emergency calls. An evaluation of the forecasting of the number of incoming calls toward the emergency system was made using open data from the Internet of things. It was shown that Random Forest was the best method for a given period from 2013 to 2016, and the best results were even better for a single quarter. The proposal makes it possible to improve situational awareness within the control room for future-based smart cities. This helps to prepare a big picture of what is happening and, in this way, prepare in advance for the coming floods.

6 References

- [1] Siol.net. Siol.net, <http://www.siol.net/>, downloaded: September 25th 2023.
- [2] Smart Cities Council. Smart Cities Readiness Guide: The Planning Manual for building tomorrow's cities today, <https://smartcitiescouncil.com/resources/smart-cities-readiness-guide>, 2015.
- [3] European Commission. Smart Cities - Smart Living. Shaping Europe's digital future, <https://ec.europa.eu/digital-single-market/en/smart-cities>, 2019
- [4] Middleton, S. E; Middleton, L; Modafferi, S. Real-time crisis mapping of natural disasters using social media. *IEEE Intelligent Systems*, 29(2):9–17, 2014.
- [4] Lamovec, P; Mikoš, M; Oštir, K. Detection of flooded areas using machine learning techniques: case study of the Ljubljana moor floods in 2010. *Disaster advances*, 6(7):4-11, 2013.
- [6] 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.
- [7] Grašič, V. Napovedovanje števila dohodnih klicev na sistem za klic v sili 112 ob uporabi odprtih podatkov interneta stvari, Doctoral dissertation, Faculty of Information Studies (FIS) in Novo mesto, 2021.
- [8] SafeCity112. SafeCity112, <https://github.com/SafeCity112/SafeCity112>, downloaded: August 5th 2022.
- [9] John, G. H; Langley, P. Estimating Continuous Distributions in Bayesian Classifiers. *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, pages 338–345, <https://dl.acm.org/doi/pdf/10.5555/2074158.2074196>, 1995.
- [10] Snoek, J; Larochelle, H; Adams, R. P. Practical Bayesian Optimization of Machine Learning Algorithms. *NIPS'12 Proceedings of the 25th International Conference on Neural Information Processing Systems*, <http://arxiv.org/abs/1206.2944>, 2:2951–2959, 2012.
- [11] Han, J; Kamber, M; Pei, J. *Data Mining. Concepts and Techniques*, 3rd Edition. The Morgan Kaufmann Series in Data Management Systems, 2011.
- [12] Hastie, T; Tibshirani, R; Friedman, J. *The Elements of Statistical Learning*. Springer, 2009.
- [13] Witten, I. H; Frank, E; Hall, M. *Data Mining: Practical machine learning tools and techniques*. Third Edit. Elsevier, 2011.

- [14] Frank, E; Hall, M. A; Witten, I. H. The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques". Fourth Edition, https://www.cs.waikato.ac.nz/ml/weka/Witten_et_al_2016_appendix.pdf, downloaded: September 23rd 2017, 2016.

Enhancing Urban Traffic Flow with AI: A Case Study of YOLO-Based Vehicle Counting

Žan Pogač*

Faculty of Industrial Engineering Novo mesto
Šegova ulica 112, 8000 Novo mesto, Slovenia

zan@pogaclab.com,

doc. Tomaž Aljaž, Ph.D.

Faculty of Information Studies

Ljubljanska cesta 31A, 8000 Novo mesto, Slovenia

tomaz.aljaz@fis.unm.si

Abstract: *Our research explored the potential of artificial intelligence to address urban traffic congestion using a vehicle counting system. The system detected and tracked vehicles, with each detection visually represented to verify its accuracy. Tests under varying conditions revealed that the system had optimal accuracy during sunny days. However, we identified limitations of the system, especially under reduced visibility conditions such as nighttime, rainy days, or situations with low lighting. In these conditions, the system proved to be less effective, and even inoperative at night. Despite challenges, such as distinguishing darker vehicles and larger vehicles in certain conditions, the system's simplicity, coupled with its low operational cost, proved efficient. Emphasizing its real-world applicability, the study confirms that computer vision, when paired with the right tools, offers significant promise for urban traffic analysis.*

Key Words: *Artificial Intelligence, Computer Vision, YOLO Algorithm, Urban Traffic Analysis*

1 Introduction

Traffic is an important part of everyday life. It allows us to move around the world, meet other people, and connect with different cultures. However, traffic can also be dangerous. Every year, millions of traffic accidents occur, causing the death or injury of millions of people [5].

One way to improve traffic safety is through automated vehicle detection and counting. Traffic counting around the globe still relies on methods that utilize different sensors, such as inductive loop detectors [4], WIM (Weigh-In-Motion) systems [3] and systems powered by infrared sensors [6]. On the other hand, computer vision systems offer numerous advantages over traditional methods, as they can be adapted to various applications and adjusted to monitor various traffic parameters. In this article, we introduce prototype system that operates with the help of artificial intelligence, more specifically, using computer vision for vehicle recognition and traffic counting. Our system can be instrumental in decision-making concerning road infrastructure, as authorities need data about the traffic flow in specific segments. Our system can be affixed to traffic lights or overpasses, ensuring a clear view of the road section we aim to analyze [7].

The paper is organized as follows. Section 2 provides a brief review of literature on which our study is based. The problem definition is given in Section 3 and the research

methodology in Section 4. Section 5 presents the results, and Section 6 summarizes and discusses the findings and implications for further practice. Section 7 draws conclusions.

2 Literature review

In this chapter, we provide an academic overview of the literature on automatic vehicle detection using computer vision, emphasizing key findings and essential studies.

Study [8] explores real-time vehicle recognition using convolutional neural networks. Source [9] clarifies the relationship between computer vision and deep learning models. Two studies in developing countries, [10] and [11], use infrared and ultrasonic sensors to improve traffic flow and monitoring. Another study [2] discusses OpenCV's role in vehicle detection, offering key insights. [12] emphasizes camera placement and vehicle occlusion. In [1] YOLO achieves a 97.67% detection rate and 89.24% speed measurement accuracy in intersection vehicle recognition.

These studies highlight advancements in automatic vehicle detection via computer vision, stressing the importance of machine learning in traffic systems - a crucial aspect of our research for better vehicle detection and road safety.

3 Problem definition

Automatic vehicle detection via computer vision is critical in various applications, including traffic management, road safety, and autonomous driving. Computer vision allows computers to interpret visual data, enabling systems to recognize, track, and analyze vehicles, enhancing road safety and efficiency. One core aspect of automatic vehicle

detection is identifying vehicles using computer vision in images or videos. This includes classifying vehicles into different types, such as cars, trucks, bicycles, and more, facilitating a better understanding of road diversity. Despite advancements, challenges persist, such as ensuring algorithm robustness in diverse weather conditions and complex traffic scenarios. Precisely determining vehicle positions remains a challenge, requiring high accuracy for road safety.

In Slovenia, manual traffic counting, often conducted by students, remains common but inefficient due to human error and reliability concerns. This method lacks real-time integration with navigation systems for traffic flow improvement. The focus of this paper is on machine learning and computer vision, with the latter being used in the empirical part of the thesis.

4 Methodology

This section outlines the methodologies used in developing the vehicle recognition and traffic counting system, along with data collection for subsequent analyses. It focuses on creating a prototype system, the algorithm for vehicle detection, and data collection, capable of analyzing traffic patterns and identifying urban vehicles using artificial intelligence.

The system comprises a computer (Intel i7-10750H CPU @ 2.60GHz, Nvidia RTX 2060, 32GB RAM), a camera, and software. The camera streams live video to a server accessible via the computer, where machine vision tools and vehicle recognition algorithms, based on CNNs (convolutional neural networks) [1], analyze the feed.

The YOLO (You Only Look Once) [13] architecture, initially pretrained with 20 convolutional layers using ImageNet, is employed for vehicle detection. The model divides the image into an $S \times S$ grid, with each grid cell predicting B bounding boxes and confidence scores. NMS (Non-maximum suppression) [13] is applied to retain unique detected entities.



Figure 1: Areas in which we detect and count vehicles

The camera was initially positioned on a balcony overlooking Dunajska Street in Ljubljana, with video streamed to the server and accessed via the computer. Four regions were defined for traffic counting: two lanes in each direction (towards Črnuče and the city center). Areas 1 and 2 were designated for vehicles heading to the city center, while Areas 3 and 4 were for the Črnuče direction (see Figure 1).

In the Črnuče direction, when a vehicle enters Area 3, the program begins detection, marking the vehicle's center with a green dot (Figure 2a). As the vehicle crosses into Area 4,

it is counted, and a red rectangle outlines it (Figure 2b), indicating successful detection and counting. These visual features assist in program adjustments for factors like lighting and weather while helping identify system errors.

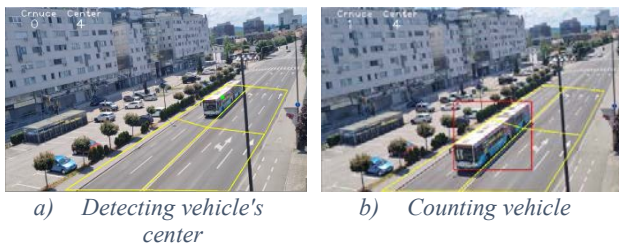


Figure 2: Example of a successful detection and counting of a vehicle

Each counted vehicle's data, including timestamp and direction, is saved in a .csv file for future statistical analysis.

5 Results

This chapter presents our study's traffic findings. The daily car average between 6 a.m. and 10 p.m. between 1st June and 28th July 2023 during the measurement period was 28,382. Major traffic peaks occurred from 7 a.m. to 9 a.m. in the morning and from 1 p.m. to 5 p.m. in the afternoon (Figure 3).

We validated measurements through one-hour manual counts at different times and weather conditions, assessing program overcounts and missed vehicles.

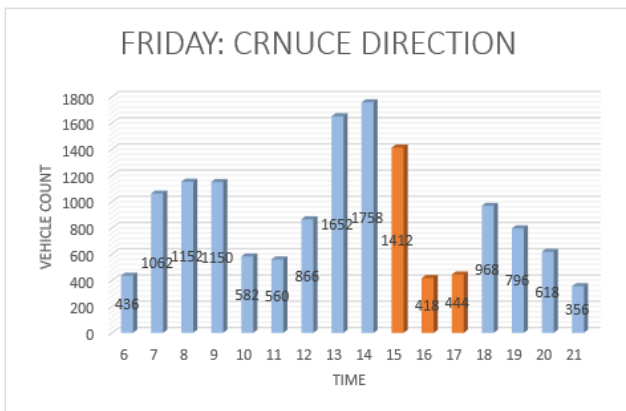


Figure 3: Example of traffic congestion on Friday x-time, y-count

Table 1 summarizes counting errors across various conditions and time frames.

Table 1: Identified counting errors

	Sunlight (n=923)	Twilight (n=512)	Rain (n=856)	Total Day (n=2291)	Night n=851
Overcounted	0	2	37	39	125
Undercounted	3	19	29	51	338
Total	3	21	66	90	463

The table segments errors by type and conditions: Midday - sunlight, Evening - twilight, and Rain - midday, representing specific scenarios affecting count accuracy. "Noon – Sun" had 3 errors (1 excess, 2 shortage) with a 0.33% error rate. Example of daytime monitoring in sunny conditions is shown in Figure 4.

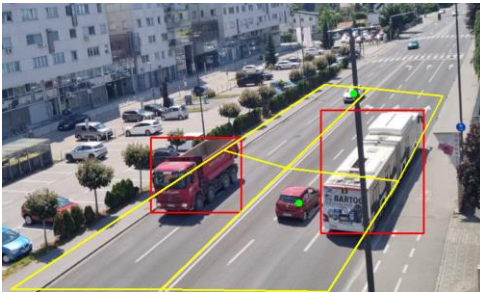


Figure 4: Example of daytime monitoring in sunny conditions

Daytime Results:

- Noon – Sun, $n = 923$ (3 errors): 0.33% error
- Evening – Dusk, $n = 512$ (21 errors): 4.1% error
- Rain – Midday, $n = 856$ (66 errors): 7.71% error

Total errors (excess + shortage), $n = 2291$ (90 errors): 3.93% error

- Total vehicles (excess – shortage), $n = 2291$ (12 vehicles counted less): 0.52% vehicles counted less

Nighttime testing involved one-hour manual counts:

- Total errors (excess + shortage), $n = 851$ (463 errors): 54.40% error
- Total vehicles (excess – shortage), $n = 851$ (213 vehicles counted less): 25.02% vehicles counted less

6 Discussion

The study's findings offer valuable insights into traffic patterns and highlight the potential of artificial intelligence in addressing urban traffic congestion challenges.

The highest accuracy was observed on sunny days with optimal lighting, resulting in a 0.33% counting error rate. Accuracy slightly declined during twilight (4.1% error rate) and further deteriorated in rainy conditions (7.71% error rate). In poorly lit nighttime conditions, the error rate surged to 54.40%. To provide context, a study [1] achieved a 97.67% vehicle detection accuracy under ideal conditions with a similar system and the same algorithm, affirming the suitability of the YOLO algorithm for object detection.

Diminished evening lighting particularly impacted the system's performance, with darker coloured vehicles posing challenges in differentiation from the dark background and asphalt. Rainy conditions sometimes led to false positives, especially for larger vehicles like trucks and buses, occasionally detecting them as two separate vehicles.

Analysing nighttime traffic proved challenging due to poor road section lighting and glare from bright headlights reflecting off the camera lens. Effective computer vision for traffic counting necessitates the combination of the right camera, object detection model, suitable road sections, and proper lighting.

The study results exceeded expectations, considering the system's simplicity, setup, and maintenance costs. Emphasis was placed on real-world applicability during equipment and software selection, prioritizing camera quality and computing power.

7 Conclusion

In this paper, we introduced a prototype traffic analysis system for larger cities based on artificial intelligence, utilizing the YOLO algorithm for computer vision and machine learning models to count vehicles and analyse their movement on Dunajska Street in Ljubljana. Our measurements demonstrated high accuracy and efficiency under daylight conditions. However, limitations emerged under reduced visibility, such as nighttime, rainy days, or poor lighting, rendering the system less effective and sometimes inoperative at night. Addressing these limitations requires model accuracy improvements and adjustments for varying lighting conditions.

As we foresee the next steps, tools like ChatGPT, developed by OpenAI, could be vital. They can offer real-time insights from traffic data, respond to user inquiries, and support data-driven decision-making. Integrating ChatGPT can enhance our system's features. However, it's crucial to ensure data privacy, avoid bias, and manage information effectively in urban traffic applications.

As AI continues to evolve, conversational AI models offer innovative solutions to urban traffic management and congestion challenges.

References

[1] Franklin, R. J. Traffic Signal Violation Detection using Artificial Intelligence and Deep Learning. In Proceedings of the V 2020 5th International Conference on Communication and Electronics Systems (ICCES), pages

839-844, Coimbatore, India, 2020.

[2] Uke, N; Thool, R. Moving Vehicle Detection for Measuring Traffic Count using OpenCV. *Journal of Automation and Control Engineering*, 1(4), 2013.

[3] Yang, X; Ahmad, S; Huang, Y; Lu, P. Automatic Vehicle Counting by Using In-Pavement Fiber Bragg Grating Sensor. In *International Conference on Transportation and Development 2022*, pages 225-234.

[4] Singh, N; Tangirala, A; Vanajakshi, L. A Multivariate Analysis Framework for Vehicle Detection from Loop Data under Heterogeneous and Less Lane Disciplined Traffic. *IEEE Access*, 2021.

[5] Chang, F. R; Huang, H. L; Schwebel, D. C; Chan, A. H. S; Hu, G. Q. Global Road Traffic Injury Statistics: Challenges, Mechanisms and Solutions. *Chinese Journal of Traumatology = Zhonghua Chuang Shang Za Zhi*, 23(4):216–218, 2020.

[6] Odat, E; Shamma, J. S; Claudel, C. Vehicle Classification and Speed Estimation Using Combined Passive Infrared/Ultrasonic Sensors. *IEEE Transactions on Intelligent Transportation Systems*, 19(5):1593-1606, 2018.

[7] Video-Based Vehicle Counting for Expressway: A Novel Approach Based on Vehicle Detection and Correlation-Matched Tracking Using Image Data from PTZ Cameras, 2020.

[8] Shrestha, A; Mahmood, A. Review of Deep Learning Algorithms and Architectures. *IEEE Access*, 7:53040-53065, 2019.

[9] Voulodimos, A; Doulamis, N; Doulamis, A; Protopapadakis, E. Deep Learning for Computer Vision: A Brief Review. *Computational Intelligence and Neuroscience*, 2018, Article ID 7068349.

- [10] Abdou, A. A; Farrag, M. H; Tolba, A. S. A Fuzzy Logic-Based Smart Traffic Management Systems. Journal of Computer Science, 18(11):1085-1099, 2022.
- [11] Akwukwaegbu, I. O; Mfonobong, E. B; Obichere, J.-K. C; Paulinus-Nwammuo, C. F. Smart Fuzzy Logic-Based Model of Traffic Light Management. World Journal of Applied Engineering and Technology Sciences, 8(2):108, 2023.
- [12] Zhang, R; Ishikawa, A; Wang, W; Striner, B; Tonguz, O.K. Using Reinforcement Learning With Partial Vehicle Detection for Intelligent Traffic Signal Control. IEEE Transactions on Intelligent Transportation Systems, 22:404-415, 2018.
- [13] Kundu, R. YOLO: Algorithm for Object Detection Explained. <https://www.v7labs.com/blog/yolo-object-detection> downloaded September 6th, 2023.

Detecting Corruption in Slovenian Public Spending from Temporal Data

Jelena Joksimović[†], Zoran, Levnajic^{†‡}, Bernard Ženko[‡]

[†]Faculty of Information studies

Ljubljanska cesta 31a, 8000 Novo mesto, Slovenia

[‡]Jožef Stefan Institute

Jamova cesta 39, 1000 Ljubljana, Slovenia

{jelena.joksimovic, zoran.levnajic}@fis.unm.si, bernard.zenko@ijs.si

Abstract. *Corruption is a pervasive issue with significant societal implications. One way to fight it is to establish mechanisms to analyse extensive data sets collected and published by governments. In this paper, we propose an innovative approach that primarily relies on unsupervised learning from time series data, to streamline the identification of suspicious transactions between public sector and private companies in Slovenia, interactively, with human involvement. This clustering technique primarily serves as a data labeling procedure, enhancing the efficiency and effectiveness of fraud detection. Our methodology is showcased using real-world data provided by the Slovenian Commission for the Prevention of Corruption, highlighting the potential for practical application in fraud detection scenarios. Apart from the methodology, we are also presenting the implementation of the framework and making it available to the scientific community to apply it on their own data sets.*

Keywords. corruption detection, time series analysis, clustering, interactive labeling

1 Introduction

Corruption is a widespread problem, often involving the misuse of public power for private gain [6]. Corruption at the intersection of the public and private sectors remains understudied, and includes opaque collusions and non-transparent deals [9].

To address this, collecting data on public-private transactions is essential, but compliance with privacy regulations is challenging. In Slovenia, the Commission for the Prevention of Corruption (KPK) is tasked with this responsibility [2, 3]. The KPK aggregates data from nine distinct sources, including the Ministry of Finance, the public procurement portal, and the Public Payments Administration [1]. Subsequently, this data is made accessible to the public via various applications on the *Erar* portal¹. The sheer volume of this data renders manual searches impractical. Collaborating with the KPK, we successfully obtained structured data on public-to-private transactions spanning 209 months. Furthermore, in partnership with the KPK, we identified a specific challenge related to

¹<https://erar.si>

corruption detection, which we have formulated as a time series analysis task. Transactions over time involving a specific private company can be effectively represented as a time series, enabling us to scrutinize shifts in its dynamics concerning events like government transitions. Exemplary instances of such dynamics are illustrated in Figure 1, sourced from the KPK.

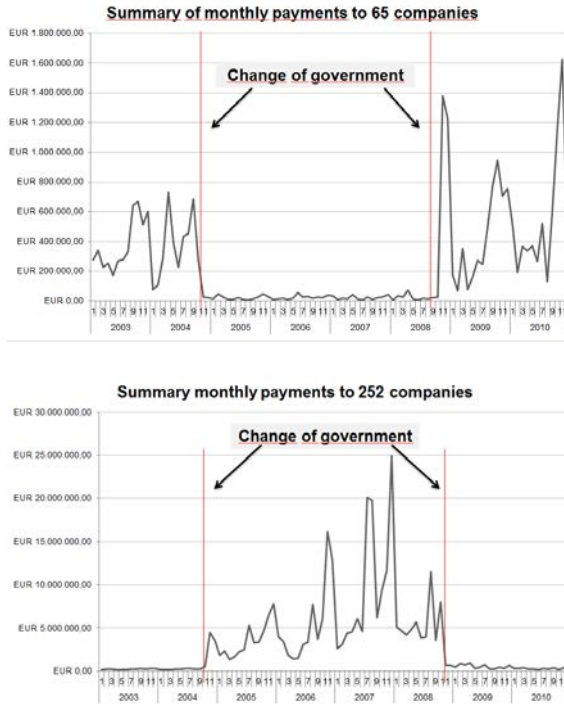


Figure 1: Illustration depicting two distinct sets of companies: The first group experiences a decline in public sector contracts during one government's tenure, while the second group secures new business opportunities. Source: Commission for the Prevention of Corruption in Slovenia.

2 Methods

Our method stems from the motivating examples depicted in Figure 1. We aim to automatically detect suspicious or "noteworthy" patterns within time series data representing a company's income from public sources. It's essential to emphasize that our objective is not to autonomously classify a company as corrupt or non-corrupt. Instead, our focus is on identifying companies that warrant specialized attention from qualified experts, such as officials from the Commission for the Prevention of Corruption (KPK). Subsequently, our framework adopts a "human in the loop" approach [7], and all final determinations are reserved for the judgment of the expert.

2.1 Data

Our data set, meticulously compiled by the KPK, encompasses all public-to-private sector transactions in Slovenia between January 2003 and May 2020. This comprehensive data set comprises records for a total of 248,989 companies. Each private company’s record includes monthly income (in euros) derived from public sources, including national and regional governments, ministries, educational institutions, hospitals, and more (for a comprehensive list, refer to [4]). Each company’s monthly income is represented as a time series. To mitigate potential bias, we intentionally anonymize the identities of the companies. Furthermore, we lack context-specific information about these companies, such as their type, ownership, contracts, partners, etc. Our primary objective is to construct a heuristic framework capable of identifying potentially corruptive companies within this data set with reasonable accuracy.

Analyzing time series data for *noteworthy* patterns² is inherently challenging. To simplify this task, we introduce a significant political event as a contextual reference for identifying these patterns. Our selection of events is guided by the examples presented in Figure 1, which showcase companies with markedly divergent income dynamics during one government. Specifically, during this period, 67 companies experienced a substantial decrease in government-related business, while 252 other companies flourished, exhibiting performance contrary to that observed before and after this government’s tenure. Thus, our quest for *noteworthy* patterns is contextualized within the framework of government transitions (cf. Figure 2). Over the span of our data set, Slovenia witnessed seven transitions of government (as detailed in Table 1).

From	To	Government	From	To	Government
12-2002	12-2004	ROP	03-2013	09-2014	BRATUŠEK
12-2004	11-2008	JANŠA	09-2014	09-2018	CERAR
12-2008	02-2012	PAHOR	09-2018	03-2020	ŠAREC
02-2012	03-2013	JANŠA	03-2020	06-2022	JANŠA

Table 1: Government transitions in Slovenia from 2002 to 2022, that represent key events in our analysis [8].

2.2 Framework overview

In this section, we introduce our framework designed to facilitate swift and precise identification of *noteworthy*, i.e., (potentially corrupt) examples. The framework encompasses the following steps:

Initial filtering. We begin by filtering the data set, removing all examples (companies) with a maximum monthly income below a predefined threshold. Subsequently, we narrow

²A note on the terminology: A pattern is considered *noteworthy* when there is a discernible shift in income dynamics (for a specific company) before and after a significant event, such as a government transition. A company is deemed *noteworthy* if it exhibits at least one such pattern. However, it is essential to reiterate that these labeled examples of companies should not be automatically construed as indicative of corruption. Instead, they serve as signals warranting further attention from KPK. As the opposite of the *noteworthy*, we have *regular* patterns. In this article, we will be using terms ‘noteworthy’ and ‘regular’ for patterns, companies (examples) or clusters in the context explained above.

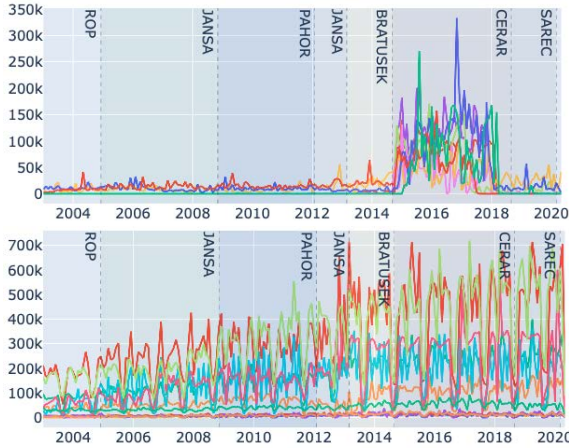


Figure 2: Illustration of labeled data. (a) *Noteworthy* examples - companies that experienced notably increased income from public bodies during the CERAR government. (b) *Regular* examples - companies demonstrating negligible income changes concerning government transitions.

down the time series data to encompass the specific interval centered around the critical event. Parameters such as the minimum income threshold, the pivotal event, and the interval width are expert-defined, based on their domain knowledge.

Time series clustering. The core objective here is to generate a set of clusters within the chosen event, each exhibiting distinct time trends that can be readily discerned by a human observer. While having the flexibility to employ various clustering methods, the choice also entails selecting the appropriate number of clusters and distance metric. Striking the right balance is essential, we opt for a sufficiently large number of clusters to capture diverse trends while ensuring the results remain manageable for expert scrutiny.

Visual inspection of clusters. Following the clustering phase, experts undertake a visual inspection of the trends, singling out the most captivating ones, those that encapsulate the most intriguing dynamics surrounding the event. However, the task extends beyond identifying only the *noteworthy* examples, equal attention is dedicated to spotting unremarkable clusters. Our future goal is namely to build a classification model capable of effectively distinguishing between these two categories of examples (more about this is explained in Section 4).

Labeling. After identifying both *noteworthy* and *regular* clusters, we proceed to visualize a predetermined number (e.g., 15) of examples within these clusters and manually label them. To ensure clarity and manageability, we limit the number of simultaneously visualized examples, as deciphering numerous trend lines within a single plot can be challenging. Additionally, we employ an interactive tool to search for and display the most similar examples to those already labeled, leveraging distance metrics. This iterative labeling process streamlines the classification of tens or even hundreds of examples efficiently. Ultimately, we curate a 'labeled examples' data set, aiming for an approximate equal split between *noteworthy* and *regular* labels.

3 Results and evaluation

To evaluate our framework, we utilized the complete KPK data set comprising 248,989 examples. Our parameters for the experiment are:

- Threshold: 1 000 000 EUR. We only choose companies with the total income of one million euros or above. Reasons for this can be found in [5].
- Event: Our focus was on the government transition that occurred in September 2014, transitioning from BRATUŠEK to CERAR. We analyzed data from the 6-month period before and after this transition, resulting in a 13-month time frame for clustering (March 2014 to March 2015).
- Clustering parameters: On the filtered data set obtained from the steps above (we were left with 9,246 companies after these steps), we applied the *k*-Means clustering algorithm with Euclidean distance and Z-Score normalization (per company), and received predefined 100 clusters. After testing multiple combinations of clustering algorithms and distance measures, we found these parameter values to be a good compromise regarding the interpretability of the results, computation time, memory requirements and clustering quality measures (e.g., silhouette score).
- Labeling: Out of *noteworthy* clusters, detected by the visual inspection, the human expert chose 15 examples from each one, and the system automatically detected 15 most similar examples (calculating the Euclidean distance and ranking the scores). Finally, the potential duplicates from the labeled data set had to be removed. The resulting labeling was inspected by a human expert.

With this procedure we collected a final data set³ of 966 companies, 483 labeled *regular* and 483 labeled as *noteworthy*.

Objective evaluation of the framework is, however, challenging. Our primary objective is to assess the efficiency gained by the experts when adopting this procedure in their workflow for identifying *noteworthy* examples. Further details on this aspect are discussed in the next section.

4 Conclusion and future work

In this paper we address a practically relevant problem of detecting potentially suspicious transaction patterns between public entities and private companies. We have developed an interactive framework that simplifies manual detection and labeling of *noteworthy* companies that authorities from KPK would ideally check further for the potential corruption involvement. The intended user of the framework is KPK, but it is open-sourced, so any expert that desires a quick dig into the dynamics of (hundreds of) thousands of time series in regard of some events, that he/she can precisely define, can use this framework for detecting and labeling the *noteworthy* patterns and examples. The next step in our investigation would be to enable training of a machine learned model that can detect such patterns automatically and evaluate the results in terms of performance measures and through manual inspection.

³Preview of all clusters and the labeled data set is available in *GitHub* repository: <https://github.com/jelenajoksa/CorrDetFramework>

5 Acknowledgements

We acknowledge support of the Slovenian Research Agency for financing this work through projects MR-Joksimović (53925) and research core funding for the programmes Complex Networks (No. P1-0383) and Knowledge Technologies (No. P2-0103). We also thank KPK for sharing the data and their challenges with us.

6 Reproducibility

The data and the code used for the analysis presented in this paper are available on *Google Colab*, url: https://colab.research.google.com/drive/1AaSnkAKNwvwEQLznXM8-NVP_Lxd9rDze?usp=sharing.

References

- [1] Commission, E.: Study on Corruption within the Public Sector in the Member States of the European Union. European Commission - Directorate-General Justice, Freedom and Security (2007)
- [2] CPC: About the Commission (2021), <https://www.kpk-rs.si/en/>, accessed: 2021-12-29
- [3] CPC: Commission for the Prevention of Corruption of the Republic of Slovenia (2021), https://en.wikipedia.org/w/index.php?title=Commission_for_the_Prevention_of_Corruption_of_the_Republic_of_Slovenia&action=history, accessed: 2021-12-28
- [4] eUprava: Javni sektor. Institucije države (2021), accessed: 2022-01-19
- [5] Joksimović, J., Perc, M., Levnajić, Z.: Self-organization in slovenian public spending. *Royal Society Open Science* **10**(8) (2023). <https://doi.org/https://doi.org/10.1098/rsos.221279>
- [6] Morris, S.D.: *Corruption & politics in contemporary Mexico*. University of Alabama Press (1991)
- [7] Mosqueira-Rey, E., Hernández-Pereira, E., Alonso-Ríos, D., Bobes-Bascarán, J., Fernández-Leal, Á.: *Human-in-the-loop machine learning: a state of the art*. Springer Netherlands (2022). <https://doi.org/10.1007/s10462-022-10246-w>
- [8] Office of the Government of the Republic of Slovenia for Communication: *Pretekle vlade* (2022), <https://www.gov.si/drzavni-organi/vlada/o-vladi/pretekle-vlade/>, accessed: 2022-01-26
- [9] Sotola, D., Pillay, P.: Private sector and Public Sector Corruption nexus: A synthesis and typology. *Administratio Publica* **28**(1), 113–135 (2020)

Mathematica and art: designs with RegionPlot

Biljana Jolevska Tuneska

Faculty of Electrical engineering and Informational Technologies
Rudjer Boskovic 18, 1000 Skopje, North Macedonia
{biljanaj}@feit.ukim.edu.mk

Abstract. *Mathematical principles have always played a crucial role in the realm of artistic creation. Throughout history, symmetry served as a fundamental tool in crafting textiles, ethnic patterns, and architectural wonders. During the Renaissance era, artists found it necessary to employ or develop mathematical reasoning to realize their artistic visions. Prominent figures among these artists included Luca Pacioli (1447-1517), Leonardo da Vinci (1452-1519), and Albrecht Durer (1471-1528).*

In today's context, mathematical tools have evolved to become more advanced, with digital technology emerging as a primary choice. Artists now harness the power of computers to produce art. Isometries, similarities, and affine transformations allow for precise or purposeful distortion of images, while projections enable the representation of three-dimensional objects on two-dimensional surfaces. These transformations can be precisely described mathematically, and the traditional use of guiding grids to aid in performing these transformations has largely been supplanted by computer software.

In this paper, we employ the mathematical program Mathematica to showcase various artistic designs.

Keywords. mathematical tools, art

1 Introduction

From a highly pragmatic standpoint, mathematical instruments have consistently played a vital role in the artistic creative process. Throughout history, humble tools like the compass and straightedge, bolstered by additional basic implements employed by draftsmen and craftsmen, have been instrumental in crafting exquisite designs that found expression in the architectural embellishments of palaces, cathedrals, and mosques since time immemorial, see [2], [3]. Here we will consider only the Golden Ratio in Art Composition and Design, and the work of the Dutch artist Maurits Cornelis Escher.

1.1 Golden Ratio in Art Composition and Design

Phi, also known as the Golden Ratio, manifests itself everywhere in the fabric of life and the cosmos, see [1]. Its presence gives a profound sense of equilibrium, concord, and beauty into the patterns we found in nature. Human beings have harnessed this very proportion, sometimes deliberately and at other times intuitively, to attain a sense of symmetry, harmony, and aesthetics in various realms such as art, architecture, design, and composition.

Much like how the Golden Section naturally appears in the aesthetics and elegance of the natural world, it serves as a versatile tool, rather than an inflexible rule, for crafting compositions. Gaining proficiency in its application can offer valuable insights into the art of arranging a painting on a canvas.

For those possessing a deeper comprehension, the Golden Ratio can be employed in more sophisticated ways to craft aesthetics and visual coherence across diverse branches of the design arts. As exemplified below, Leonardo Da Vinci extensively utilized the Golden Section in his works. Observe how all the fundamental dimensions of the room, the table, and ornamental shields in Da Vinci's masterpiece "The Last Supper" were rooted in the Golden Ratio. The lines in the artwork unveil Da Vinci's intricate deployment of the Golden Ratio.

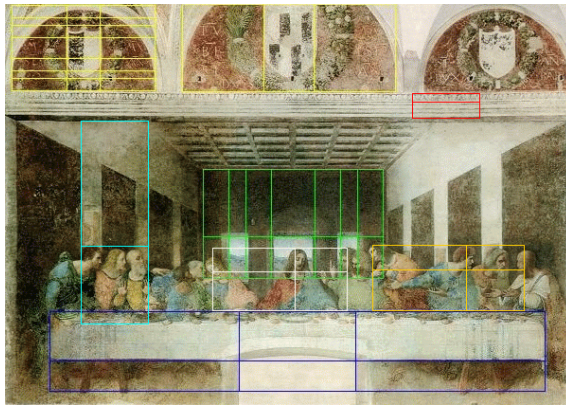


Figure 1: The Last Supper

But not only back in the 15th century, the Golden Ratio is widely used in today's art. I am presenting you two logo's of a well known successful companies:

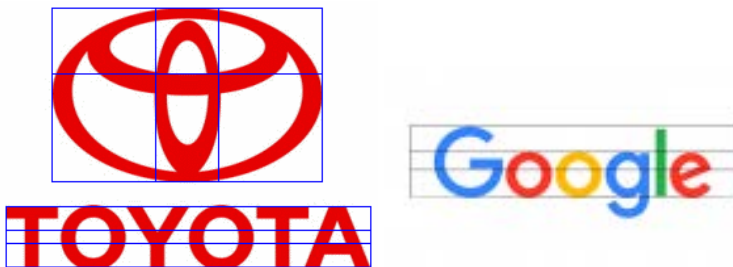


Figure 2: Today's well known logo's

1.2 The Art of Escher

Maurits Cornelis Escher (1898-1972) stands as one of the world's preeminent graphic artists, enjoying admiration from countless individuals globally, a fact evident in the multitude of websites dedicated to his work, see [4]. He achieved widespread fame for his iconic "impossible drawings," such as "Ascending and Descending" and "Relativity," as

well as his captivating metamorphic works, including "Metamorphosis I, II, and III," "Air and Water I," and "Reptiles."

In the course of his lifetime, Escher crafted a staggering 448 lithographs, woodcuts, and wood engravings, accompanied by over 2000 drawings and sketches. What set Escher apart was his ability to bridge the divide between art and science, melding intricate mathematics with precise draftsmanship. His body of work marries intricate realism with fantastical elements, with his renowned "impossible constructions" being prime examples artworks that ingeniously employ mathematical shapes, architectural elements, and perspective to craft visual puzzles. Notably, his art primarily took the form of prints, encompassing lithographs and woodcuts, exhibiting a distinct appearance and subject matter that defied the prevailing norm of abstract art during his era.

Despite not having a formal mathematical training, Escher had an intuitive and nuanced understanding of the discipline. He used geometry to create many of his images and incorporated mathematical forms into others. Additionally, some of his prints provide visual metaphors for abstract concepts particularly that of infinity, the depiction of which Escher became interested in later in his career. During his lifetime Escher kept abreast of current ideas in the field and corresponded with several eminent mathematicians on the subjects of interconnecting and impossible shapes incorporating their ideas directly into his work.

Escher highlighted the contradiction of representing three dimensional objects on a two-dimensional plane and this is particularly clear in images such as Drawing Hands (1948) in which two hands (seemingly simultaneously) engage in the paradoxical act of drawing each other into existence.

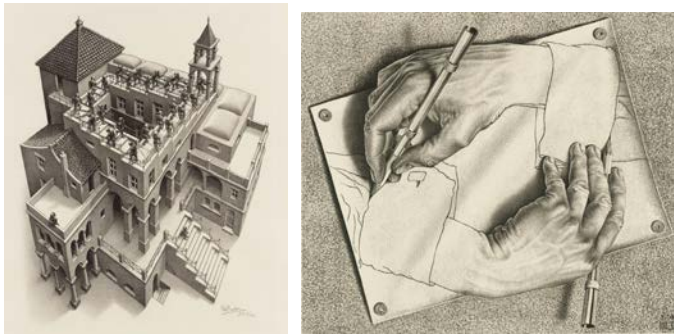


Figure 3: Some of Escher's work

2 Mathematica and art

Wolfram Mathematica, is a powerful computational software system that offers a captivating avenue for the creation of art that seamlessly merges mathematics, creativity, and technology. While Mathematica is renowned for its extensive capabilities in fields like mathematics, physics, and data analysis, it also serves as an exceptional tool for artists seeking to express their vision in unique and mathematically inspired ways.

In this digital age, artists can harness Mathematica's rich set of functions and libraries to delve into the realm of generative art, algorithmic design, and visual experimentation.

This software empowers artists to transform mathematical concepts, equations, and algorithms into captivating visual representations, giving rise to a dynamic and evolving form of artistic expression.

Whether you are intrigued by fractals, parametric curves, 3D graphics, or even data-driven art, Mathematica offers an array of tools and functions that can be employed to create stunning and thought-provoking artworks. With its user-friendly interface and the ability to seamlessly integrate mathematical computations into artistic projects, Mathematica opens up new horizons for artists to explore, enabling them to push the boundaries of their creativity and produce art that is not only visually compelling but also intellectually engaging.

In this exploration of Mathematica's potential as a tool for artistic creation, we will delve into some of the captivating ways in which artists can utilize its features to bring their artistic visions to life.

2.1 Designs with RegionPlot

Wolfram Mathematica's RegionPlot function provides an exciting avenue for artists and designers to create captivating and visually striking art designs. RegionPlot allows you to graphically represent regions defined by mathematical inequalities or logical conditions, transforming abstract mathematical concepts into visually appealing artwork. By harnessing the power of Mathematica, you can explore the limitless possibilities of geometric patterns, abstract compositions, and algorithmically generated designs. In this artistic journey, we will delve into how you can leverage RegionPlot to craft mesmerizing art designs and provide examples to inspire creativity.

In the following we are giving some directions for further art exploration.

1. Geometric Abstraction: We begin with the basics by experimenting with geometric abstraction. We use RegionPlot to depict circles, squares, or other basic shapes in unique and artistic arrangements. For instance, one can create a mesmerizing pattern of concentric circles using the Mathematica code in Figure 4. Then we can adjust the parameters and styling options to achieve the desired visual effect.

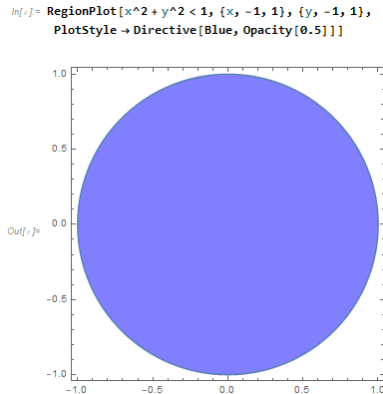


Figure 4: Geometric Abstraction

2. Fractal Art: We can dive into the world of fractals, where self-similar patterns create intricate and mesmerizing designs. Using RegionPlot, one can render famous

fractals like the Sierpinski Triangle or the Mandelbrot Set. An example of generating a Sierpinski Triangle is given in Figure 5. Further we can tweak the parameters to explore various fractal forms and complexities.

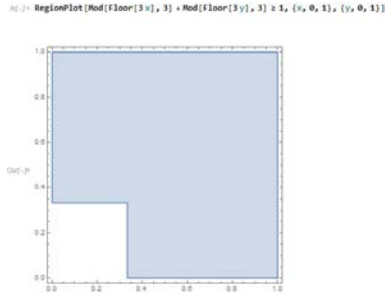


Figure 5: Fractal Art

3. Abstract Art with Inequalities: Embrace abstract art by defining mathematical inequalities and create unique and expressive compositions. One can combine multiple RegionPlot instances to layer patterns and shapes, achieving visually stimulating results. For instance, we can create an abstract artwork with a blend of circles and squares, see Figure 6. We can experiment with colors, opacities, and shapes to craft your own abstract masterpiece.

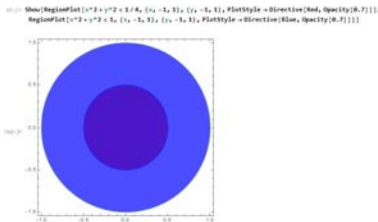


Figure 6: Abstract Art with Inequalities

4. Mathematical Concepts as Art: Turn mathematical concepts into visually appealing art. For example, we can create a captivating piece by representing prime numbers as shaded regions within a grid, see Figure 7. This transforms a mathematical concept into a thought-provoking and artistic composition.

By exploring Mathematica’s RegionPlot function, you can unlock endless possibilities for crafting art designs that blend mathematics, creativity, and aesthetics, providing a unique platform for your artistic expression. The examples provided here are just the beginning of your artistic journey with Mathematica.

At the end of this paper we give an example how we can merge all of previous items in one, and make interesting designs as the following. In Figure 8 the code on Mathematica is given, and in Figure 9 we have the output design.

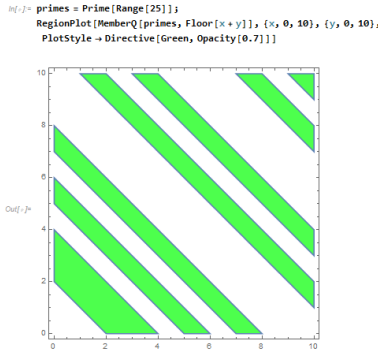


Figure 7: Mathematical Concepts as Art

```

RegionPlot[(x - 1)^2 + (y + 3)^2 < 0.82 || (x - 3)^2 + (y + 3)^2 < 0.82 || (x - 3)^2 + (y - 1)^2 < 0.82 || (x + 1)^2 + (y - 3)^2 < 0.82 || (x + 3)^2 + (y + 1)^2 < 0.82 || (x + 3)^2 + (y - 3)^2 < 0.82 || (x - 1)^2 + (y - 3)^2 < 0.82 || (x - 3)^2 + (y - 1)^2 < 0.82 || (x + 1)^2 + (y - 1)^2 < 0.82 || (x - 1)^2 + (y + 1)^2 < 0.82 || (x - 1)^2 + (y - 3)^2 < 0.82 || (x - 3)^2 + (y - 1)^2 < 0.82 || (x + 1)^2 + (y - 3)^2 < 0.82 || (x + 3)^2 + (y + 1)^2 < 0.82 || (x + 3)^2 + (y - 3)^2 < 0.82, {x, -4, 4}, {y, -4, 4}, BoundaryStyle -> None, Mesh -> 7, MeshShading -> {{Darker[RGBColor[0.1, 0.9, 1], 0.5], Darker[RGBColor[0.1, 0.9, 1], 0.5]}, Gray, Gray, Darker[RGBColor[0.1, 0.9, 1], 0.5]}, Darker[RGBColor[0.1, 0.9, 1], 0.5]}, Gray, Gray}, {Lighter[RGBColor[1, 0.5, 0], 0.2], Lighter[RGBColor[1, 0.5, 0], 0.2]}, Black, Black, Lighter[RGBColor[1, 0.5, 0], 0.2], Lighter[RGBColor[1, 0.5, 0], 0.2]}, Black, Black}, {Gray, Gray, Darker[RGBColor[0.1, 0.9, 1], 0.5]}, Darker[RGBColor[0.1, 0.9, 1], 0.5]}, Darker[RGBColor[0.1, 0.9, 1], 0.5]}, Darker[RGBColor[0.1, 0.9, 1], 0.5]}, {Gray, Gray, Lighter[RGBColor[1, 0.5, 0], 0.2], Lighter[RGBColor[1, 0.5, 0], 0.2]}, Gray, Gray, Lighter[RGBColor[1, 0.5, 0], 0.2], Lighter[RGBColor[1, 0.5, 0], 0.2]}, {Lighter[RGBColor[1, 0.5, 0], 0.2], Lighter[RGBColor[1, 0.5, 0], 0.2]}, Gray, Gray, Lighter[RGBColor[1, 0.5, 0], 0.2], Lighter[RGBColor[1, 0.5, 0], 0.2]}, {Darker[RGBColor[0.1, 0.9, 1], 0.5], Darker[RGBColor[0.1, 0.9, 1], 0.5]}, Black, Black, Darker[RGBColor[0.1, 0.9, 1], 0.5], Darker[RGBColor[0.1, 0.9, 1], 0.5]}, Black, Black, Lighter[RGBColor[1, 0.5, 0], 0.2], Lighter[RGBColor[1, 0.5, 0], 0.2]}, Black, Black, Darker[RGBColor[0.1, 0.9, 1], 0.5]}, Darker[RGBColor[0.1, 0.9, 1], 0.5]}, {Black, Black, Lighter[RGBColor[1, 0.5, 0], 0.2], Lighter[RGBColor[1, 0.5, 0], 0.2]}, Black, Black, Lighter[RGBColor[1, 0.5, 0], 0.2], Lighter[RGBColor[1, 0.5, 0], 0.2]}], MeshStyle -> {None, GrayLevel[0.2]}, Background -> GrayLevel[0.8], Frame -> False}

```

Figure 8: Code in Mathematica

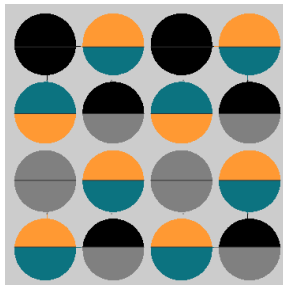


Figure 9: Design in Mathematica

3 Conclusion

The intricate relationship between mathematics and art has been a cornerstone of artistic expression throughout history. From the foundational use of symmetry in ancient designs to the Renaissance masters employing mathematical reasoning, the two disciplines have been intertwined.

In the modern era, this relationship has been further amplified with the advent of digital technology, particularly through tools like Wolfram Mathematica. The software not only facilitates the representation of complex mathematical concepts in visual forms but also offers artists a platform to experiment, innovate, and express their visions in unprecedented ways. The RegionPlot function in Mathematica exemplifies this, allowing artists to transform abstract mathematical ideas into tangible, visually appealing artworks.

Further work on this topics can investigate other functions like GeometricTransformation, or RotationTransform from Wolfram Mathematica to create inspiring and creative

artworks.

References

- [1] G. Maisner, *The Golden Ratio: The Divine Beauty of Mathematics*, Quatro's Race Point Publishing Group, 2018.
- [2] D. Schattschneider, *Mathematics and Art: So Many Connections*, <http://www.mathaware.org/mam/03/essay3.html>, April 2003.
- [3] J.J.O'Connor, E.F.Robertson, "Mathematics and art: perspective", 2003, <http://www-history.mcs.st-and.ac.uk/HistTopics/Art.html>
- [4] <https://mcescher.com/>

Diagnosing depression from physiological data using AI

Albert Zorko, Zoran Levnajic

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

Abstract: *Depression is a major burden on our society. It reduces both the quality of life and the capacity to work for millions around the globe. Diagnosing depression is traditionally done via psychiatric tests and scales. However, the accuracy of such diagnosis can be subjective, since some tests leave room for varying interpretations. This calls for diagnostics that is grounded in physiological variables such as heart rate and breathing. They are simple to measure, and the measured values are far more objective. If depression could be diagnosed from them – even if imperfectly or partially – it would offer an important new tool for psychiatry. This paper has two aims. First, we briefly review the existing literature and key results in this direction. Second, we use Machine Learning (ML) to examine a labelled dataset of depressed and healthy subjects, whose heart and respiratory rates were precisely recorded. Results of this analysis, albeit preliminary, are promising and worthy of further research.*

Key Words: *depression, heart rate, respiratory rate, machine learning, artificial intelligence*

1 Introduction

Depression often begins in adulthood (National institute of Mental Health, 2023), yet recent studies show that rates are rising, and that depression often begins in childhood. As many as 3% of children and 8% of adolescents in the U.S. have depression (Brennan, 2022).

There are seven warning signs that are most diagnostic of impending depression: decreased interest in or enjoyment of activities, depressed mood, hopelessness, decreased concentration, decreased self-esteem, worry/moodiness, and irritability (Rottenberg, 2010).

The earliest written accounts of what we now consider depression appeared in Mesopotamia 2 000 years BC (Schimelpfening, 2023). In these accounts, depression is seen as a spiritual rather than a physical condition. It was dealt with by priests rather than doctors. In many cultures, depression was believed to be caused by demons and evil spirits and was treated with methods such as: beatings, starvation and physical restraint. All methods aimed at exorcising demons (Schimelpfening, 2023). Hippocrates¹ developed the theory of the four bodily juices - the so-called humoral theory. An imbalance between the body juices (blood, mucus, yellow bile and black bile) leads to

¹ 460 - 370 BC

disease. Depression (then called melancholia²) is caused by an excess of black bile (Hoppe, 2019; Radden, 2002). The Persian physician Rhazes³ believed that the disease originated in the brain. He recommended baths and behavioral therapy as treatment, as well as rewards for appropriate behavior (Schimelpfening, 2023). Avicenna or Ibn Sīnā⁴, who gave melancholia quite a lot of attention with his medical encyclopedia *The Canon of Medicine* (Clarke & O'Malley, 1996) - a standard medical text in medieval universities. He divided the disease into an early and a chronic phase (Pies, 2020). The early phase contained symptoms: suspicion of evil, fear for no reason, involuntary muscle movements, dizziness, tinnitus and quick anger. The chronic phase, on the other hand, has symptoms of sadness, suspiciousness, moaning, restlessness, and 'abnormal' fear. Importantly, Avicenna identified a clinical identity that is now referred to as a mixed state (Ghaemi & Mauer, 2020).

The term depression appears in the 1680s and is used to mean "a deterioration of affairs", which is not necessarily related to the mind and body (Rousseau, 2000). Johann Christian Reil⁵, professor of medicine at the University of Halle, introduced psychiatry as a medical specialty, the title of psychiatric hospital and psychotherapy as a therapeutic method on a par with surgery and pharmacotherapy (for mental and somatic illnesses) (Marneros, 2008). He also stressed the need to examine the potential criminal liability of mentally ill persons. In 1856, the French psychiatrist Louis Delasiauve first used the term depression in connection with a psychiatric symptom. In 1860, however, the word appears in medical dictionaries in connection with a physiological and metaphorical decline in emotion.

Pharmacological treatment (in addition to psychotherapy) appears in the 1950s. First appeared the first-generation antidepressants (Imipramine, the first tricyclic antidepressant) (Abdallah, Sanacora, Duman, & Krystal, 2018). During this period, Iproniazid, the first monoamine oxidase inhibitor (MAOI⁶) for depression, was also discovered (Blackwell, 2011). The next period covers the 1970s. This period is marked by the development of selective serotonin reuptake inhibitors (SSRIs) and the evidence of their efficacy in the treatment of depression (Murphy, et al., 2021). The first antidepressant in the class of selective serotonin reuptake inhibitors (SSRIs) is Fluoxetine, known as Prozac (Wenthur, Bennett, & Lindsley, 2014). The 1990s are characterised by the development of selective serotonin-norepinephrine reuptake inhibitors (SNRIs) in the treatment of various disorders, including depression (Gorman & Justine, 1999). Among others, Duloxetine is widely used under the name Cymbalta.

In addition to the development of various pharmacological agents for the treatment of depression, understanding the role of neurotransmitters in the pathophysiology of depression is of paramount importance. An important question in the study of this area is whether receptor antagonists can improve the symptoms seen in depression (Abdallah, Sanacora, Duman, & Krystal, 2018).

A diagnosis of this mental disorder is needed before treatment can begin. This now consists of a physical examination, laboratory tests, a psychiatric assessment, and the

² feelings of sadness, apathy and other symptoms now associated with depression.

³ 865 - 925 AD

⁴ 980 - 1037 AD

⁵ 1759 - 1813

⁶ Monoamine oxidase inhibitor

DSM-5 questionnaire (Mayo Clinic, 2022). The physical examination also includes health questions that the doctor asks the patient, as depression may be associated with a physical health problem. In a laboratory test, the doctor may decide to do a complete blood count or a thyroid function test to see if the thyroid is functioning properly. In a psychiatric assessment, a mental health professional obtains information about thoughts, symptoms, emotions, and behaviour patterns. The final step involves using the internationally recognised criteria for depression from the Diagnostic and Statistical Manual of Mental Disorders DSM-5 (American Psychiatric Association, 2023). The first version was published in 2013 by the American Psychiatric Association (APA). The manual, which is used mainly in the USA, guides the physician in establishing a hierarchy of diagnoses (from most important to least important). In other countries, the International Classification of Diseases (ICD), published by the World Health Organisation (ICD-11 version⁷) (World Health Organisation, 2022), is used. Mental disorders have their own chapter in this classification. Depressed individuals point to bodily sensations that indicate changes in the autonomic nervous system. These changes refer to a decrease in sympathetic and parasympathetic nervous system activity (Goodwin & Jamison, 2007).

2 Methods

2.1 Diagnosing depression using physiological data

As the diagnosis of depression is more subjective with existing methods, there is interest in methods using biomarkers such as HRV to obtain more objective results. The remainder of this chapter lists the main results in this field.

The ability to detect depression from linear and non-linear heart rate features during a mental task protocol using machine learning was described by Sangwon Byun and his team of authors (Byun, et al., 2019). They achieved 74.4% accuracy and 73% sensitivity. A model for depression detection based on heart rate variability parameters and the effect of therapy on the parameters (Xing, et al., 2019) uses a method in which the ECG is measured before treatment and then again at three points during the treatment itself. Statistical t-test to examine the significant HRV parameters of the depressed and control groups. They use a support vector machine (SVM) in addition to the characteristic HRV parameters to build a detection model. They achieved 89.66% accuracy in detecting depression. Using resting heart rate variability and skin conductance response to detect depression in adults (Smith, et al., 2020) is a method in which they obtain 3-minute resting heart rate and skin conductance response (SCR) data with eyes closed. Different classifiers were used for the training set. SCR performed better than for depression detection, with a strong correlation with suicidal tendencies. Using HRV alone, they achieved 81% test accuracy. However, when HRV and SCR were used, only 78% accuracy was achieved. In the article Smart Devices and Wearable Technologies to Detect and Monitor Mental Health Conditions and Stress: A Systematic Review (Hickey, et al., 2021), the authors present a variety of smart devices designed to identify depression, anxiety, and stress. They use HRV to identify the device, and better results have been achieved using HRV and EEG together (Ahn, Ku, & Kim, 2019). However, it would be appropriate to investigate the applicability of the dual-device method for long-term stress monitoring. Author Hickey and others also state that there are currently no EEG devices

⁷ Effective from 1.1.2022

available on the market but mostly use heart rate. However, devices using the average heart rate are not as reliable in detecting depression as those using HRV, electrodermal activity (EDA) and possibly respiratory rate. The EDA method is based on changes in electrical parameters on the skin (Boucsein, 2012). These changes may have a psychological significance, as when an emotional stimulus is given, the endocrine glands produce sweat, which improves the conductivity of the electrical current. The greater the stimulus, the greater the sweat secretion and consequently the better the skin conductivity (Stern, Ray, & Quigley, 2001). A very promising method is suspected by a group of authors who state that the method is problematic due to the physical activity of the individual and the consequent movement of the sensors on the skin (Chen, Abbod, & Shieh, 2021). Detecting depression using wearable devices remains a challenge. Currently in use are EEG monitors (Li, Hu, Shen, Xu, & Retcliffe, 2015) and accelerometers (Narziev, et al., 2020)

Challenges of depression: One of the serious problems associated with mental health is the lack of consensus on classification, diagnosis and treatment, resulting from an incomplete understanding of the processes of these disorders (Vivekanantham, Strawbridge, Rampuri, Ragunathan, & Young, 2016). This is evident in mood disorders - major depressive disorder (MDD), which is a complex disorder in which as many as 60% of patients experience some degree of treatment resistance. This results in prolonged and worsening episodes (Fava, 2003). This fact suggests the need for a robust method to define subtypes within diagnostic categories and thus to refine treatment (Insel, et al., 2010).

Depressed patients without anxiety, especially those who show signs of psychomotor arrest, have reduced sympathetic tone, resulting in reduced tonic and phasic skin conductance levels (Christie, Little, & Gordon, 1980). However, if depression is associated with anxiety and agitation, it can cause an increase in heart rate, blood pressure and blood flow in the forearm, indicating sympathetic activation (Zahn, 1986).

An additional challenge is posed by the fact, cited in an article (Čukić, Savić, & Sidorova, 2023), that psychiatrists and psychologists are not very enthusiastic about the use of objective methods using the measurement of physiological parameters in daily diagnosis (Porges, 2011). This, of course, raises a number of ethical questions, e.g., until when will the subjective method continue to be in use, if there are methods that offer an unbiased diagnosis?

In this paper, we take a preliminary step in this direction by having physiological data (HRV and respiration) for healthy and depressed subjects. We will process the data using AI and look at the results. We will publish the full study below.

2.2 Subject

The data was collected to study mental disorders in humans. Respiration and ECG were measured in 73 participants, 26 males (age 19-52 years, mean \pm SD: 34.5 \pm 10.5) and 47 females (age 16-60 years, mean \pm SD: 32.7 \pm 11.2). All participants were classified into three groups: a control group, patients diagnosed with depression and patients treated for depression with tricyclic antidepressants.

The control group consisted of 14 males (age 20-52 years: 36.6 ± 10.4) and 24 females (age 19-60 years: 32.8 ± 11.0). The depression group consisted of 6 males (age 19-36 years: 24.8 ± 6.3) and 11 females (age 16-37 years: 25.5 ± 7.5). The treatment group consisted of 6 males (age 27-51 years: 39.3 ± 9.0) and 12 females (age 20-60 years: 39.0 ± 11.0).

Depression in the patients was diagnosed by means of a questionnaire. For the measurements, these patients were divided into two groups: those adequately treated with tricyclic antidepressants and those not yet treated with antidepressants.

2.3 Measurement and data analysis

The whole group of participants had their heart rate measured using a chest-mounted ECG and respiration measured using a nasal sensor (thermistor), which allows fast and dynamic measurement. The measurement was performed in the supine position to allow the individual to calm down before the measurement. The large differences in HRV between sleep and wakefulness are also shown in a paper (Zorko, Frühwirth, Goswami, Moser, & Levnajić, 2020). The heart rate and respiration data are stored at a sampling rate of 200 Hz and at a resolution of 10 bits. For the purposes of the study, a temporal resolution of 1ms for the detection of the peak of the R wave and 5ms for the detection of the onset of inspiration is sufficient. For each individual, the time elapsed from the last beat (Rpeak value) to the start of respiration is calculated. To avoid subjective data, we recalculated the absolute value of this time against the total elapsed time between two beats and obtained a relative dependence. The total observation time of the measurement values is limited to 9 minutes. The following is the analysis procedure.

The data obtained were processed and a histogram of the calculated values was plotted for each individual. The results differ for each group and are shown in Figure 1.

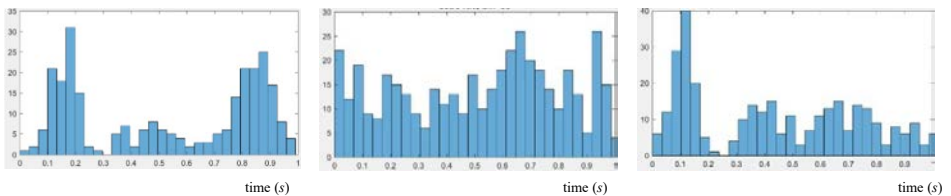


Figure 1: histograms of the calculated inspiratory delay times versus the previous heart rate. From left to right: a 28 year old from the control group, a 30 year old from the depressed group and a 27 year old from the treated depressed group.

Using this technique, it can be seen that three peaks appear in the control group. The reason for this is that at rest, in healthy subjects, there is a coupling between respiration and heart rate, which is called pulse-respiratory coupling (PRC) (Kralemann, et al., 2013). In depressed patients, no coupling is observed, and the inspiration appears to start completely randomly with respect to the heart rate (Rpeak value). A good PRC represents a higher degree of autonomic coordination, but the functions underlying PRC are not yet fully understood. However, it appears that PRC is a measure reflecting the centrally controlled functional tone of the autonomic nervous system (ANS) in its polarities of

ergotropy and trophotropy (Matić, et al., 2022). In depressed individuals, there is apparently a lower degree of autonomic coordination and, consequently, a lower pulse-respiratory coupling (PRC), which is reflected as a completely random onset of inspiration relative to the heart rate. The question is: What about the treated individuals? It would be difficult to put all those treated in the same category, as the length of treatment and the exact diagnosis are extremely important. In the case of the treated depressed person in Figure 1 (right figure), we can see that the pulse-respiratory coupling is improved, as the outlines of the three peaks in the histogram are already visible.

Having finished examining the data, we set about detecting depressed individuals from the dataset. To do this, we first had to create a data model. To organise the data model, we used the following data for the participants: gender, age, healthy/unhealthy, treated/untreated. The data model organised in this way was used as input to Weka, an open-source software for data pre-processing, machine learning implementation and visualisation (WekaIO Inc., 2023).

As the data was already cleaned and edited, no pre-processing was necessary. As Weka offers several different machine learning methods, we decided to use only the random forest method in the preliminary study, which gave promising results. In parallel, we also tested the same data model with the predictive clustering trees method built by ClusPlus. The user interface for this tool is ClowdFlows, which was provided to us by the Jožef Stefan Institute for the purposes of the study. The ClusPlus tool allows the construction, execution and sharing of interactive workflows and the modification of various parameters of the data model and the data mining kernel. The most important features of this tool are a web-based user interface for building and managing workflows, a cloud-based architecture, a service-oriented architecture, a large list of workflow components, and a real-time processing module (Jožef Stefan Institute, 2016). The tool has a service-oriented architecture that allows for parallelisation, remote execution, high availability, access to large public (and proprietary) databases and easy integration of third-party components.

3 Results

Using the Weka tool, one decision tree (`weka.classifiers.trees.J48 -C 0.25 -M 2`) with 89.04% of the data correctly classified between healthy and depressed. Using a random forest ensemble (`weka.classifiers.trees.RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1`), the correct classification improved to 95.89%.

Different observation intervals were tested using ClusPlus and the ClowdFlows user interface. We tested from just five data points to all data points captured for an individual. We found that after six minutes of observation, the data was always correctly classified between healthy and depressed individuals. We have not yet determined why, in the decision tree shown in Figure 2, the separation between treated and untreated occurs very low in the tree. This fact is most likely due to the fact that the pattern of the time delays of the onset of inspiration from heart rate differs very little from depressed subjects. Another unknown is the difference between males and females. Contrary to our expectations, in healthy subjects the separation to females occurs only in the fourth level,

and in depressed subjects lower than the fourth level. Apparently, age plays a more important role than gender.

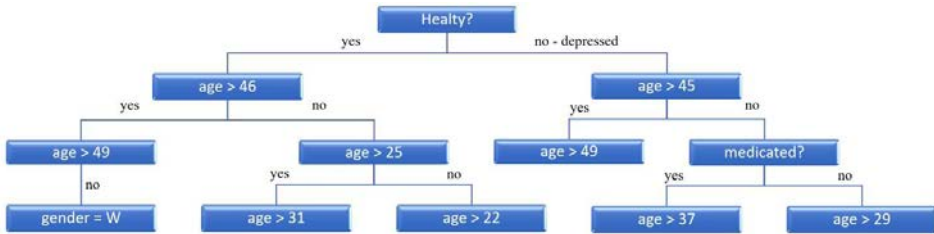


Figure 2: Decision tree generated for the data model by the ClusPlus tool with the CloudFlows user interface.

4 Discussion and conclusion

Depression is increasingly becoming an illness faced by a very large number of people. Consequently, it has a very large impact on the working population, ranging from sickness absenteeism from work or school to reduced work performance. The first problem starts with the diagnosis of the disease, which currently relies on the judgement of a doctor using a general examination, a blood test, and a questionnaire to identify mental disorders. This approach can too often be subjective, resulting in misclassification and, consequently, in incorrect treatment.

The overwhelming desire for a non-subjective method is correspondingly quite strong, as can be seen in many of the methods mentioned in the article. We have also tried to add our own little stone to the mosaic of methods, with quite encouraging results. As the study is still at a preliminary stage, we are still waiting for experiments with the different methods offered by the tools we have been working with in order to get even better results. At the same time, we wonder if it is possible to obtain similar results with other analytical tools, thus confirming the basic research question that it is possible to use a data mining tool on non-invasively acquired physiological data of a person. The results obtained are more than promising, with an accuracy of 95%. But these are preliminary results and a full study with all results will be published later. We believe that the time is not far off when, instead of numerous examinations and visits to the laboratory, it will be enough to measure heart rate and respiration and, with the help of a device, obtain a very reliable answer as to whether a person is depressed or not. This would have a very significant impact both on the speed and reliability of diagnosis and on the appropriate method of treatment. All of this has an extremely significant impact on society and its development.

5 References

- [1] Abdallah, C. G., Sanacora, G., Duman, R. S., & Krystal, J. H. (August 2018). The neurobiology of depression, ketamine and rapid-acting antidepressants: Is it glutamate inhibition or activation? *Pharmacology & Therapeutics*, pp. 455-470. doi:10.1016/j.pharmthera.2018.05.010.

- [2] Ahn, J. W., Ku, Y., & Kim, H. C. (May 2019). A Novel Wearable EEG and ECG Recording System for Stress Assessment. *Sensors*, 9, pp. 1991. doi:10.3390/s19091991.
- [3] American Psychiatric Association. (2023). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5-TR)*. Retrieved September 2023 from psychiatry.org: <https://www.psychiatry.org/psychiatrists/practice/dsm>.
- [4] Blackwell, B. (2011). Volume 9: Update. In A. T. Ban, *An oral history of neuropsychopharmacology the first fifty years Peer interviews* (pp. x-xxiii). Brentwood, USA: ACNP.
- [5] Boucsein, W. (2012). *Electrodermal Activity*. New York, USA: Springer.
- [6] Brennan, D. (April 2022). *Depression in Children*. Retrieved September 2023 from Depression Guide: <https://www.webmd.com/depression/depression-children>.
- [7] Byun, S., Kim, A. Y., Jang, E. H., Kim, S., Choi, K. W., Yu, H. Y., & Jeon, H. J. (August 2019). Detection of major depressive disorder from linear and nonlinear heart rate variability features during mental task protocol. *Computers in Biology and Medicine*, p. 103381. doi:10.1016/j.combiomed.2019.103381.
- [8] Chen, J., Abbod, M., & Shieh, J.-S. (February 2021). Pain and Stress Detection Using Wearable Sensors and Devices-A Review. *Sensors (Basel)*, 4, p. 1030. doi:10.3390/s21041030.
- [9] Christie, M. J., Little, B. C., & Gordon, A. M. (1980). Peripheral indices of depressive states. In H. M. van Praag, M. H. Lader, O. J. Rafaelsen, & E. J. Sachar, *Handbook of Biological Psychiatry Part II: Brain Mechanisms and Abnormal Behavior Psychophysiology* (pp. 145-182). New York, USA: Marcel Dekker.
- [10] Clarke, E., & O'Malley, C. D. (1996). *The human brain and spinal cord: a historical study illustrated by writings from antiquity to the twentieth century*. San Francisco, USA: Norman Publishing.
- [11] Čukić, M., Savić, D., & Sidorova, J. (January 2023). When Heart Beats Differently in Depression: Review of Nonlinear Heart Rate Variability Measures. *JMIR Mental Health*, p. e40342. doi:10.2196/40342.
- [12] Fava, M. (April 2003). Diagnosis and definition of treatment-resistant depression. *Biological Psychiatry*, 8, pp. 649-659. doi:10.1016/s0006-3223(03)00231-2.
- [13] Ghaemi, S. N., & Mauer, S. (2020). Diagnosis, classification, and differential diagnosis of mood disorders. In J. Geddes, N. Andreasen, & G. Goodwin, *New Oxford Textbook of Psychiatry 3rd ed.* (pp. 681-690). Oxford, UK: Oxford University Press. Retrieved September 2023 from <https://doi.org/10.1093/med/9780198713005.003.0066>.
- [14] Goodwin, F. K., & Jamison, K. *Manic-Depressive Illness: Bipolar Disorders and Recurrent Depression*. New York, USA: Oxford University Press.
- [15] Gorman, J. M., & Justine, K. SSRIs and SNRIs: a broad spectrum of efficacy beyond major depression. *Journal of Clinical Psychiatry*, pp. 33-39.
- [16] Hickey, B. A., Chalmers, T., Newton, P., Lin, C.-T., Sibbritt, D., McLachlan, C. S., Lal, S. (May 2021). Smart Devices and Wearable Technologies to Detect and Monitor Mental Health Conditions and Stress: A Systematic Review. *Sensors*, 10, p. 3461. Retrieved from <https://doi.org/10.3390/s21103461>.
- [17] Hoppe, C. (January 2019). citing Hippocrates on depression in epilepsy. *Epilepsy & Behavior*, pp. 31-36. doi:10.1016/j.yebeh.2018.10.041.
- [18] Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., Research Domain Criteria (RDoC): Toward a New Classification Framework for Research on Mental Disorders. *The American Journal of Psychiatry*, pp. 748-751. doi:

- 10.1176/appi.ajp.2010.09091379.
- [19] Jožef Stefan Institute. (2016). ClowdFlows Data mining workflows on the cloud. Retrieved September 2023 from clowdflows.org: <http://clowdflows.org/>.
- [20] Kraleman, B., Frühwirth, M., Píkovský, A., Rosenblum, M., Kenner, T., Schaefer, J., & Moser, M. (September 2013). In vivo cardiac phase response curve elucidates human respiratory heart rate variability. *Nature Communications*, 1, p. 2418. doi:10.1038/ncomms3418.
- [21] Li, X., Hu, B., Shen, J., Xu, T., & Retcliffe, M. (December 2015). Mild Depression Detection of College Students: an EEG-Based Solution with Free Viewing Tasks. *Journal of Medical Systems*, 12, p. 187. doi:10.1007/s10916-015-0345-9.
- [22] Marneros, A. (January 2008). Psychiatry's 200th birthday. *The British Journal of Psychiatry*, 1, pp. 1-3. doi:10.1192/bjp.bp.108.051367.
- [23] Matic, Z., Kalauzi, A., Moser, M., Platiša, M. M., Lazarević, M., & Bojić, T. (December 2022). Pulse respiration quotient as a measure sensitive to changes in dynamic behaviour of cardiorespiratory coupling such as body posture and breathing regime. *Frontiers in Physiology*, pp. 1-15. doi:10.3389/fphys.2022.946613.
- [24] Mayo Clinic. (October 2022). Depression (major depressive disorder). Retrieved September 2023 from [Mayoclinic.org](https://www.mayoclinic.org/diseasesconditions/depression/diagnosis-treatment/drc-20356013): <https://www.mayoclinic.org/diseasesconditions/depression/diagnosis-treatment/drc-20356013>.
- [25] Murphy, S. E., Capitão, L. P., Giles, S. L., Cowen, P. J., Stringaris, A., & Harmer, C. J. (September 2021). The knowns and unknowns of SSRI treatment in young people with depression and anxiety: efficacy, predictors, and mechanisms of action. *The Lancet Psychiatry*, 9, pp. 824-835. doi:10.1016/S2215-0366(21)00154-1.
- [26] Narziev, N., Goh, H., Toshnazarov, K., Lee, S. A., Chung, K.-M., & Noh, Y. (March 2020).
- [27] STDD: Short-Term Depression Detection with Passive Sensing. *Sensors (Basel)*, 5, p. 1396. doi:10.3390/s20051396.
- [28] National institute of Mental Health. (April 2023). Depression. Retrieved from *Mental Health Information*: <https://www.nimh.nih.gov/health/topics/depression>.
- [29] Pies, R. W. (September 2020). How Avicenna Recognized Melancholia and Mixed States-1000 Years Before Modern Psychiatry. Retrieved September 2023 from *Psychiatric Times*: <https://www.psychiatristimes.com/>.
- [30] Porges, S. W. (2011). *The Polyvagal Theory: Neurophysiological Foundations of Emotions, Attachment, Communication, Self-Regulation*. 1st ed. New York, USA: W.W. Norton & Company.
- [31] Radden, J. (2002) *The Nature of Melancholy: From Aristotle to Kristeva*. Oxford, England: Oxford University Press.
- [32] Rottenberg, J. (December 2010). How Does Depression Start? Retrieved September 2023 from *Psychology Today*: <https://www.psychologytoday.com/us/blog/charting-the-depths/201012/how-does-depression-start>.
- [33] Rousseau, G. (April 2000). Depression's forgotten genealogy: notes towards a history of depression. *History of Psychiatry*, 41, pp. 71-106. doi:10.1177/0957154x0001104104.
- [34] Schimelpfening, N. (March 2023). The History of Depression. Retrieved September 2023 from *Verywell Mind*: <https://www.verywellmind.com/who-discovered-depression-1066770>.
- [35] Smith, L. T., Levita, L., Amico, F., Fagan, J., Yek, J. H., Brophy, J., Arvaneh, M.

(2020). Using Resting State Heart Rate Variability and Skin Conductance Response to Detect Depression in Adults. 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) (pp. 5004-5007). Montreal: QC, Canada. doi:10.1109/EMBC44109.2020.9176304.

[36] Stern, R. M., Ray, W. J., & Quigley, K. Psychophysiological Recording second edition. Oxford, England: Oxford University Press.

[37] Vivekanantham, S., Strawbridge, R., Rampuri, R., Raganathan, T., & Young, A. H. (September 2016). Parity of publication for psychiatry. *The British Journal of Psychiatry*, 3, pp. 257261. doi:10.1192/bjp.bp.115.165118.

The interconnection of Artificial Intelligence and Property Technology: literature review

Marijana Ribičić

DOBA Faculty

Innovation and Sustainable Business Management in Digital Society

Maribor, Slovenia

{marijana.ribicic}@net.doba.si

Abstract: *This paper presents a systematic literature review on the link between artificial intelligence (AI) and property technology (PropTech). PropTech means any technological solution in the real estate sector, be it 3D visualization, a platform to connect buyers and sellers of real estate, crowdfunding, facility management, property management, the sharing economy, smart cities, smart homes, AI (artificial intelligence) machine learning (ML), smart contracts or construction management BIM (building information modelling). This thesis begins with the overview of artificial intelligence, machine learning and PropTech referring to academic papers, textbooks and online journals.*

Key Words: *Property technology (PropTech), AI (artificial intelligence), ML (machine learning), VR (virtual reality), Property Management, Real Estate Agent Tools, Indoor Mapping, The IoT Home, literature review*

1 Introduction

PropTech is characterized by the massive implementation of emerging technology such as home matching tools, drones, virtual reality (VR), BIM, data analytics tools, AI, Internet of Things (IoT) and blockchain, smart contracts, crowdfunding in the real estate sector, FinTechs related to real estate, smart cities, regions, smart homes and shared economy. Categories of IT - platform PropTech, proposed by Oxford University, are divided into 12 types [1]:



Figure 1. Twelve Types of IT PropTech platforms

This article gives an overview of the impact of AI on each of the mentioned types of PropTech.

The term artificial intelligence was coined in 1956 and machine learning (ML) is a subset of AI, which mainly studies with the algorithms and computational statistics to perform high-level intellectual tasks without human intervention [2].

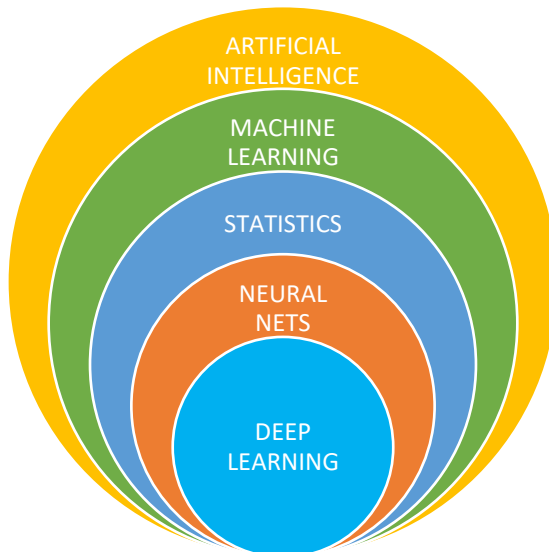


Figure 2: AI can be understood as an umbrella that consists of both ML and deep learning

2 Methodology

Using Google Scholar on 09/01/2023, I found 1070 results that referenced PropTech and artificial intelligence in the same article. I determined the most important papers and examined them in further detail based on the number of times a paper mentions AI and PropTech (or real estate). If AI was mentioned more than 10 times in connection with PropTech the paper was considered as highly relevant.

Majority of Google Scholar articles mentioned AI and PropTech together insignificantly, or quite by accident as I scrolled down the search results. Therefore, I analysed 18 scientific articles which closely referenced AI and PropTech (see attachment: Literature review).

3 Research

In the following sections I will show how AI relates to a specific type of PropTech.

3.1 The interconnection of Artificial Intelligence and Property Management

Modern Property Management Technology companies are using advanced technologies like AI and ML to overcome challenges in the property management sector. The section of the thesis [2] explores areas of property management, such as facility management and autonomous property systems, where robust technologies like ML and AI can make a difference. Finland have just started implementing AI and ML to fully automate aspects of property management. This section of research explores the potential application of AI with respect to Property Management System (PMS). The current PMS of Resimator [2], features a built-in service desk with a live chat client that supports video, audio and screen share options. Also, the live-messaging platform, connects easily with every web chat, app, social media messaging or voice interface. Property management companies are moving towards sensor deployment and automating work through the implementation of IoT devices, aiming to provide operational efficiency and improved energy management. AI can also undertake analyses and make decisions which were once routine undertakings carried out by people [3]. AI can change the idea of office work and the way property information is exchanged between stakeholders of commercial office space by replacing the humans with robots and ML tools to perform the routine tasks and reducing manning requirements in different property management functions.

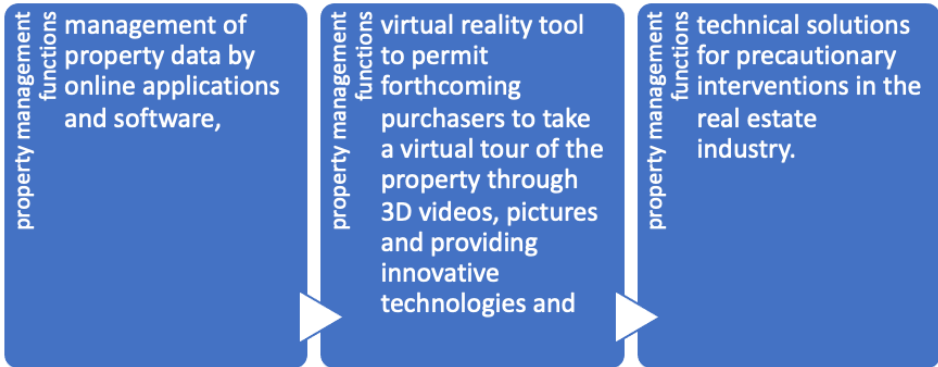


Figure 3. Property management functions

3.2 The interconnection of Artificial Intelligence and Construction Management

Wouter Truffino has been named in ‘the top 10 International PropTech Influencers 2017’, He started to build the first ecosystem in ConTech (Construction Technology) and PropTech (Property Technology) in the Netherlands in May 2015. After this Holland ConTech and PropTech were born.

The most important global development [4] is the speed of development in AI. ML and AI are likely to outpace the rate of improvement seen in transistor chips (following Moore’s Law), which has been the benchmark for technological progress for several decades. The spectrum of innovation in the Architecture-Engineering-Construction-Operation (AECO) industry is called PropTech [5]. In their research [5] authors wrote about *Netix Controls* which developed a centralized system based on ML, to collect and integrate data from different manufacturer chillers, fire alarms, and lighting systems, and applied it to two separate networks of 77 and 14 buildings, respectively, which correspond to the two case studies in this paper.

Wouter Truffino

- Truffino thinks AI is the major development that everybody needs to pay attention to.
- An example is the start-up *Octo*, which uses data analytics and ML to improve decision-making in asset management. For housing corporations, this would mean that short-term savings can be made on maintenance and inspection costs.

AECO industry

- *Netix Controls* has analysed and adopted in the international market (such as, USA, Finland and United Arab Emirate – UAE) the traditional Building Management System (BMS) for years to implement a new software, called *Konnect*, an AI powered platform that allows energy analytics, remote connectivity, and predictive maintenance.
- *Konnect* adds a building performance management system composed of AI and ML technologies, which receives data from the IoT sensors set in the buildings

Figure 4. PropTech Construction solution based on ML

3.3 The interconnection of Artificial Intelligence and Facility Management

Facility management has gained traction with the advancement in big data, IoT and AI with predictive repair and maintenance, and incident management being a focus area for many PropTech firms in recent [6]. New sensors lead to more stand-alone applications offered as Software as a service (*SaaS*). Examples are applications such as *Kapacity.io*, *BrainBox AI*, and *Facilio*. Those solutions use artificial intelligence and machine learning to optimize heating and cooling devices. They also need to be integrated into the existing Building Management System (BMS).

3.4 The interconnection of Artificial Intelligence and Portfolio Management

Portfolio management is a means of overseeing a group of investments that meet the long-term financial objectives and risk tolerance of a client, a company, or an institution (Wikipedia). ML and AI can and will be used to facilitate real estate investment in myriad ways, spanning all aspects of the real estate profession, from property management and investment decisions to development processes transforming real estate into a more efficient and data-driven industry [7]. At least 65% [7] of these companies are in early to advanced stages of machine learning analytics and are contending with the potential downfalls of machine learning, but finding new insights that customers need. The range of other machine learning and AI applications is evidence of great ingenuity on the part of entrepreneurs in finding ways to reshape the industry. 73% of companies use machine learning for data aggregation, cleaning, generation, and integration which implies significant improvements in future data analytics.

3.5 The interconnection of Artificial Intelligence and Long-Term Rentals / Sale Search

The aim of the paper [8] is to identify the attitudes and preferences of primary apartment buyers regarding the use of modern technology (PropTech) by developers in the local

residential property market. The paper [9] proposes a machine learning approach for solving the house price prediction problem. The paper [10] explores two key aspects: first, how platform entrepreneurs' views shape algorithms and adapt to venture capital funding, and second, how technology surveils landlords and tenants, creating inequalities through automated processes.

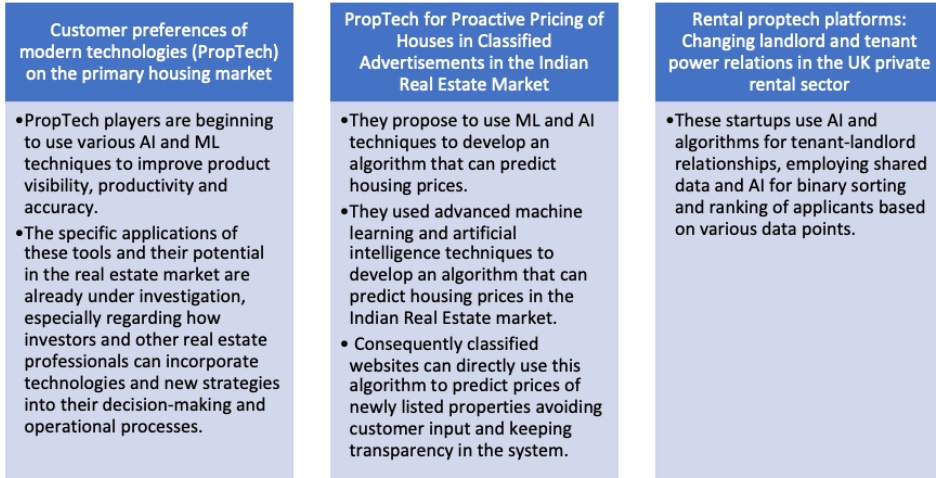


Figure 5. Different AI and ML solutions for Long-Term Rentals

3.6 The interconnection of Artificial Intelligence and Real Estate Agent Tools

The article [11] describes the development and importance of innovative companies offering solutions in the Polish real estate market. The *Skyconcept* survey confirms that Polish commercial real estate companies are interested in and willing to use more of Blockchain and AI technologies. Very good example of implementation of AI is the case study of *ImmoPresenter 3D* (IP3D)1 [12], Proptech startup company from Munich, Germany.

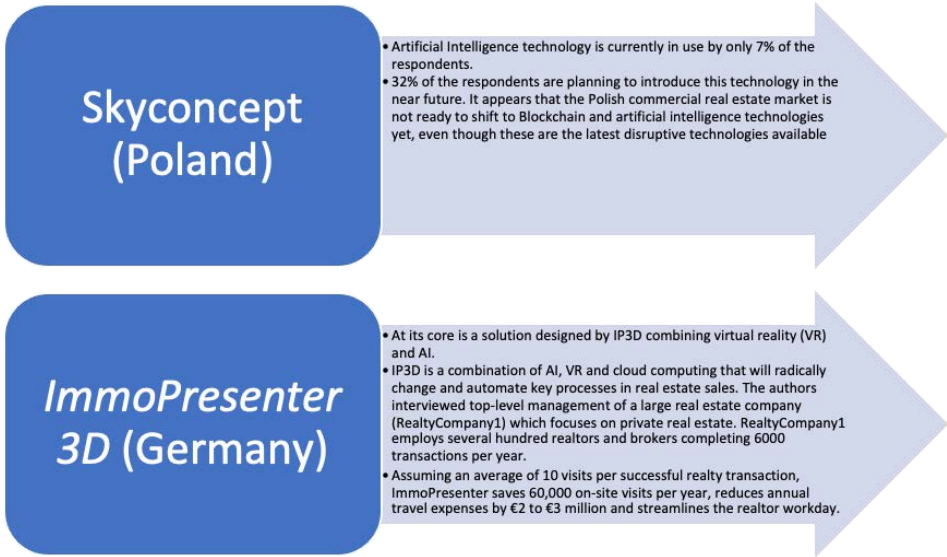


Figure 6. Real Estate PropTech companies using AI

3.7 The interconnection of Artificial Intelligence and Indoor Mapping

AI and ML technologies are enabling a wide range of new approaches to mapping, designing, and constructing the built world [7]. With machine learning technology 3D augmentation and space planning can be performed quickly and efficiently.

In order to analyze real estate within its physical context it is useful to employ geospatial analysis, the gathering, display, and manipulation of imagery, GPS, satellite photography and historical data, described explicitly in terms of geographic coordinates or implicitly, in terms of a street address, postal code, or forest stand identifier as they are applied to geographic models. The importance of location in real estate is not lost on the ML scientists who have developed many tools that integrate with geospatial information to help us better understand locational relationships and predict outcomes with this important layer of information in the mix [7].

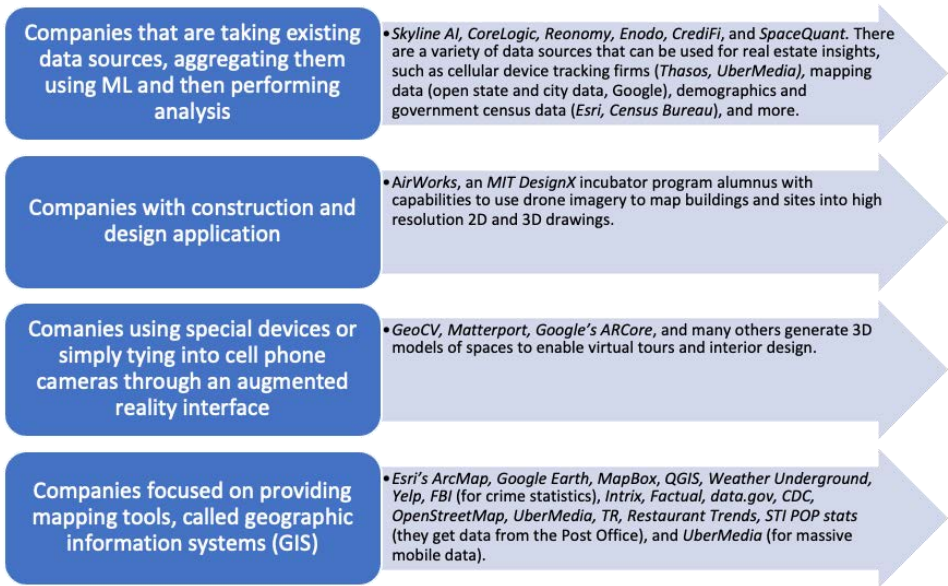


Figure 7. PropTech Indoor Mapping companies

3.8 The interconnection of Artificial Intelligence and the IoT Home

The extensive use of digital technologies in real estate, including the Internet of Things (IoT), cloud computing, decision automation, ML and AI is redefining the way people live, work and invest [8]. Property management companies are moving towards sensors deployment and automating work through the implementation of IoT devices, aiming to provide operational efficiency, improved energy management and cost-effective management [2].

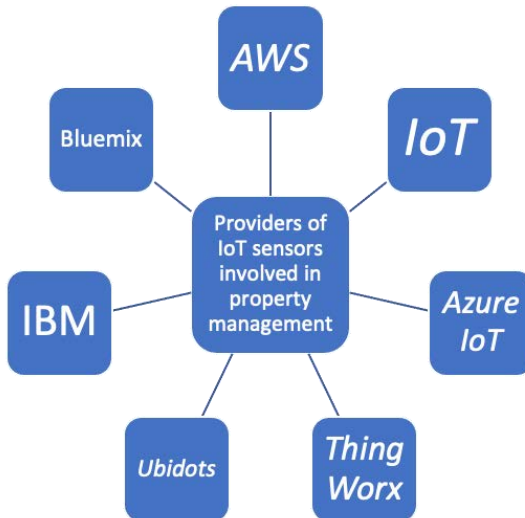


Figure 8. Providers of IoT sensors involved in property management

IoT data integrated monitoring platform will help IoT devices to collect property data and decide how data can be transferred and monitored in an integrated dashboard through mobile and desktop devices.

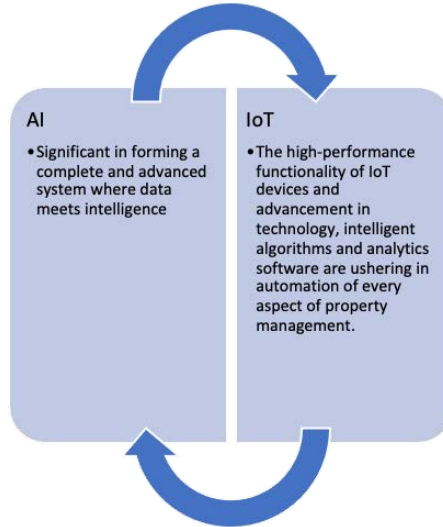


Figure 9. The relationship between IoT device data and AI

Netix Controls Company [5] commercializes a network of technologies based on IoT and AI by adopting a proprietary framework (called Niagara Framework) for facility and building management. *Netix Controls* is testing BMS and AI together since 2018 with the objective to overcome the traditional BMS with an “intelligent”- BMS.

4 Discussion

I didn't find scientific articles about connection between AI and:

- Home Services
- Short-Term Rental / Vacation Search
- Life, Home, Property & Casualty
- Commercial Real Estate Search

There is, maybe, no need for some PropTech types of to develop a special IT solution. The term “Home Services” refers to the upkeep, maintenance, repair, and improvement of residential properties. Repairs are done by an individual, but in the future we can fully expect to one day work side-by-side with some the type of robot. Keenly integrated with artificial intelligence and machine learning software, robotics is a rapidly-emerging field. When it comes to sectors such as property management, this will be revolutionary. But at the moment AI could only be applied maintenance call outs, or if the house is a "smart house", to automatically call an engineer when a fault is detected.

Long or short stay/search and commercial search both come under the IT PropTech umbrella of Real Estate Search. Digital nomads cannot be classified as either long or short stay visitors so for the purposes of AI and ML can be considered long stay search applications. There are a large number of agencies that deal with real estate sales as well as the short-term tourist rental of real estate.

IT PropTech solution for Life, Home, Property & Casualty helps cover privately owned items i.e. your home or your car. A casualty insurance policy includes liability coverage. Here we can use the same AI and ML solutions as in long term rentals. This could also range from data on hobbies, medical conditions, personalised data on smart watches, and wealth data such as income and assets. An interesting question here concerns the impact of AI and algorithms on personal data collection and analysis. Insurance companies would often note how an individual would have their data anonymised and could have it removed from the system under the General Data Protection Regulation (GDPR). However, AI and algorithms would continue to function and apply sorting criteria based on previous learning, which as highlighted earlier, may be based on incomplete data topologies.

From the research, we can conclude that AI and ML permeates almost every aspect of PropTech.

AI and ML are already present in most areas of life. AI and ML is on an upward trajectory. The technical solutions they offer facilitate business activity, reduce individual and business costs, give us more free time to enrich our lifestyles for example by exploring nature and being with family. The arrival of AI and ML makes our old way of life redundant. The future, especially with the advent of robotics (5th Industrial revolution), brings us a simpler life with less stress and more free time.

4. 1 ChatGPT as a part of PropTech in the future

ChatGPT has been investigated in the field of construction management, mapping, as artificial lawyer, designer and in the field of FinTech which can be also used in PropTech segment.

Generative Pre-Trained Transformer (GPT) language models such as ChatGPT have the potential to revolutionize the construction industry by automating repetitive and time-consuming tasks. The paper [13] presents a study in which ChatGPT was used to generate a construction schedule for a simple construction project. The output from ChatGPT was evaluated by a pool of participants that provided feedback regarding their overall interaction experience and the quality of the output. The results show that ChatGPT can generate a coherent schedule that follows a logical approach to fulfill the requirements of the scope indicated. The participants had an overall positive interaction experience and indicated the potential of such a tool in automating many preliminary and time-consuming tasks. However, the technology still has limitations, and further development is needed before it can be widely adopted in the industry. Overall, this study highlights the advantages of using large language models and Natural Language Processing (NLP) techniques in the construction industry and the need for further research.

Despite its exceptional ability to generate natural-sounding responses, I believe that ChatGPT doesn't possess the same level of understanding, empathy, and creativity as a human yet and cannot fully replace them in most situations. Future research needs to be conducted to further explore the applicability and capabilities of Natural Language Processing tools in the construction industry.

My conclusion is that we can involve ChatGPT to PropTech in the field of mapping. This despite is being at an early stage, ChatGPT has the potential to revolutionize the way they design and produce maps. However, it is crucial to acknowledge that mapping with ChatGPT is still an immature solution. While being a new alternative, it will not necessarily replace existing popular mapping solutions in the near future. Users should remain mindful while employing ChatGPT for mapping, particularly on the unvalidated

data sources suggested by ChatGPT, and the likely non-straightforward map improving process.

In terms of future works [14], one promising direction is to integrate ChatGPT with other AI tools with drawing capability, e.g., DALL·E 2, as an intelligent and autonomous mapping agent. The agent should include a decision-making core to overcome the current limitations, such as dependence on external conditions and unsatisfactory initial [18]. Another future direction is to comprehensively evaluate ChatGPT's spatial thinking capability and exploit its usefulness in a broad range of GIScience research and applications.

Given the different trade-offs between the approaches of JusticeBot and ChatGPT, an interesting approach could be combining the two. Tools such as the JusticeBot could be used to inject verified and accurate knowledge to ChatGPT. For example, ChatGPT could be used as the communications layer [15] that communicates with the user and makes the information accessible to them.

Development Bots (DevBots) [16] trained on large language models can help synergise architects' knowledge with AI decision support to enable rapid architecting in a human-bot collaborative Association Control Service Element (ACSE). An emerging solution to enable this collaboration is ChatGPT, a disruptive technology not primarily introduced for software engineering, but is capable of articulating and refining architectural artifacts based on natural language processing.

Research on AI and ChatGPT in the FinTech sector is lacking. ChatGPT have the potential to enhance fraud detection, customer loyalty, financial inclusion, data analysis, back-office process automation, and customer service. ChatGPT [17] can become a powerful tool for financial institutions and contribute to the growth of the financial sector.

6 References

- [1] N. Siniak, T. Kauko, S. Shavrov, and N. Marina, "The impact of proptech on real estate industry growth," *IOP Conference Series: Materials Science and Engineering*, vol. 869, no. 6, 2020. [Online]. Available: <https://doi.org/10.1088/1757-899X/869/6/062041>.
- [2] K. Sapkota, "ARTIFICIAL INTELLIGENCE IN PROPERTY MANAGEMENT AUTOMATION-Technologies, Current Applications and Challenges," n.d.
- [3] T. Kaur and P. Solomon, "A study on automated property management in commercial real estate: a case of India," *Property Management*, vol. 40, no. 2, pp. 247–264, 2022. [Online]. Available: <https://doi.org/10.1108/PM-05-2021-0031>.
- [4] W. Truffino, "Smart Tech developments changing the real estate industry," *Oligschlager. An Interview with Wouter Truffino of Holland ConTech & PropTech*, 2018.
- [5] A. P. Pomè, C. Tagliaro, A. Celani, and G. Ciaramella, "Is Digitalization Worth the Hassle? Two cases of Innovation Building Operation and Maintenance," *IOP Conference Series: Earth and Environmental Science*, vol. 1176, no. 1, 2023. [Online]. Available: <https://doi.org/10.1088/1755-1315/1176/1/012031>.
- [6] Z. Tan and N. G. Miller, "Connecting Digitalization and Sustainability: PropTech in the Real Estate Operations and Management," 2023. [Online]. Available: <https://doi.org/10.1080/19498276.2023.2203292>.
- [7] J. Conway and B. A. Architecture, "Artificial Intelligence and Machine

- Learning: Current Applications in Real Estate," 2018.
- [8] A. Górska, A. Mazurczak, and Ł. Strączkowski, "Customer preferences of modern technologies (PropTech) on the primary housing market," *Scientific Papers of Silesian University of Technology. Organization and Management Series*, 2022. [Online]. Available: <https://doi.org/10.29119/1641-3466.2022.162.12>.
- [9] S. Putatunda, "PropTech for Proactive Pricing of Houses in Classified Advertisements in the Indian Real Estate Market," n.d.
- [10] T. Wainwright, "Rental proptech platforms: Changing landlord and tenant power relations in the UK private rental sector?" *Environment and Planning A*, vol. 55, no. 2, pp. 339–358, 2023. [Online]. Available: <https://doi.org/10.1177/0308518X221126522>.
- [11] M. Jakub Merkel, "Key technologies and evolution of PropTech companies," in *Zarządzanie. Teoria i Praktyka*, vol. 1, 2023
- [12] R. Hagl and A. Duane, "Applying Business Model Innovation to Real Estate Distribution by Employing Virtual Reality and Artificial Intelligence A Report from Praxis Das 3D-Druck-Kompodium View project Impact of Augmented Reality and Virtual Reality Technologies on Business Model Innovation View project Applying Business Model Innovation to Real Estate Distribution by Employing Virtual Reality and Artificial Intelligence A Report from Praxis," 2020. [Online]. Available: <https://www.researchgate.net/publication/346037644>.
- [13] S. A. Prieto, E. T. Mengiste, and B. García de Soto, "Investigating the Use of ChatGPT for the Scheduling of Construction Projects," *Buildings*, vol. 13, no. 4, 2023, doi: 10.3390/buildings13040857.
- [14] R. Tao and J. Xu, "Mapping with ChatGPT," *ISPRS International Journal of Geo-Information*, vol. 12, no. 7, 2023, doi: 10.3390/ijgi12070284.
- [15] J. Tan, H. Westermann, and K. Benyekhlef, "ChatGPT as an Artificial Lawyer?" [Online]. Available: <https://chatpdf.com>.
- [16] A. Ahmad, M. Waseem, P. Liang, M. Fahmideh, M. S. Aktar, and T. Mikkonen, "Towards Human-Bot Collaborative Software Architecting with ChatGPT," in *ACM International Conference Proceeding Series*, 2023, pp. 279-285, doi: 10.1145/3593434.3593468.
- [17] H. Ali, A. F. Aysan, and R. Associate, "What will ChatGPT Revolutionize in Financial Industry?" *Foundation Non-Resident Fellow Middle East Council for Global Affairs (ME Council) What will ChatGPT Revolutionize in Financial Industry*, n.d.
- [18] Z Li, H. Ning, *Autonomous GIS: The next-generation AI-powered GIS*. arXiv 2023, arXiv:2305.06453.

The Role of Artificial Intelligence in Circular Economy

Milica Stanković, Gordana Mrdak, Tiana Anđelković
Academy of Applied Technical and Preschool Studies
Department Vranje, Filipa Filipovića 20, Serbia
{milica.stankovic, gordana.mrdak, tiana.andjelkovic-
vr}@akademijanis.edu.rs

Abstract: *In order to save the „Earth capital“, we must introduce a model of circular economy in which each of us have responsibility for sustainable use of resources and elimination of waste. The circular economy changes business models, habits and way of thinking, both of producers and consumers, because the new eco-design of the product extends its life through repair, modification and recycling. Artificial Intelligence facilitates a circular economy and supports the design, development and maintenance of circular products. AI can assist product designers in creating environmental friendly products and monitoring products in order to generate possible product improvements. In general, artificial intelligence can be used to analyze data collected during the product life cycle to improve efficiency in real time. The aim of the paper is to emphasize the role of artificial intelligence in circular economy.*

Key Words: *circular economy, artificial intelligence, linear economy*

1 Introduction

The circular economy changes business models, habits and way of thinking, both of producers and consumers, because the new eco-design of the product extends its life through repair, modification and recycling. All processes take place with the use of renewable energy sources. In short, the circular economy implies the circulation of materials and their reuse, which simultaneously uses drastically less energy and water.

The aim of the paper is to emphasize the role of artificial intelligence in circular economy. In the first part of the paper, we will point out the necessity of moving from a linear to a circular economy. The second part of the paper is dedicated to Europe's transition to Circular Economy that will contribute to a number of advantages for people and society as a whole. The third part of the paper focuses on the role of artificial intelligence in the circular economy. After a comprehensive analysis, we provide relevant conclusions.

2 From linear to circular economy

The linear model of production is based on the transformation of resources into products and their conversion into waste after use. The main paradigm in "linear economy" is: take - make/use – dispose. This means that we don't only consume natural resources uncontrollably, but also we produce huge amounts of hazardous waste, which nature

cannot decompose and absorb. [1] The consequence of such action is the reduction of natural resources, large amounts of waste and environmental pollution. The situation is getting exponentially worse. This is why the linear economy has to change. It's very important to think about what kind of products we buy, how that product affects us and the environment, and whether we really need it. [2]

According to estimates of the Global Footprint Network, current model of economic growth, based on the use of natural resources, has put humanity in a position to use as many resources in just seven months as all ecological systems on the planet can renew in a year. In other words, our generation uses the "Earth capital" of future generations. Linearity implies that by 2060 we will need at least two planets to meet the demand for materials. [3] [4]

The question is what we can do and what we must do to save the „Earth capital“? We must introduce a new model of economy in which each of us have responsibility for our relationship with natural resources. The main goal of the circular economy (CE) is sustainable use of resources and elimination of waste [3]. The main source of economic growth is the greatest possible reuse of materials from products that have completed their "life cycle" and the least possible use of new resources. Products are designed so that they can be easily reused, disassembled, repaired or recycled. In the concept of circular economy, waste does not exist, but only raw material that can be reused for the same or other production processes. In the circular economy, the waste of one industry is the raw material of another. According to the Circularity Gap Report, the world is only 9% circular and that trend is negative. In the EU, this share is slightly higher, around 12%. [1]

Table 1: Comparison between the linear economy and circular economy

Linear economy	Circular economy
Linear supply chains follow a Take-Make-Waste model	Circular supply chain follows a Reduce-Reuse-Recycle model
Focus is on producing as much as possible, quickly and at a low cost	Focus is on reducing waste and maximizing the value of resources
Suppliers are chosen based on the lowest cost and quickest lead time	Suppliers are chosen based on sustainability criteria, including use of recycled materials, low waste generation, and reduced carbon emissions
Products are designed for single-use	Products are designed to be used for longer periods increasing its durability, reparability and recyclability
Waste is generated at each stage of the supply chain	Waste is minimized at every stage of the supply chain There is a focus on recycling and repurposing materials

3 Europe's circular economy transition

In December 2015, European Commission adopted Circular Economy Action Plan to accelerate the transition from a linear to a circular economy. Key elements of the revised framework for waste management from 2018 are the goals that the European Union should reach by 2035. Among others, these goals include: [5] [6]

- Recycling 65% of municipal waste by 2035;
- Recycling of 70% of packaging waste by 2030, with specific goals for the recycling of individual packaging materials. Paper and cardboard - 85%; aluminum - 60%; glass - 75%; plastic - 55%; wood - 30%;
- Disposal of no more than 10% of produced waste in landfills by 2035;
- Extending the obligation of separate waste collection to hazardous waste from households until the end of 2022; bio-degradable waste by the end of 2023; and textiles by the end of 2025;
- Strengthening prevention and taking special measures to minimize food waste, as a contribution to achieving the EU's obligations towards the goals of sustainable development.

In 2019, European Commission presented the "European Green Deal" as the most ambitious package of measures to make Europe the first climate-neutral continent in the world and a world leader in circular economy and clean technologies by 2050. With the Green Agreement, the EU committed itself to meet the goals of the 2030 Agenda and the Paris Agreement from 2015. For Europe by 2030, the potential economic gain from the transition to a circular economy is estimated at 1.8 billion euros. [5] [6]

4 Artificial Intelligence within circular economy activities

Artificial intelligence (AI) is a term used to describe intelligent systems that attempt to acquire human-like cognitive skills, such as the ability to reason, understand meaning, generalize, and understand from previous experiences. [7] [8] Digital innovations facilitate a circular economy that ensures the maximum use of limited resources through the use of digital platforms, smart devices and artificial intelligence. [9]

AI is one of the paramount technologies which can provide numerous advantages with the assistance of various algorithms for a smooth transition towards CE. For instance, real-time data analysis for supply chain management, cost reduction and carbon footprint reduction for sustainable development, automate the processes for reverse logistics, assessing the impact of waste generation for waste management, sorting different materials for recycling purposes etc. [7]

AI can support the design, development and maintenance of circular products. Products should be designed to ensure a long product life and improve recycling potential. AI can assist product designers by suggesting initial designs and prototypes for environmentally friendly products. [10] Also, AI can help design new materials to replace unsustainable resources, such as harmful chemicals. This design could improve product durability and facilitate end-of-life recycling. AI could be used to monitor products and study performance over time to generate possible product improvements. [11]

AI could facilitate circular businesses by supporting the recycling infrastructure needed for a functioning circular economy. Effective sorting is required, because CE involves reusing, repairing, and recycling products. AI-powered image recognition can identify and differentiate waste, minimising resource loss. For instance, Unilever and Alibaba recently partnered to trial an AI-enabled sorting machine that distinguishes between different types of plastic. Similarly, Apple's Daisy robot can "take apart up to 200 iPhone devices per hour, removing and sorting components to recover materials that traditional recyclers can't—and at a higher quality". This facilitates higher value recovery of materials, creating secondary product markets. [11]

8 Conclusion

In order to save the „Earth capital“, we must introduce a model of circular economy in which each of us have responsibility for sustainable use of resources and elimination of waste. In the concept of circular economy, the waste of one industry is the raw material of another and raw material can be reused for the same or other production processes. Artificial Intelligence facilitates a circular economy and supports the design, development and maintenance of circular products. AI can assist product designers in creating environmental friendly products and monitoring products in order to generate possible product improvements. In general, artificial intelligence can be used to analyze data collected during the product life cycle to improve efficiency in real time.

9 References

- [1] Kowszyk, Y., Maher, R. Case studies on Circular Economy models and integration of Sustainable Development Goals in business strategies in the EU and LAC, EU-LAC foundation, Hamburg, 2018.
- [2] GIZ. Osnove cirkularne ekonomije, Deutsche Gesellschaft für Internationale Zusammenarbeit GIZ GmbH, Beograd, 2016.
- [3] The World Bank. Squaring the circle: Policies from Europe's Circular Economy Transition, International Bank for Reconstruction and Development/The World Bank, Washington, 2022.
- [4] Ministarstvo zaštite životne sredine Republike Srbije. Mapa puta za cirkularnu ekonomiju u Srbiji, Beograd, 2020.
- [5] Janković, E., Vukomanović, Ž., Stojinović, G., Cvejanov, K., Jezdimirović, I. Polazne osnove za tranziciju ka cirkularnoj ekonomiji, Inženjeri zaštite životne sredine, Novi Sad, 2022.
- [6] Gluščević, M., Kaluđerović, Lj. Analiza kapaciteta jedinica lokalne samouprave u pogledu stvaranja uslova za prelazak na cirkularnu ekonomiju, GIZ, Nemačka, 2019.
- [7] All Noman, A., Habiba Akter, U., Pranto, T.H., Bahalul Haque, A, Machine Learning and Artificial Intelligence in Circular Economy: A Bibliometric Analysis and Systematic Literature Review, Annals of Emerging Technologies in Computing (AETiC), 6(2):13-40, 2022.
- [8] Pathan, M.S.; Richardson, E.; Galvan, E.; Mooney, P. The Role of Artificial Intelligence within Circular Economy Activities—A View from Ireland. Sustainability, 2023.
- [9] Cioffi, R., Travaglioni, M., Piscitelli, G., Petrillo A., Parmentola, A. Smart Manufacturing Systems and Applied Industrial Technologies for a Sustainable Industry: A Systematic Literature Review, Applied Sciences, 10(8), 2020.

[10] Acerbi, F., Forterre, D.A., Taisch, M. Role of artificial intelligence in circular manufacturing: a systematic literature review. 54(1):367–372, 2022.

[11] Roberts, H., Zhang, J., Bariach, B., Cowls, J., Gilbert, B., Juneja, P., Tsamados, A., Ziosi, M., Taddeo, M., Floridi, L., Artificial intelligence in support of the circular economy: ethical considerations and a path forward, AI & SOCIETY, 2022.

Test Environment for the Implementation of the Circular Economy

Erika Džajić Uršič

Rudolfovo – Science and Technology Centre Novo mesto
Podbreznik 15, 8000 Novo mesto, Slovenia
{erika.ursic}@rudolfovo.eu
Faculty of Information Studies
Ljubljanska cesta 31A, 8000 Novo mesto, Slovenia
{erika.ursic}@fis.unm.si
School of Advanced Social Studies in Nova Gorica
Gregorčičeva 15, 5000 Nova Gorica
{erika.ursic}@fuds.si

Urška Fric

Faculty of Information Studies
Ljubljanska cesta 31A, 8000 Novo mesto, Slovenia
{urska.fric}@fis.unm.si
Rudolfovo – Science and Technology Centre Novo mesto
Podbreznik 15, 8000 Novo mesto, Slovenia
{urska.fric}@rudolfovo.eu

Alenka Pandiloska Jurak

Rudolfovo – Science and Technology Centre Novo mesto
Podbreznik 15, 8000 Novo mesto, Slovenia
{alenka.pandiloska}@rudolfovo.eu
Faculty of Information Studies
Ljubljanska cesta 31A, 8000 Novo mesto, Slovenia
{alenka.pandiloska}@fis.unm.si

Abstract: *The transition from linear to Circular Economy (CE) is one of the priorities in the economy and one of the global trends – in Slovenia, in the European Union (EU), and beyond. Considering the EU is too slow in the transition in line with the post-2015 roadmap, the work will need to be accelerated. At the same time, the EU is moving into the digital decade, as the digital transition is one of the critical elements of the EU's development and strategic autonomy. Digitalization has the potential to provide solutions to a number of challenges, including the transition from linear to CE. We must act digitally if we want to catch up in the transition. The paper presents a pioneering research project aimed at establishing a dedicated test environment for the potential implementation of CE. It enables an impact assessment of the key challenges (properties) that most often influence the implementation of CE. The test development should provide answers to critical questions on the impact of actor relations in CE implementation and its long-term implications for contributing to the Sustainable Development Goals.*

Key Words: *Sustainability, Circular Economy, Transition, Test Environment, EU*

1 Introduction

The concept of CE has gained substantial traction in recent years as a strategy to enhance companies' competitiveness while concurrently mitigating societal environmental impact [1]. CE advocates for a regenerative production system wherein resource cycles are closed, and the intrinsic value is maximized by prolonging usage through closed-loop value chains, circular business models, and designs emphasizing durability [2]. Within the literature, CE serves as an overarching concept [3] characterized by an ongoing absence of consensus regarding its precise definition [4] or the terminology associated with its various operationalization strategies [5].

In scientific literature, there has been a notable emphasis on developing methodologies and tools to facilitate the transition of companies towards CE [6]–[8]. Some scholars have concentrated on gaining a deeper understanding of the static perspective of CE business models, constructing representation frameworks that describe or model the discrete components of the CE business model [8], which means creating tools to figure out how CE business models and their parts should change over time to match CE principles [2], [7], [9]–[11].

Considering that the EU is too slow in the transition from linear to CE in line with the post-2015 roadmap, the work on the transition will need to be accelerated. A possible way toward effectuating this holistic transition could be the test environment for implementing CE, discussed in the subsequent section. The test environment will assume a transformative role by channeling research-driven insights into actionable frameworks. A test environment typically refers to a controlled setting or platform where researchers, businesses, or policymakers can experiment with and evaluate various strategies, technologies, and processes related to CE. This environment facilitates testing, analysis, and simulation of CE practices without directly impacting real-world operations [11]. Our test environment is expected to offer significant advantages, including cost savings, broad accessibility, adaptability to various actor needs, and the potential to provide customizable control services for regions and local communities.

To address the issues in the transition from linear to CE, the following research question has been defined to guide research: “To what extent does dedicated test software/environment facilitate optimization of CE processes in Slovenia?”. By understanding the profound and lasting effects of embracing CE practices we aim to empower actors with the insights needed to make informed decisions, foster sustainable collaborations, and contribute to the broader goals of environmental and economic resilience.

2 Methods and Methodology

2.1 Test environment

The proposed methodology involves developing and using test software. The envisioned software is a versatile platform designed to test and simulate specific attributes closely tied to the physical, environmental, technological, and social aspects of exchanges within

the CE framework, including appropriate properties gained from the empirical research analyses of case studies.

This study includes three key steps: conducting a literature review of scientific bases (WoS, Scopus ...) and related secondary research to identify CE-related properties, using case studies from Slovenia to illustrate these properties, and analysing the outcomes to assess the effectiveness of CE indicators in evaluating environmental impacts within various CE strategies.

The development of the test environment will later be used to develop a methodological outline – model with CE approaches. These approaches could be based on physical, organizational, and social conditions, such as accessibility and availability of primary raw materials, and proximity to business actors, learning organizations, established networks, etc. Building a new model would be possible through meta-analysis to “guess” the cause-and-effect relationships between business actors within the CE process. Emphasis could also be placed on social interactions between business actors, as social interactions lay the foundations for creating synergies between entities. After the model is established, the simulation in the test environment will follow.

The dedicated physical space for the test environment must house the necessary hardware and software infrastructure, along with personnel equipped with the required technical capabilities. The solution of the test environment encompasses:

- A data repository (architecture and implementation) serves as a collection of data concerning existing flows (material, service, financial, knowledge, technology, and experience flows) among actors in the CE. This repository is intended to support decision-making processes.
- An ontology for the CE process (architecture and implementation) that will facilitate integration of actors into complex networks and assess their potential within economic, technological, and environmental contexts.
- A module for simulating the characteristics of CE.
- Additional services, such as setup of hosting on hosted servers.

An intuitive content input platform..

2.2 Steps to perform the test process

A structured testing process is crucial for ensuring reliability and functionality of software or systems. It involves several systematic steps. Initially, the test team acquaints themselves with requirements and constraints. Next, a comprehensive test plan is developed, encompassing unit tests for individual components and functionalities with results documented in collaboration with project manager. Detailed test specifications are later prepared for system and integration tests, including various scenarios and performance evaluations. When a test environment is tailored to the product's architecture, it is set up followed by unit testing to confirm each module's correctness. Test cases are defined, and the software is deployed on the test environment. The test team rigorously executes test cases, recording results and addressing any identified errors. Error analysis and reporting are conducted, followed by preparation of the final report assessing product suitability and compliance with user or client requirements. If excessive defects are found, testing cycles are repeated until defect levels meet predetermined criteria [12], [13].

3 Conclusion

CE has emerged as a crucial strategy to enhance company competitiveness while mitigating environmental impacts. The research addresses the challenges Slovenia and the EU face in managing its waste and resources effectively, especially in light of climate change. It acknowledges that waste management is strictly regulated and requires companies to adhere to guidelines for waste handling. However, transitioning to CE represents almost a voluntary decision by companies, distinguishing it from traditional waste management.

To address this challenge, this research proposes the development of a test environment to explore the dynamics of CE implementation in Slovenia's ecosystem. The test environment combines data storage, ontology, and simulation modules, offering a comprehensive platform to investigate CE practices' impact on actor relationships and facilitate decision-making processes.

By promoting collaboration among actors throughout Slovenia, this initiative aims to generate invaluable insights into the effects of CE introduction on relationships within the ecosystem. The crucial research question focuses on understanding the long-term consequences of CE adoption, highlighting its potential to foster economic, technological, and environmental sustainability.

4 References

- [1] M. Geissdoerfer, P. Savaget, N. M. P. Bocken, and E. J. Hultink, The Circular Economy – A New Sustainability Paradigm?, *Journal of Cleaner Production*, vol. 143, pp. 757–768, Feb. 2017, doi: 10.1016/j.jclepro.2016.12.048.
- [2] P. Ghisellini, C. Cialani, and S. Ulgiati, A Review on Circular Economy: The Expected Transition to a Balanced Interplay of Environmental and Economic Systems, Feb. 2016, doi: 10.1016/j.jclepro.2015.09.007.
- [3] F. Blomsma and G. Brennan, The Emergence of Circular Economy: A New Framing Around Prolonging Resource Productivity, *Journal of Industrial Ecology*, vol. 21, no. 3, pp. 603–614, 2017, doi: 10.1111/jiec.12603.
- [4] J. Kirchherr, D. Reike, and M. Hekkert, Conceptualizing the Circular Economy: An Analysis of 114 Definitions, *Resources, Conservation and Recycling*, vol. 127, pp. 221–232, Dec. 2017, doi: 10.1016/j.resconrec.2017.09.005.
- [5] D. Reike, W. J. V. Vermeulen, and S. Witjes, The Circular Economy: New or Refurbished as CE 3.0? – Exploring Controversies in the Conceptualization of the Circular Economy through a Focus on History and Resource Value Retention Options, *Resources, Conservation and Recycling*, vol. 135, pp. 246–264, Aug. 2018, doi: 10.1016/j.resconrec.2017.08.027.
- [6] N. M. P. Bocken and S. W. Short, Towards a Sufficiency-driven Business Model: Experiences and Opportunities, *Environmental Innovation and Societal Transitions*, vol. 18, pp. 41–61, Mar. 2016, doi: 10.1016/j.eist.2015.07.010.
- [7] M. P. P. Pieroni, T. C. McAloone, and D. C. A. Pigosso, From Theory to Practice: Systematising and Testing Business Model Archetypes for Circular Economy, *Resources, Conservation and Recycling*, vol. 162, p. 105029, Nov. 2020, doi: 10.1016/j.resconrec.2020.105029.
- [8] P. Rosa, C. Sassanelli, and S. Terzi, Towards Circular Business Models: A

- Systematic Literature Review on Classification Frameworks and Archetypes, *Journal of Cleaner Production*, vol. 236, p. 117696, Nov. 2019, doi: 10.1016/j.jclepro.2019.117696.
- [9] F. J. Diaz Lopez, T. Bastein, and A. Tukker, Business Model Innovation for Resource-efficiency, Circularity and Cleaner Production: What 143 Cases Tell Us, *Ecological Economics*, vol. 155, pp. 20–35, Jan. 2019, doi: 10.1016/j.ecolecon.2018.03.009.
- [10] European Commission, “A European Green Deal,” European Commission – European Commission. 2022, https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal_en, downloaded: Mar. 10, 2022.
- [11] C. D. Tupper, “10 – Models and Model Repositories,” in *Data Architecture*, C. D. Tupper, Ed., Boston: Morgan Kaufmann, 2011, pp. 191–205. doi: 10.1016/B978-0-12-385126-0.00010-3.
- [12] W. L. Oberkampf, T. G. Trucano, and C. Hirsch, Verification, Validation, and Predictive Capability in Computational Engineering and Physics. *Applied Mechanics Reviews*, vol. 57, no. 5, pp. 345–384, Dec. 2004, doi: 10.1115/1.1767847.
- [13] S. G. Charlton and T. G. O’Brien, *Handbook of Human Factors Testing and Evaluation*. CRC Press, 2019.

Breathing Easy in North Macedonia: The Effect of Green Infrastructure and Movement Restrictions on The Air Quality

Mare Srbinovska, Vesna Andova, Aleksandra
Krkoleva Mateska, Maja Celeska Krsteska
Faculty of Electrical Engineering and Information Technologies,
Ss Cyril and Methodius University in Skopje
Ruger Boskovik 18, 1000 Skopje, Macedonia
{mares, vesnaa, krkoleva, celeska} @feit.ukim.edu.mk

Abstract: *Air quality is a significant issue in urban areas, marked by poor air quality characterized by elevated levels of particulate matter (PM). This particulate matter includes black carbon, volatile organic compounds, and assorted pollutants that pose threats to both human health and the overall environment on a global scale. Consequently, there exists an urgent need to reduce air pollution by adopting various strategies. In this paper, we analyze the effect of green areas and movement restrictions on the concentrations of PM. The tests are based on the data collected during the period 2018-2022 at the technical campus of Ss Cyril and Methodius University.*

Key Words: *air quality monitoring system; green area; covid, sensor network; particulate matter.*

1 Introduction

Air pollution presents an ever-expanding challenge with far-reaching consequences for the well-being of all life forms inhabiting our planet. This issue has garnered increasing attention in recent years, primarily due to its detrimental effects on both human health and the environment. According to estimates by the World Health Organization (WHO) in 1999 [1], air pollution contributes to approximately 7 million premature deaths annually, accounting for over one in every ten recorded deaths.

Extensive research, including studies conducted by Dockery et al. in 1992 [2], as well as Dockery and Schwartz [3] in the same year, has unveiled a clear correlation between elevated concentrations of air pollutants like Ozone (O₃), Particulate Matter (PM), and Sulfur Dioxide (SO₂), and an elevated mortality rate.

Furthermore, a 2019 report by the European Environment Agency (EEA) [4] underscores the gravity of the situation, revealing that air pollution stands as one of the primary causes of premature death in numerous European nations, responsible for over 400,000 such fatalities.

Beyond its direct impact on human health, these pollutants pose a significant threat to cultural heritage, particularly monuments and artworks located in urban centers, as noted by Samet et al. [5]. In urban environments, major air pollutants include PM, SO₂, Carbon Monoxide (CO), O₃, Nitrogen Oxides (NO_x), and Volatile Organic Compounds (VOCs), as highlighted by Mayer [6].

The initial objective of this paper is to examine the impact of air quality improvement measures, such as installations of green infrastructures, and reduction of traffic and

activities. This is done by evaluating the effects of existing green infrastructure and the movement restrictions introduced to fight the Covid-19 pandemic.

The data is collected by three sensor nodes positioned in the technical campus of Ss Cyril and Methodius University in Skopje. Although the series of measurements started in May 2018, due to technical challenges and data insufficiency in 2019, the measurement results for that year are not included in this paper.

This paper is organized as follows. Section two introduces the measurement sensors description and methods used to collect the data. The third section presents the results obtained with statistical analysis of the data. The last section concludes the paper.

2 Method

2.1 Measurement system description

The measurement monitoring system was originally designed employing wireless sensor network technology. It comprises of multiple sensor nodes, with each node incorporating four sensors and a Wi-Fi module integrated onto a single board controller. These sensors are responsible for measuring various parameters, including PM_{2.5}, PM₁₀, CO, and NO₂. The controller's primary responsibility involves processing data before transmitting it across the network. The sensor nodes, equipped with Wi-Fi modules, communicate data to the nearest routers situated within the Faculty building. The collected data from these routers is uploaded onto an accessible platform. Users can monitor the data online or opt to download it for in-depth analysis. More information regarding the hardware specifications and key characteristics of this configuration can be found in [7] and [8].

2.2 Data Acquisition

The data utilized in this study was sourced from two primary origins: sensor data and publicly available weather information. Throughout the study duration, the data was obtained in CSV format and then structured and readied for in-depth analysis. Originally, sensor data was gathered at 30-second intervals. To enhance the data's manageability and significance, frequent readings were transformed into hourly averages, calculated by determining the mean value of readings within each hourly period.

To ensure uniformity and comparability across the data, sensor readings were converted from ohms (Ω) to micrograms per cubic meter ($\mu\text{g}/\text{m}^3$).

3 Results

In the observed period, the setup provides continuous measurements for all four parameters but, in this article, the concentration of PM_{2.5} and PM₁₀ are analyzed. The aim is to evaluate the effect of different environmental and societal conditions on the concentrations of PM in the air. The next subsections analyze each of these questions separately.

In the studied period, results show that PM concentration is much higher during the winter months. Similarly, for a 24 hour period, it is reported the highest concentration of PM is recorded during nighttime, i.e., around 2-3 a.m. The graphs in Figure 1 show the repetitive occurrence of the PM_{2.5} and PM₁₀ peaks and their duration. Furthermore, the correlation coefficient between measured PM_{2.5} and PM₁₀ concentrations shows that there is a

strong positive correlation (>0.9) between the concentration of PM_{2.5} and PM₁₀ obtained from each node.

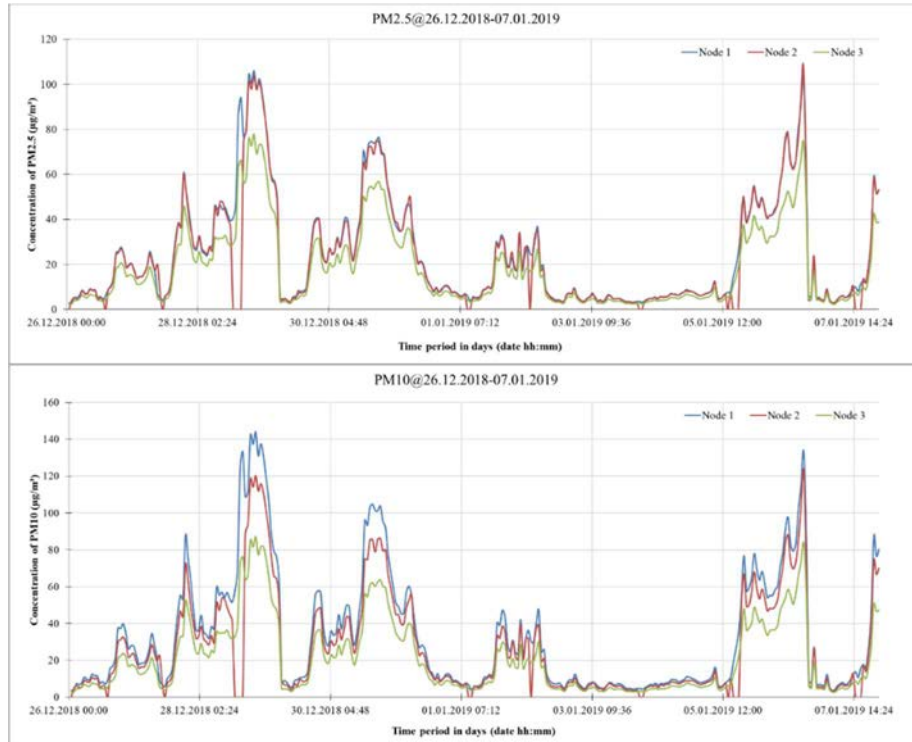


Figure 1: Concentration of PM_{2.5} (top) and PM₁₀ (bottom) for 13 days, average hourly data.

3.1 The effect of green infrastructure on PM

During our study's initial phase, the influence of vegetation on PM concentration levels is analyzed. This investigation covered the period from May 2018 to April 2019. The results reveal that the sensor located near the green area consistently registers lower PM concentration values compared to the other sensors. The performed statistical test confirms that the reported difference in the concentration of PM at different locations is statistically significant (p -value=0.000). Combined with corresponding post hoc test, the data confirms that variations in measured data are correlated with the placement of the sensor nodes, with the node situated farthest from the green area exhibiting the highest recorded PM concentrations. Further data analysis shows that the difference in measurements between sensor located near the green area and the other two sensors is apparent in the periods of extreme PM concentrations (higher than 50 mg/m^3 for PM_{2.5} and higher than 100 mg/m^3 for PM₁₀). The results presented in [9] show that the impact of the nearby green area results in a nearly 25% reduction in PM_{2.5} when compared with sensors surrounded solely by grassy patches. Throughout the entire analyzed timeframe, the reduction in PM₁₀ pollution at location near the green area amounts to a substantial 37%. Specifically, during periods of normal concentration (below 50 mg/m^3), the PM₁₀

concentration at the same location is 35% lower, and this difference increases to an even more significant 43.5% during periods of extreme concentration.

3.2 The effect of the Covid pandemic

For estimating the effect of the Covid pandemic, i.e., the imposed restrictions introduced during spring 2020 we analyzed the data collected in May 2018, 2020, 2021, and 2022 [10]. In May 2020 several precaution measures were introduced: kindergartens, schools and universities were closed, there were lockdowns and curfews, and some factories, stores, and construction fields were closed. The meteorological records reveal that May 2020 and May 2021 experienced cooler temperatures compared to May 2018 and May 2022. Additionally, it is noteworthy to observe that while the temperatures were similar between May 2020 and May 2021, May 2020 exhibited substantially higher levels of humidity, cloud coverage, and precipitation.

Figure 2 shows the boxplot of the data over the observed period. They suggest the data is not normally distributed, which is also confirmed by normality test. They also indicate better air quality in 2020 and 2021 compared to 2018 and 2022. Furthermore, it is evident that the data exhibits significantly greater variability in 2018 and 2022 when compared to the years 2020 and 2021. Notably, the disparity in recorded data between 2020 and 2021 is quite minimal for PM_{2.5}, whereas a more pronounced difference is observed for PM₁₀. In 2018, in contrast to the other three observed periods, numerous outliers are documented, and the maximum concentration significantly exceeds that recorded in the remaining three analyzed time frames.

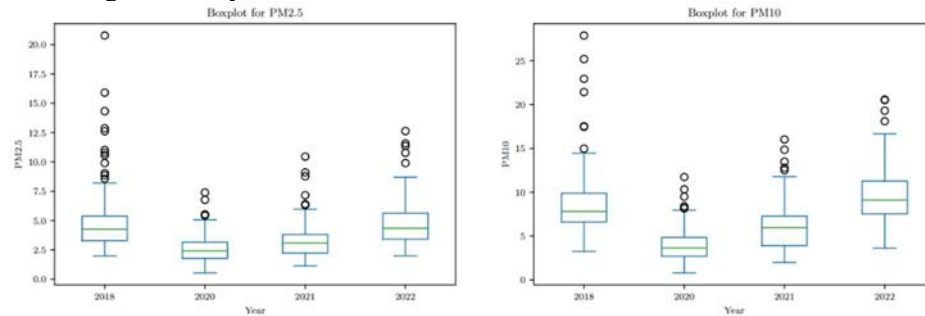


Figure 2: Boxplot for PM_{2.5} (left), and PM₁₀ (right) for the data in the observed period.

With an appropriate hypothesis test, we analyzed if there was a difference in the concentration of PM_{2.5} (resp. PM₁₀) in May 2018, 2020, 2021, and 2022, i.e., pre-Covid, during the pandemic, and post-Covid. The test rejected the Null hypothesis with a p-value 0.000, so there is a significant difference in the PM_{2.5} (resp. PM₁₀) concentration through the years in the observed period. To locate the difference a post hoc test was performed. It showed a significant reduction in the concentrations of both PM_{2.5} and PM₁₀ in 2020 when compared to the years 2018 and 2022. While there is a concentration variation in favor of 2022, the results suggest it is not significantly different from the concentration observed in 2018. Similarly, the concentration of PM in 2021 is considerably lower than that in 2018 and 2022. When comparing PM concentrations between 2020 and 2021, it becomes evident that the concentration in 2021 surpasses that of 2020, indicating that this difference is statistically significant.

4 Conclusion

In conclusion, this paper presents a comprehensive overview of the current air pollution situation in North Macedonia and its environmental repercussions. Over several years of data collection, the impact of specific events on air quality is evaluated, notably considering the effects of the COVID-19 pandemic, the prohibition of waste burning, and the implementation of the electricity block tariff system for households and small consumers.

The presented findings form the foundation for a discussion on potential strategies and measures aimed at enhancing air quality within the country. The findings indicated a noteworthy reduction in air pollution within regions featuring green spaces when contrasted with regions lacking such spaces. The magnitude of this effect ranged from modest to substantial. In summary, the outcomes imply that the presence of green areas contribute to a decline in air pollution levels, though it is important to acknowledge that additional factors could potentially influence this association. Furthermore, our data demonstrated a rise in air pollutant levels during the winter season, particularly at night. This increase is likely attributable to heightened usage of fossil fuel-based heating systems.

In conclusion, it is crucial to emphasize that ongoing measurement campaigns provide a means to evaluate the influence of diverse factors on air quality. The data gathered through these campaigns, along with their statistical analyses, should serve as the cornerstone for crafting efficient strategies to combat air pollution. Given that air pollution mitigation measures encompass both broad, general approaches and tailored, site-specific interventions, the importance of accumulating data over an extended timeframe cannot be overstated in devising site-specific measures.

6 References

- [1] WHO; Monitoring Ambient Air Quality for Health Impact Assessment. WHO Regional Publications, European Series, 85, 1999.
- [2] Dockery, D; Schwartz, W; Spengler, J; Air pollution and daily mortality: Associations with particulates and acid aerosols. *Environ. Res.*, 59, 362-373, 1992.
- [3] Dockery, D; Schwartz, J; Increased mortality in Philadelphia associated with daily air pollution concentrations. *Am. Rev. Respir. Dis.*, 145, 600-604, 1992.
- [4] EEA, E. E. (2019). Air quality in Europe — 2019 report. European Environment Agency.
- [5] Samet, J; Zeger, S; Dominici, F; Curriero, F; Coursac, I; Dockery, D.; Zanobetti, A; The national morbidity, mortality, and air pollution study. Part II: Morbidity and mortality from air pollution in the United States. . *Res. Rep. Health Eff. Inst.* , 94, 5-79, 2000.
- [6] Mayer, H; Air pollution in cities. *Atmospheric Environment*, 33, 4029-4037, 1992.

- [7] Srbinovska, M.; Krkoleva, A; Andova, A; Celeska, M; Wireless Sensor Networks Implemented in Vertical Green Walls for Air Quality Improvement. In 12th Conference on Sustainable Development of Energy, Water and Environment Systems, pages. 4-8, Dubrovnik, Croatia, 2017.
- [8] Velkovski, B; Srbinovska, M; Dimchev, V; Implementation of a green wall structure in particulate matter reduction using an air quality monitoring system. In IEEE EUROCON 2019-18th International Conference on Smart Technologies, pages 1-6, Novi Sad, Serbia, 2017.
- [9] Srbinovska, M; Andova, V; Krkoleva Mateska, A. Celeska Krstevska, M; The effect of small green walls on reduction of particulate matter concentration in open areas. *Journal of Cleaner Production*, 279: 123306, 2021.
- [10] Andova, V; Andonović, V; Celeska Krstevska, M; Dimcev, V; Krkoleva Mateska, A; Srbinovska, M; Estimation of the Effect of COVID-19 Lockdown Impact Measures on Particulate Matter (PM) Concentrations in North Macedonia. *Atmosphere*. 14(2):192, 2023.

The Ecological Impact of Server-Side Rendering

Mohamed Abdel Maksoud

Codoma.tech Advanced Technologies OÜ

Tallinn, Estonia

mohamed@amaksoud.com

Abstract. *The web platform with its sheer scale plays a vital role in modern society but is energy-intensive. Addressing the urgency of reducing global carbon emissions, we investigate the ecological impact of a rendering technology and estimate potential carbon emission reductions. While web development has evolved considerably over two decades, the ecological consequences remain unclear. Our empirical study reveals substantial carbon emission savings potential, even with conservative estimates. Enhancing the precision of our analysis augments the magnitude of actual emission reductions, promising a more substantial impact.*

Keywords. carbon footprint, sustainability, green computing

1 Introduction

The web platform stands as one of the most ubiquitous and pervasive technologies used today. With over 200 million active websites in existence, its scale is immense, as is its corresponding energy consumption. Given the urgent imperative to curtail global carbon emissions, it is crucial to ensure the optimal operation of such expansive platform, minimizing inefficiencies wherever feasible.

Over the last two decades, web development has evolved significantly from a simple information display in web browsers to a complex process involving various abstractions, frameworks, and server components. These advances offer benefits like faster-loading apps and improved search engine visibility. However, the ecological consequences of these changes remain unclear.

Within this work, we provide empirical evidence concerning the ecological cost associated with a particular rendering technology. Subsequently, we leverage these findings to construct an estimation of the potential aggregate reduction in carbon emissions achievable.

The next section we describe the current technologies and practices in web development with a focus on rendering. Section 3 details our alternative to the current rendering technology and lays out the assumptions we use to carry out the aggregate analysis. Section 4 specifies the experiment performed to compare the two approaches. The empirical results and an aggregate analysis is detailed in Section 5. We finally discuss the results broadly in Section 6.

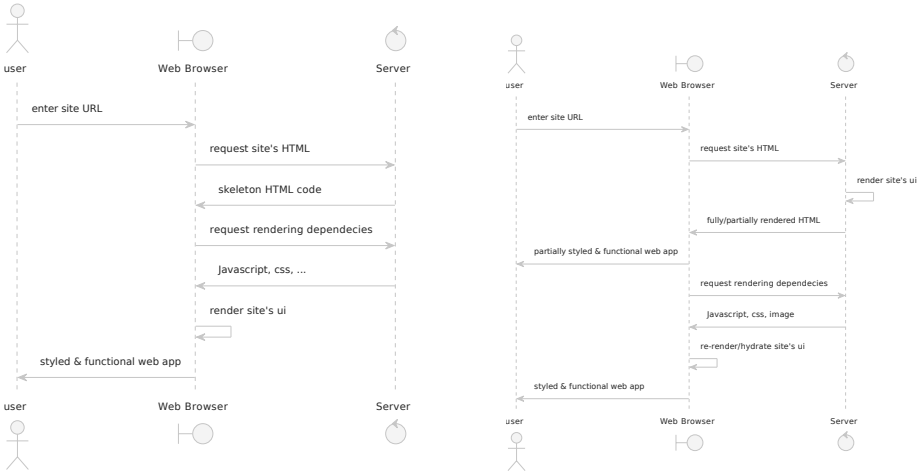


Figure 1: Sequence diagram of client-side rendering (left) and server-side rendering (right).

2 Background

In a three-tier architecture [10], the view layer refers to the visible part of the application. In web development context the view layer is referred as the *frontend* and is developed predominantly in JavaScript, an interpreted language which executes on a run-time engine. The continuous improvements in JavaScript engines made it a viable option to also run on the server as well as the client side. This gave rise to *isomorphic JavaScript*: writing applications which run on both the client and the server. This evolution affected frontend development practices as described in the following subsection.

2.1 Modern Frontend Technologies

Frontend developers currently have multiple options to do *rendering*, the process of converting an application state to an HTML representation. Broadly, the rendering options are:

Client-Side Rendering (CSR): The server sends a blank HTML page along with the JavaScript code and other static dependencies. The browser uses these to display and make the application interactive. The rendering process takes place completely on the client side.

Server-Side Rendering (SSR): The server renders a complete HTML page and sends it to the browser with the application code and assets. The browser does the rendering again, either completely or partially (in a process called *hydration*). This makes it easier for search engines to index the site, and arguably improves the responsiveness of the application.

Figure 1 shows the sequence of operations performed in both rendering techniques. Examining SSR, we can notice a certain degree of redundancy: the application is rendered twice. While this can be beneficial for SEO and responsiveness, there is a price for this repeated computation.

2.2 The Ecological Cost of SSR

Computation has quantifiable impact on the environment. This impact is computed in terms of the equivalent Carbon emissions ($kgCO_2e$) resulting from performing a computation.

In the context of web applications, the Carbon footprint of an application is decided by two factors:

1. The CPU energy the application consumes. We discard the RAM energy consumption since it is negligible in comparison of that of the CPU.
2. The cost of manufacturing the server hardware.

The repeated computation in SSR increases the Carbon footprint of applications. The energy a CPU spends to render the application on the server can be counted as a complete waste. Moreover, rendering JavaScript applications is demanding task and requires larger and more capable servers in comparison to serving static assets.

3 Methodology

3.1 An Efficient SSR Alternative

The work in [3] shows that JavaScript typically falls in the middle range concerning energy efficiency. However, energy efficiency varies significantly depending on the specific application. Therefore, we set to collect concrete measurements specifically addressing SSR.

To quantify the energy consumption associated with SSR implementation in JavaScript, we developed a minimalistic alternative known as *gr8s* [5]. Gr8s is a static asset server with integrated templating capabilities which enable the core SSR feature: serving relevant content based on the page's URL.

A key distinction offered by *gr8s* over conventional SSR is its avoidance of JavaScript code execution for content rendering. Instead, it relies on template replacement using special tags within the HTML code and a designated *data source* that defines the required replacements for each page.

In the development of *gr8s*, we selected Go as the programming language. Go is a compiled language known for its small executable sizes and is widely regarded as performant. Furthermore, Go provides robust primitives for harnessing hardware capabilities, including multithreading and multiprocessor support.

3.2 Quantifying the Cost of Inefficient SSR

Now that we have an alternative implementation, we can compute the energy wasted on current SSR.

We expose both implementations to the same workload and take measurements of energy consumption by CPU and memory utilization.

Empirical Analysis

In this section, we conduct an empirical examination of the experimental data to quantitatively assess the distinctions between the two architectural approaches within the context of a single site.

Deductive Aggregate Analysis

Based on the empirical results, we extrapolate using a set of publicly available statistics and reasonable assumptions to estimate the aggregate cost of a large segment of the sites that employ SSR.

In the pursuit of generating a convincing estimate, we consistently err on the side which favors conventional SSR, thereby ensuring that our estimate approximates its minimum ecological impact. The aggregate analysis uses the following assumptions:

- Due to its prevalence and the availability of usage statistics, our analysis is confined to the `Next.js` framework.
- We factor in the use of content delivery networks (CDN), which significantly reduces computation time for the SSR server.
- In assessing the manufacturing carbon footprint, we assume the all sites are hosted on a pool of servers that are optimally shared and utilized.

4 Experiment Setup

We developed a simple application using `Next.js` and subjected it to a load test of 100,000 requests with a concurrency factor of 100 using the `ApacheBench` tool. We evaluated two server-side rendering (SSR) approaches: first with `Next.js` SSR and then with `gr8s`. Both servers were containerized using `Docker`, facilitating the collection of `RAPL` [1] metrics using the `docker-activity`¹ tool.

Our test was conducted on a server equipped with 8 gigabytes of RAM and Intel Xeon E3-1230 v5 CPU with 8 cores operating at 3.40GHz.

5 Evaluation

The experiment run was concluded over a span of 495 seconds. The `Next.js` server took most of the experiment's time to handle the payload while the `gr8s` server took the last few seconds. Figure 2 shows the CPU energy consumption and memory usage throughout the experiment and zooms in on the last 50 seconds where the `gr8s` server is active.

5.1 Empirical Analysis

Per second, `gr8s` consumes 9.9 times more CPU energy than `Next.js` on average. Nevertheless, `gr8s`'s overall energy consumption is 9.8 times lower than `Next.js`. This is due to the former's considerable reduction in the processing time needed to handle the same payload. From these numbers, we define *per-hour-saving* as the energy saved per hour if we use `gr8s` instead of `Next.js`. From the measurements, we derive it as follows:

$$\begin{aligned} \text{energy-total-nextjs} &= 3527.90 \text{ } j & \text{energy-total-gr8s} &= 360.43 \text{ } j \\ \text{energy-total-saving} &= 3167.47 \text{ } j & \text{time-nextjs} &= 0.135 \text{ } \textit{hours} \end{aligned}$$

¹<https://github.com/jdrouet/docker-activity/>

$$\begin{aligned} \text{per-hour-saving} &= \frac{\text{energy-total-saving}}{\text{time-nextjs} \times 3.5 \times 10^6} \\ &= \mathbf{0.00653 \text{ kWh}} \end{aligned}$$

One more metric relevant to energy consumption is the maximum memory usage. This factor decides the number of servers needed to power a set of websites. While gr8s’s memory usage capped at 82.54 MB, Next . js server reached 411.91 MB to serve the 100,000 requests.

Other performance metrics such as requests per second and latency at 99% strongly favor gr8s over Next . js. They are not discussed here though as they are not directly relevant to the ecological impact.

5.2 Deductive Aggregate Analysis

We use the measurements of per-hour-saving and maximum memory consumption from the previous section to estimate the total energy saving of all Next . js SSR sites. Table 1

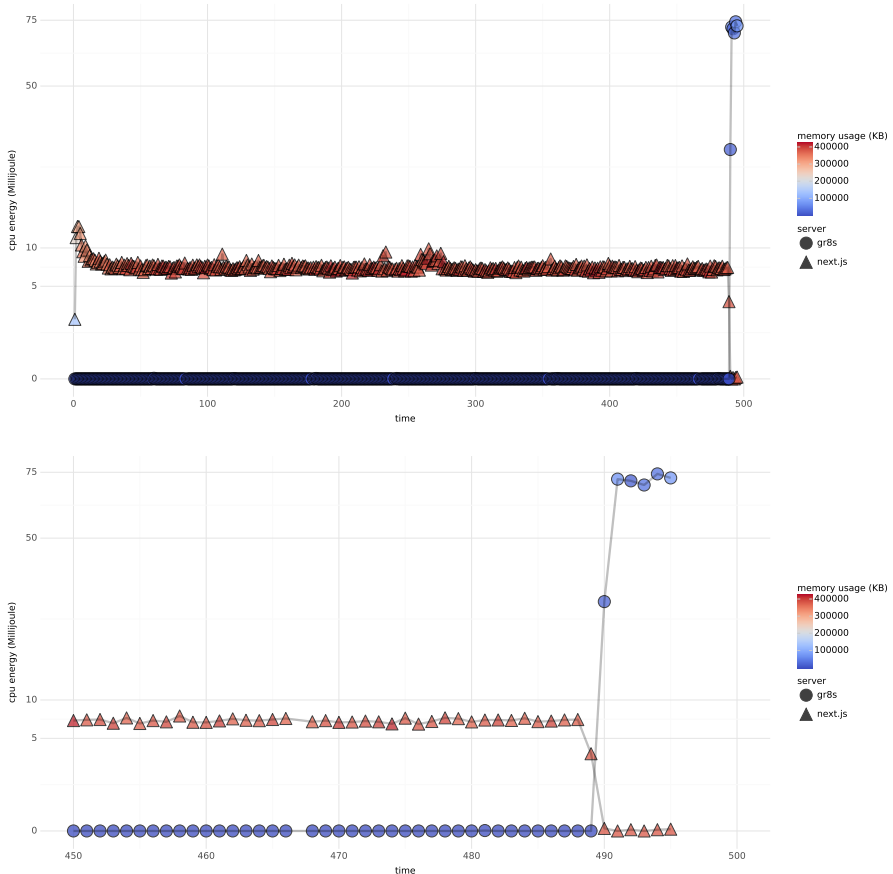


Figure 2: Memory usage and CPU energy for the entire experiment (left) and for the last 50 seconds of the experiment (right) .

lists the relevant figures needed to compute this estimate. We assumed the average data center carbon intensity to be equal to that of the US: the country with the most data centers [4]. We the server’s useful life from [8] as the server’s lifetime.

We deduced some of the numbers listed in this table, namely the number of Next . js virtual servers, the percentage of Next . js sites which use a CDN, and the average server utilization.

To approximate the number of virtual servers required to serve all the Next . js sites, we assume that the number of servers needed by a site is proportional to its PageRank, employing a commonly recognized logarithmic base of 5. For instance, a site with a PageRank of 5 necessitates 25 times more servers than a site with a PageRank of 3. We utilizing this rationale and the data from OpenPageRank [11] to reach the number 1,370,153.

We analyzed around 2,000 sites of the top Next . js sites from [9] and found that 97.9% of them use a CDN. A good CDN hit rate is conventionally between 95% and 99%. We assume the lower end in our analysis.

The 40% server utilization is the average between the 15% (on-premises) and 65% (cloud) as per [2].

Saving related to CPU consumption

Given the CPU-based per-hour-saving and an estimate of the total compute hours of Next . js servers, we can calculate the total saving in carbon emissions.

We start by calculating the compute hours of sites with CDN. A server is active only when there is a CDN miss.

The number of hours per year a CDN site’s server would be active is

$$365 \times 24 \times (1 - 0.95) = 438 \text{ hours}$$

The number of Next . js servers using CDN is

$$1,370,153 \times 0.979 \simeq 1,341,380 \text{ servers}$$

The aggregate number of hours per year is

$$1,010,889 \times 438 = 587,524,440 \text{ hours}$$

We use saving-per-hour to calculate the aggregate annual energy saving of CDN sites

$$587,524,440 \times 0.00653 \simeq 3,836,535 \text{ kWh} \tag{1}$$

Table 1: Relevant aggregate figures of Next . js sites.

Sites using Next . js [6]	1,032,573.00	
Next . js (virtual) servers	1,370,153.00	
Next . js sites with CDN	97.90	%
CDN hit rate	95.00	%
Data center carbon intensity	367.00	gCO2e/kWh [4]
Average server utilization	40.00	%
Shared physical server [7]		
RAM size	128.00	GB
Manufacturing footprint	1,283.18	kgCO2e
Server lifetime	5.00	years

A site without a CDN need significantly more compute time to do server-side rendering.

The number of hours per year a non-CDN site's server would be active can be approximated using an estimate of a server's utilization:

$$365 \times 24 \times 0.4 = 3,504 \text{ hours}$$

The number of Next.js servers without a CDN is

$$1,370,153 \times (1 - 0.979) \simeq 28,773 \text{ servers}$$

The aggregate number of hours per year is

$$28,773 \times 3,504 = 100,820,592 \text{ hours}$$

We use saving-per-hour to calculate the aggregate annual energy saving of non-CDN sites

$$100,820,592 \times 0.00653 \simeq 658,358 \text{ kWh} \quad (2)$$

Finally, the total energy saving is the sum of (1) and (2):

$$\begin{aligned} \text{CPU Annual Saving} &= 4,494,893 \text{ kWh} \\ &\simeq \mathbf{1,649,626 \text{ kgCO}_2\text{e}} \end{aligned} \quad (3)$$

Saving related to server manufacturing

Manufacturing the physical servers has a certain carbon footprint. Here we estimate the saving in carbon emissions if all Next.js sites were to use gr8s for SSR. We use the maximum memory usage to decide the number of physical servers needed. We assume a perfect utilization of the cloud server we consider [7]. This ensures our estimate is conservative.

The number of servers needed to power Next.js sites is

$$\frac{1,370,153 \times 411.91 \text{ MB}}{128 \text{ GB}} \simeq 4,306 \text{ servers}$$

The number of servers needed to power gr8s sites handling the same workload:

$$\frac{1,370,153 \times 82.54 \text{ MB}}{128 \text{ GB}} \simeq 863 \text{ servers}$$

Given manufacturing carbon footprint and the server's life-time, we estimate the annual carbon saving as follows:

$$\begin{aligned} \text{Manufacturing Annual Saving} &= \frac{(4,306 - 863) \times 1,283.18}{5} \\ &\simeq \mathbf{883,598 \text{ kgCO}_2\text{e}} \end{aligned} \quad (4)$$

5.3 Total Annual Carbon Saving

Adding the savings from (3) and (5):

$$\text{Total Annual Carbon Saving} \simeq \mathbf{2,533,224 \text{ kgCO}_2\text{e}} \quad (5)$$

6 Discussion

In this work we explored a potential inefficiency in the widely-used technology SSR. Although we kept our estimates and assumptions as conservative as possible, we still saw a significant potential saving in carbon emissions should SSR be implemented more efficiently.

It is prudent to conduct a more exact analysis on a real-life, large-scale web application. This has the potential of producing relevant, convincing evidence for the industry to adopt more efficient methods to reduce their Carbon footprint.

References

- [1] Huazhe Zhang and Helmut Hoffmann, A Quantitative Evaluation of the RAPL Power Control System. 2014.
- [2] The Natural Resources Defense Council. Data Center Efficiency Assessment. Issue Paper August 2014, IP:14-08-a, <https://www.nrdc.org/sites/default/files/data-center-efficiency-assessment-IP.pdf>.
- [3] Pereira, et. al. Energy Efficiency across Programming Languages: How Do Energy, Time, and Memory Relate? In *Proceedings of the 10th ACM SIGPLAN International Conference on Software Language Engineering*, pages 256-267, 2017.
- [4] Ember. European Electricity Review 2022. <https://ember-climate.org/insights/research/european-electricity-review-2022/>, downloaded: September 5th 2023.
- [5] Codoma.tech Advanced Technologies, gr8s - the graceful smart frontend server. <https://github.com/codomatech/gr8s-server>, downloaded: September 5th 2023.
- [6] BuiltWith Pty Ltd. Next.js Framework Usage Statistics, <https://trends.builtwith.com/framework/Next.js>, downloaded: September 5th 2023.
- [7] Dell Inc. PowerEdge R640 Carbon Footprint, January 2019. https://i.dell.com/sites/csdocuments/CorpComm_Docs/en/carbon-footprint-poweredge-r640.pdf.
- [8] Amazon Inc. SEC Filings Details, Filing Date: Feb 04, 2022, Page 41, <https://d18rn0p25nwr6d.cloudfront.net/CIK-0001018724/f965e5c3-fded-45d3-bbdb-f750f156dcc9.pdf>, downloaded: September 5th 2023.
- [9] Majestic SEO. The Majestic Million, <https://majestic.com/reports/majestic-million>, downloaded: September 5th 2023.
- [10] Eckerson, Wayne W. Three Tier Client/Server Architecture: Achieving Scalability, Performance, and Efficiency in Client Server Applications. *Open Information Systems*, 10(1), 1995.
- [11] DomCop. What Is Open PageRank?, <https://www.domcop.com/openpagerank/what-is-openpagerank>, downloaded: September 5th 2023.

Application of Simulation Modelling for Decision Support in Traffic Safety

Blaž Rodič, Matej Barbo

Faculty of Information Studies

Ljubljanska cesta 31A, 8000 Novo mesto, Slovenia

{blaz.rodic@fis.unm.si, matej.barbo@student.fis.unm.si}

Abstract: *In this contribution we present the case of a real-life decision problem of selecting the best performing collision warning activation algorithm to be used for rear-end collision prevention on motorcycles and other powered two-wheelers (PTW). We have developed a hybrid i.e., multi-method simulation model that allows simulation of various situations in which a vehicle may collide with the rear end of a PTW and used this model to estimate the effectiveness of seven collision warning activation algorithms found in research literature. Simulation results have shown that the Hirst & Graham algorithm consistently outperformed other activation algorithms in all emergency braking scenarios and for all tested speed limits, and in most incidents enabled the following vehicle to brake on time and thus avoid hitting the rear-end of the PTW.*

Key Words: *Agent Based Modelling, Hybrid Simulation Modelling, Road Traffic Simulation, Collision Warning Activation Algorithms, Powered Two-wheeler Safety, Rear-end Collision Warning, Accident Prevention*

1 Introduction

Globally, there are several hundred thousand deaths among motorcyclists annually [1]. The risk of death for all PTW users is 20 times higher than that for car drivers or passengers [2]. To support the case for introduction of new safety technologies as standard equipment on PTWs, we have developed and patented an ESS (Emergency Stop Signal) + RECAS (Rear-End Collision Alert Signal) system called MEBWS (Motorcycle Emergency Braking Warning System) in previous research at the Faculty of Information Studies in Novo mesto [3]. The main innovative aspect of MEBWS is the implementation of combined ESS and RECAS functionality in a self-contained device designed for PTWs.

Emergency Stop Signal or ESS are brake or hazard lights that flash quickly when the driver applies full braking power. ESS is currently offered by some car manufacturers but will become mandatory for new passenger and cargo vehicles in the EU by 2024. Rear-End Collision Alert Signal or RECAS is a supplementary technology that utilizes ESS warning lights in the event of an impending collision with a FV i.e., when time to collision (TTC) drops below the critical threshold.

A common safety technology in new cars is Forward Collision Warning (FCW). Although there is an overlap of functionalities between FCW, ESS and RECAS technologies regarding prevention of rear-end accidents, ESS and RECAS can help prevent accidents

in situations where the driver of FV misinterprets or fails to detect a FCW warning signal, or the FCW system is absent or fails to detect a PTW.

An essential component of a collision warning system is the use of appropriate warning activation algorithm. The biggest technical challenge in maximizing the potential of a collision warning system, and consequently improving the drivers' confidence in the system, is defining an "optimal" activation algorithm that ensures reliable and consistent activation while minimizing the number of false alarms. In this contribution we present the method used to select the best performing algorithm to be used in MEBWS.

2 Methodology

2.1 Simulation Modeling

Our approach to estimating the effectiveness of collision warning activation algorithms is based on simulation modelling using Agent-Based Modeling (ABM), Discrete Event Simulation (DES) and System Dynamics (SD) simulation modelling methodologies. For this purpose, we have developed a hybrid simulation model of rear-end collisions involving PTW as the LV using multiple simulation methodologies, including ABM and SD. The speeds of vehicles are stochastically generated using actual speeds data on EU roads, while the response of the FV and its driver is modeled using the data on driver perception response times to emergency braking with and without flashing emergency stop lights and reaction times of braking system from scientific literature. The collision warning algorithms were compared against a baseline scenario to determine their effectiveness i.e., the ratio of reduction of rear-end collisions. Each simulation run contains one car (FV) and one leading PTW on a straight road section and models exactly one potential rear-end collision situation. For further information about the simulation model please refer to [3].

2.2 Collision Warning Activation Algorithms

The most widely used algorithms are relatively simple temporal-type algorithms [5] based on time to collision (TTC), which is an effective safety indicator for potential collision detection. Collision warning systems using temporal-type algorithms activate when TTC drops below a predefined critical threshold, which in literature varies between 2 and 5 seconds. In this research we have compared the following algorithms: Fujita&Akuzawa&Sato [7], Bella&Russo [8], Hirst&Graham [9], UN/ECE spec. [10] and Modified UN/ECE Spec. [3].

In order to compare the effectiveness of each activation algorithm, we have had to set a baseline scenario for each simulation run. For this purpose, we defined the algorithm "MEBWS Disabled", which represents the natural response of a FV driver by taking into account total reaction time t_f^{total} when setting critical TTC threshold for each simulation incident. The basis of this algorithm is the TTC value of 1.5 seconds which represents the lower perceived limit for the driver to avoid collision by braking, regardless of the vehicle speed.

4 Results

4.1 Comparing Collision Warning Algorithms

Results of *Moto, normal drive* emergency braking scenario are presented in Figure 1, which demonstrates that the “Hirst&Graham” algorithm is the most effective algorithm at all tested speed limits with 73.93% average effectiveness of on-time brakings. Over the whole range of speed limits in the EU, i.e., from 30 km/h to 130 km/h, the probability of an accident was reduced on average by 35.61% (percentage points) when using the “Hirst&Graham” algorithm compared to the baseline driver response of “MEBWS Disabled” with 38.32% average effectiveness, followed by the “Bella&Russo” algorithm with a 22.27% reduction of accident probability (60.59% average effectiveness), the “Mod. UN/ECE Spec.” algorithm with a 19.43% reduction (57.75% average effectiveness), “Fujita&Akuzawa&Sato” algorithm with a 12.57% reduction (50.89% average effectiveness) and “UN/ECE Spec.” algorithm with an 8.20% reduction (46.52% average effectiveness) of accident probability.

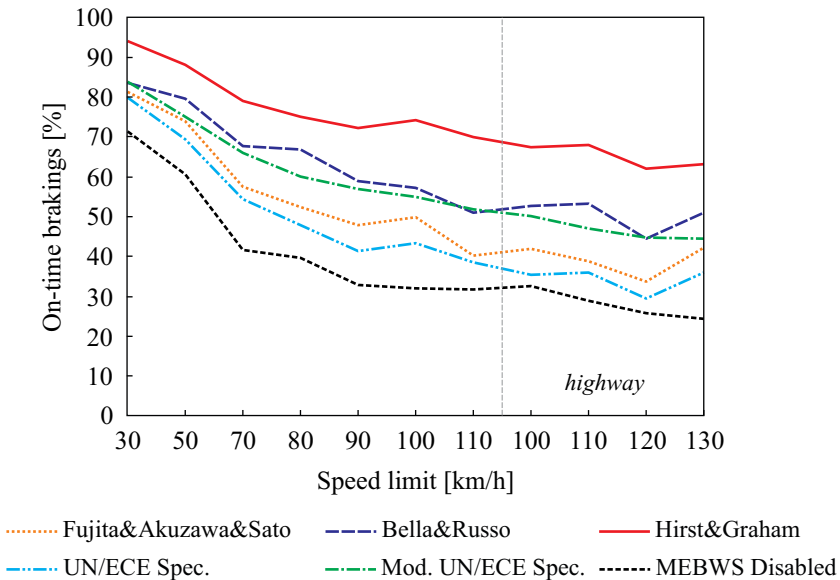


Figure 1: Effectiveness of activation algorithms in *Moto, normal drive* emergency braking scenario, sample size 400, precision level $\pm 5\%$

In the *Moto, emergency stop* emergency braking scenario (Figure 2) most algorithms performed similarly, except for the “Hirst&Graham” algorithm, which again performed significantly better than others, with 59.16% average effectiveness. The “Hirst&Graham” algorithm reduced the probability of an accident on average by 28.50% (percentage points) compared to the baseline driver response of “MEBWS Disabled” with average effectiveness of 30.66%. It is followed by the “Mod. UN/ECE Spec.” algorithm with an 13.43% improvement (44.09% average effectiveness) over the “MEBWS Disabled”, “Bella&Russo” algorithm with a 9.00% improvement (39.66% average effectiveness),

“Fujita&Akuzawa&Sato” algorithm with a 7.73% improvement (38.39% average effectiveness) and “UN/ECE Spec.” algorithm with a 7.70% improvement (38.36% average effectiveness) in number of avoided accidents over the “MEBWS Disabled” baseline driver response.

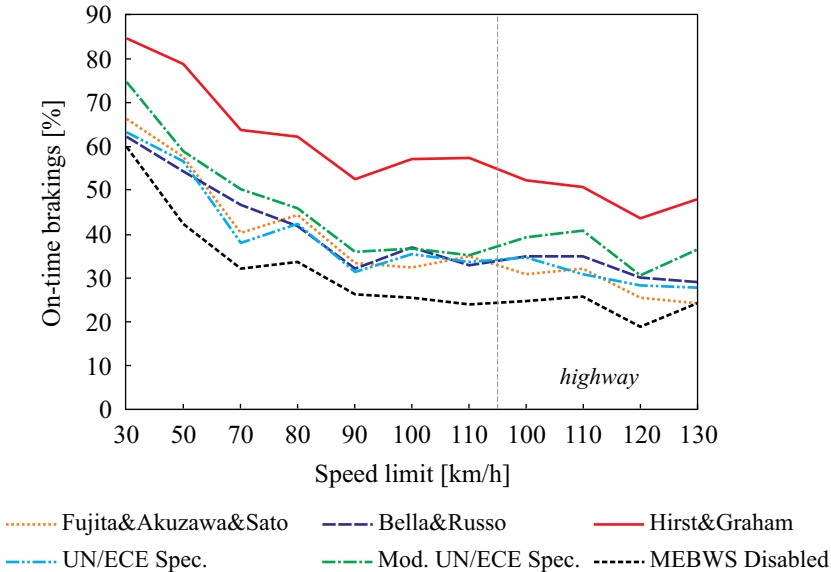


Figure 2: Effectiveness of activation algorithms in *Moto, emergency stop* emergency braking scenario, sample size 400, precision level $\pm 5\%$

Results of *Moto, not moving* emergency braking scenario are presented in Figure 3. The “Hirst&Graham” algorithm was with 60.34% average effectiveness the most effective also in this scenario and has prevented on average 29.52% (percentage points) of accidents more than the “MEBWS Disabled” baseline driver response with average effectiveness of 30.82%. Algorithm “Mod. UN/ECE Spec.” performed 14.77% (45.59% average effectiveness) better than “MEBWS Disabled”, while “Fujita&Akuzawa&Sato” showed a 9.98% (40.70% average effectiveness), “Bella&Russo” a 7.80% (38.61% average effectiveness) and “UN/ECE Spec.” a 7.50% (38.32% average effectiveness) improvement over the “MEBWS Disabled” baseline driver response.

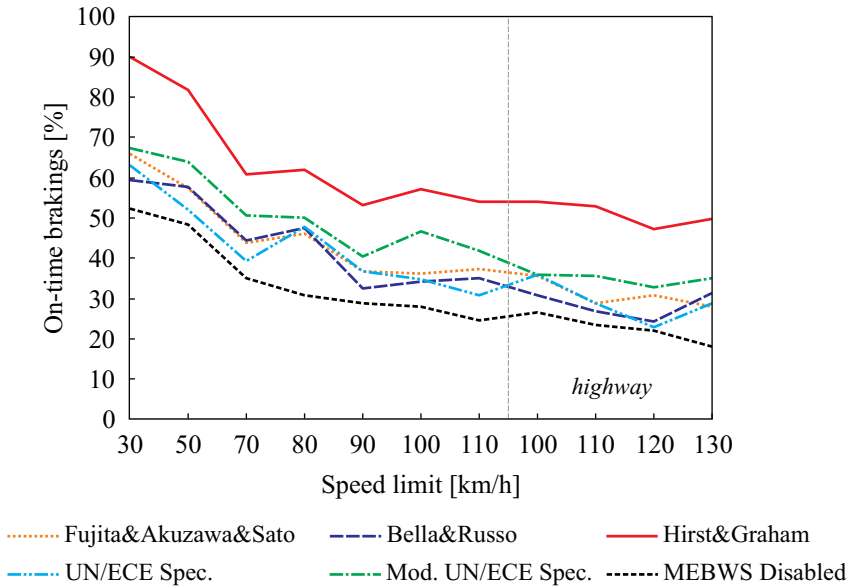


Figure 3: Effectiveness of activation algorithms in *Moto, not moving* emergency braking scenario, sample size 400, precision level $\pm 5\%$

The analysis of the effectiveness of the activation algorithms showed that “Hirst&Graham” algorithm consistently outperformed other activation algorithms in all emergency braking scenarios and for all tested speed limits, and in most incidents enabled the FV to brake on time and thus avoid rear-ending the PTW. Therefore we chose the “Hirst&Graham” algorithm to be implemented into MEBWS and conduct further simulation experiments. The number of false alarms would be likely higher with the “Hirst&Graham” algorithm than with other algorithms, but as the consequences of even a low speed collision can result in hospitalization of PTW users [13], accident prevention is arguably more important than minimization of false alarms.

5 Conclusion

To support the introduction of safety features such as MEBWS as standard equipment on PTWs, we have carried out a study on the effectiveness of collision warning activation algorithms to be used in motorcycle safety technology. The “Hirst&Graham” algorithm has proved to be overall the best of the tested algorithms; hence we chose it as the algorithm to be implemented into MEBWS.

5.1 Limitation of results

Limitations of the current version of the model include the focus on rear-end collisions and the modeling of the road transport system as a set of scenarios using probability distributions of traffic amount per vehicle category and actual speeds, based on publicly

available EU statistical data. Both limitations are the results of a conscious choice. We believe that the selected level of abstraction of the model yields sufficiently accurate results while keeping the model transparent and allowing it to run on all main personal computing platforms.

8 Acknowledgements

This work was supported by the Slovenian Research and Innovation Agency (Research program No. P1-0383, Complex networks).

9 References

- [1] H. Ospina-Mateus, L. A. Quintana Jiménez, F. J. Lopez-Valdes, and K. Salas-Navarro, ‘Bibliometric analysis in motorcycle accident research: a global overview’, *Scientometrics*, vol. 121, no. 2, pp. 793–815, Nov. 2019, doi: 10.1007/s11192-019-03234-5.
- [2] S. de Craen, M. J. A. Doumen, N. Bos, and Y. van Norden, ‘The roles of motorcyclists and car drivers in conspicuity-related motorcycle crashes’, SWOV, Leidschendam, Dec. 2011. Accessed: Mar. 30, 2022. [Online]. Available: <https://www.swov.nl/file/15659/download?token=ahORThzg>
- [3] M. Barbo and B. Rodič, ‘Modeling the influence of safety aid market penetration on traffic safety: Case of collision warning system for powered two-wheelers’, *Accid. Anal. Prev.*, vol. 192, p. 107240, Nov. 2023, doi: 10.1016/j.aap.2023.107240.
- [4] J. F. Lenkeit and T. Smith, ‘Preliminary Study of the Response of Forward Collision Warning Systems to Motorcycles’, in *Proceedings of the 11th International Motorcycles Conference*, in Forschungshefte Zweiradsicherheit, vol. 17. Cologne, Köln: Institut für Zweiradsicherheit, Oct. 2016, pp. 1–24. Accessed: Feb. 02, 2022. [Online]. Available: https://lindseyresearch.com/wp-content/uploads/2019/05/NHTSA-2018-0092-0017-Preliminary_Study.pdf
- [5] S. M. S. Mahmud, L. Ferreira, M. S. Hoque, and A. Tavassoli, ‘Application of proximal surrogate indicators for safety evaluation: A review of recent developments and research needs’, *IATSS Res.*, vol. 41, no. 4, pp. 153–163, Dec. 2017, doi: 10.1016/j.iatssr.2017.02.001.
- [6] F. Bella, A. Calvi, and F. D’Amico, ‘An empirical study on traffic safety indicators for the analysis of car-following conditions’, *Adv. Transp. Stud.*, vol. 2014 Special Issue, pp. 5–16, Jan. 2014, doi: 10.4399/97888548735442.
- [7] Y. Fujita, K. Akuzawa, and M. Sato, ‘Radar Brake System’, *JSAE Rev.*, vol. 16, no. 1, p. 113, Jan. 1995, doi: 10.1016/0389-4304(95)94875-N.
- [8] F. Bella and R. Russo, ‘A Collision Warning System for rear-end collision: a driving simulator study’, *Procedia - Soc. Behav. Sci.*, vol. 20, pp. 676–686, 2011, doi: 10.1016/j.sbspro.2011.08.075.
- [9] S. Hirst and R. Graham, ‘The Format and Presentation of Collision Warnings’, in *Ergonomics and Safety of Intelligent Driver Interfaces*, Y. I. Noy, Ed., in Human factors in transportation. , Mahwah, New Jersey: Lawrence Erlbaum Associates, 1997, pp. 203–219.
- [10] Economic Commission for Europe of the United Nations, ‘Regulation No 48 of the Economic Commission for Europe of the United Nations (UNECE) — Uniform provisions concerning the approval of vehicles with regard to the installation of lighting and light-signalling devices [2019/57]’, *Off. J. Eur. Union*, vol. 62, no.

- L14, pp. 42–146, Jan. 2019, Accessed: May 04, 2021. [Online]. Available: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:42019X0057>
- [11] J. Gail, M. Lorig, C. Gelau, D. Heuzeroth, and W. Sievert, *Optimization of Rear Signal Pattern for Reduction of Rear-End Accidents during Emergency Braking Maneuvers*. Bergisch Gladbach: Federal Highway Research Institute, 2001. Accessed: Aug. 05, 2020. [Online]. Available: https://bast.opus.hbz-nrw.de/opus45-bast/frontdoor/deliver/index/docId/288/file/emergency_braking.pdf
- [12] R. van der Horst and J. Hogema, ‘Time-to-collision and Collision Avoidance Systems’. TNO Institute for Human Factors, Jan. 1994. Accessed: Jul. 17, 2022. [Online]. Available: https://www.researchgate.net/publication/237807114_TIME-TO-COLLISION_AND_COLLISION_AVOIDANCE_SYSTEMS
- [13] E. Hardy, D. Margaritis, J. Ouellet, and M. Winkelbauer, ‘The Dynamics Of Motorcycle Crashes: A Global Survey of 1578 Motorcyclists’, Feb. 2020. Accessed: Mar. 03, 2023. [Online]. Available: <https://investigativeresearch.org/the-dynamics-of-motorcycle-crashes-2020/>

Automating privacy compliance

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

Boštjan Delak,
Faculty of information studies
Ljubljanska cesta 31a, 8000 Novo Mesto, Slovenia
bostjan.delak@fis.unm.si

Abstract: *Organizations are facing increasing demands of personal data legislation that require allocating resources for tackling privacy issues or risking very high fines. As information privacy presents a complex legal, organizational, and technical challenge, organizations must decide how to effectively deal with such issues. One possible approach is to implement principles of privacy by design into personal data processing operations to achieve compliance with a clear set of criteria. The authors of this article have developed a conceptual model that can be used in the context of any organization. The proposed model was successfully tested on four different organizations within the health sector. This paper outlines privacy by design concept, its transformation into a clear conceptual model, the process of implementation, and finally possibilities to automate the process with the use of available AI tools.*

Key Words: *privacy by design, conceptual model, personal data, artificial intelligence*

1 Introduction

In the modern information society, increasing emphasis is placed on the field of personal data protection as it expresses the concern for the right to privacy of individuals and represents an effort to comply with legal requirements. Compliance became particularly important in the EU after the introduction of the General Data Protection Regulation (GDPR). The USA currently does not have similar legislation on the federal level; however, an increasing number of states, for example, California, Virginia, Colorado, Connecticut, and Utah [1] are adopting information privacy laws. One promising feature of GDPR is promoting the concept of privacy by design which originates from Cavoukian's [2] influential essay.

However, as the concept is vaguely defined, we analyzed existing literature to find specific elements that could then be composed into a coherent model. The next challenge was to develop a procedure for the implementation of the model on data processing operations and use various case studies to test the model. Results were encouraging and were published in two articles [6], [7].

In this article, we summarize key findings regarding the development and implementation of conceptual model. We also reflect on possibilities to automate the processes with the use of widely available AI tools, and address some challenges that arise from such approach.

2 Methods

2.1 Literature review

The conceptual model of privacy by design [6] is based on an analysis of several approaches to implementing privacy protection measures in IT. The starting point was Cavoukian's [2] key principles of privacy by design, followed by three extensive meta-analyses [14], [11], [16] which enabled us to explore existing attempts to tackle privacy issues. For example, Cavoukian's principles of privacy by design were preceded by authors such as Denning [5] and Chaum [3] who as early as the eighties wrote about Privacy-Enhancing Technologies, which became formalized in Fair Information Privacy Principles promoted by the US Federal Trade Commission. Other authors [17], [9], [10], [4] focused their efforts on developing the concept of Privacy impact assessment which later became formalized in GDPR under the term DPIA. Hoepman [10] also emphasized the importance of privacy protection strategies, Foukia et al. [8] developed a privacy framework called PISCES, Jensen et al. [12] developed project STRAP, Kalloniatis et al. [13] did similar work with the framework PRIS. Even the EU Commission undertook a similar challenge by forming a consortium of research institutions and within the PRIPARE project developed a handbook called Privacy and Security by Design Methodology [15] with privacy guidelines for IT developers. A lot of work was done by different standard organizations, for example, ISO/IEC, which adopted personal data protection standard ISO/IEC 27701:2019.

2.2 Conceptual model

The purpose of this analysis and comparison of different approaches was to determine whether there exist elements of personal data protection that are common to all compared approaches. Common elements were identified and used as building blocks for the conceptual model of privacy by design. They were grouped into one of three sets: "legal elements", "security elements", and "privacy by design and by default elements" [6]. The sets are consistent with the structure of GDPR where legal elements occupy a central position, followed by data security, and finally by privacy by design and by default provisions.

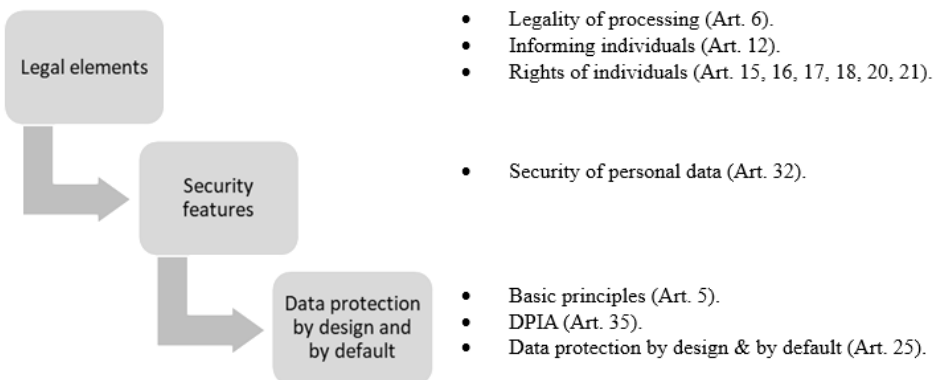


Figure 1: Representation of the conceptual model of privacy by design [6]

2.3 Implementation procedure

A general outline of the model was then operationalized by adding an implementation procedure and compliance matrix for scoring audited processes. Implementation consists of 5 steps which start with information gathering, proceeding with analysis of legal, security, and data protection by design and by default elements, and finally closing with a report that includes gap analysis and recommendations on how to improve information privacy compliance.

Table 1: Representation of the conceptual model of privacy by design with groups of elements that correspond with GDPR provisions [6]

Sequence of steps	Step description
1 Information gathering	Information was gathered from legal documents, company website, intranet, internal documents, DPIA, contracts with processors, interviews with the head of the IT, data protection officer, IT administrators, security consultant, process owners, and data processor contract managers. Personal data processing operations and registers were then determined.
2 Analysis of legal elements	Information was analyzed for the presence of legal elements in the following sequence: Determining the purpose of data processing- Finding an appropriate legal basis for data processing. Determining how the transparency of data processing is ensured. Determining mechanisms for exercising the rights of individuals according to GDPR. Determining how contractual processing is arranged.
3 Analysis of security elements	Information was analyzed for the presence of security elements. Each data process was checked from the viewpoint of ensuring proper confidentiality, integrity, and availability of personal data. Audit measures were also checked.
4 Analysis of privacy by design and by default elements	Information was analyzed for the presence of privacy by design and by default elements. Analysis was performed in the following sequence: Determining how data encryption is ensured both when transferring and storing data. Determining if and how data minimization is ensured. Determining if and how data pseudonymization is ensured. Analysis of (existing) DPIA.
5 Final Report with gap analysis and recommendations	Doing gap analysis (difference between the actual state of data processing and benchmarks). Preparing recommendations for assuring a higher level of compliance. Preparing and presenting a final report.

Table 2: Representation of the conceptual model of privacy by design with groups of elements that correspond with GDPR provisions [6]

Processing operation / presence of privacy by design model elements	Legal elements					Security elements			Data protection by design and by default			Compliance	
	Legality of processing	Informing individuals	Rights of individuals	Processing agreement	Transferring data outside EU	Confidentiality	Integrity	Accessibility	Encryption	Pseudonymity	Data minimization	Basic (GDPR)	Upgraded (PbD)
Legend: 1 – element not present 2 – element is present, mayor inadequacy 3 – element is present, minor inadequacy 4 – element is fully present													
Processing operation													

3 Results

Implementation of the model was first tested on the Slovenian central health information system (eHealth) in 2021 [7]. In 2023, the model was tested on three other healthcare organizations, however, the results have not yet been published. In all cases, the implementation procedure was done according to the steps in Table 1. For each personal data process, elements were identified as per Table 2 and scored from the perspective of compliance with GDPR. In the final report, recommendations were made on how to improve data processing operations.

During the implementation of the model, several issues arose. Scoring of security elements proved difficult as GDPR is not clear on giving detailed guidelines about what level of information security is sufficient. On top of this, larger and more complex organizations usually require extensive security measures. However, taking the same route as ISO/IEC standards with detailed checklists would inevitably reduce the simplicity and usability of the model which should apply to all types and sizes of organizations.

We therefore tried to identify a key set of criteria for assessing security elements. This set was much more limited than ISO/IEC control lists, thus sacrificing precision and depth. But in the end, it did prove to be easier and more flexible to use.

The same challenge arose when identifying and scoring privacy by design and by default elements as GDPR only briefly outlines what those are. Here the relatively new ISO/IEC 27701:2019 standard proved to be very useful as it enabled us to pinpoint those elements.

4 Discussion

The first case study was the most extensive. As one of the authors of this article is employed by the National Institute of Public Health which maintains eHealth, the scope of available information for assessing the model was much greater, however, there was also a greater possibility of bias when scoring the elements and assessing state of compliance. In the other three studies, the situation was different. Those studies were less extensive since not all data processing operations were analyzed and data was gathered through representatives. All case studies showed that the conceptual model can be used in a relatively straightforward, predictable, and systematic way. These qualities encouraged reflection on automating the process of implementation. Two approaches seem especially promising. One is the use of special software that enables quick assessment of GDPR compliance by using the proposed model as an algorithm. Another is the use of available AI tools to perform assessments based on specified criteria. This option seems most interesting; however, some issues should be carefully reconsidered. One is the potential disclosure of personal data that is stored on web AI platforms. Another is loss of transparency when implementing the model since it is much more difficult to determine how did the AI interface interpret the information and assess the level of compliance. Though with adequate measures, those risks could be minimized.

5 Conclusion

The implementation of the proposed conceptual model was tested on four different organizations. Results have practical and theoretical implications. From a practical perspective, it is clear that a systematic approach to ensuring privacy compliance is possible, and it can be relatively straightforward. It is also possible to determine the gap between actual and target levels of privacy, although there have been difficulties attempting to quantify the gap. From a theoretical point of view, case studies were important because they provide partial answers to specific research questions, whether the proposed conceptual model is sufficiently general for different contexts of personal data processing, whether it allows determining the actual situation of personal data processing and related identification of compliance with the target state, and whether it allows more effective data protection against alternative approaches. The results are also promising regarding automation of the process with the use of software applications or available AI interfaces, where last option seems most promising, however, it should be noted it is not without risks, especially when one is dealing with sensitive personal data.

6 References

- [1]. 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.
- [2]. Cavoukian, A. Privacy by Design. The 7 Foundational Principles, <https://www.ipc.on.ca/wp-content/uploads/Resources/7foundationalprinciples.pdf>. 2009.
- [3]. Chaum, D. Security without Identification Card Computers to make Big Brother Obsolete. *Communications of the ACM* 28(10): 1030-1044,1985.
- [4]. 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.
- [5]. Denning, D.E. *Cryptography and Data Security*. Addison-Wesley Publishing Company, Boston, USA, 1982.
- [6]. Drev, M.; Delak, B. Conceptual Model of Privacy by Design. *Journal of Computer Information Systems*, 62(5), 888-895, 2021.
- [7]. 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.
- [8]. 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.
- [9]. 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.
- [10]. Hoepman, J. H. Privacy Design Strategies. *IFIP International Information Security Conference*, pages 446–459, Berlin, Germany, 2014.
- [11]. 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.
- [12]. 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.
- [13]. Kalloniatis, C.; Kavakli, E.; Gritzalis, S. Addressing Privacy Requirements in System Design: The PriS Method. *Requirements Engineering*, 13(3):241-255, 2008.
- [14]. Kurtz, C.; Semmann, M.; Böhm, 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.
- [15]. 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.
- [16]. 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.
- [17]. Spiekermann, S.; Cranor, L. F. Engineering privacy. *IEEE Transactions on Software Engineering*, 35(1), 67–82, 2009.

The Necessity of Digital Transformation of Developing Countries: The Case of Bosnia and Herzegovina

Dino Arnaut, Damir Bećirović
International Business-Information Academy Tuzla
Kulina bana br. 2, 75000 Tuzla, Bosnia and Herzegovina
{dino, direktor}@ipi-akademija.ba

Abstract: *Digitalization is a key driver of economic development, promoting regional cooperation and investor attraction. The Covid-19 pandemic has highlighted digital transformation and energy transition as future directions. This paper aims to demonstrate the need for digital transformation and emphasize its benefits for developing countries like Bosnia and Herzegovina. Analyzing digital transformation in Bosnia and Herzegovina using nine international indices helps evaluate a country's digital maturity and provides insights for interventions and investments. Understanding a country's digital readiness is crucial for creating a more inclusive digital future. Bosnia and Herzegovina lags in digital transformation compared to neighboring countries due to insufficient data collection methodologies and lack of inclusion in indices. The country faces challenges in public administration, innovation, and business sector disruption. To improve, the government must create a conducive environment for investment and innovation in digital technologies.*

Key Words: *Digital Transformation, Digital Economy, Public Management, Change Management, Bosnia and Herzegovina.*

1 Introduction

Organizational change implementation is a challenging task. Regardless of the measures achieved in the past towards more effective organizational change management, change programs consistently demonstrate a high level of failure. It is not surprising that most recent research focuses on the examination of the traits of successfully implemented organizational change initiatives in the public and private sectors. Many empirical studies on digital governance and the transformation of the public sector are qualitative in nature and based on case studies, which can only provide a theoretical viewpoint.

Digitalization is a key driver of economic development, impacting macroeconomic and microeconomic aspects. It promotes regional economic cooperation and enhances a country's attractiveness for investors. The Covid-19 pandemic has highlighted digital transformation and energy transition as key future directions. Bosnia and Herzegovina need stronger pressure to catch up with digital trends, potentially increasing growth, and economic cooperation with other regions.

By generating new kinds of value based on digitalization, innovation, and digital technology, digitalization primarily aims to contribute to the transformation of the economy, public sector, and society. The public sector's performance and efficiency are also goals of digital transformation. Therefore, the purpose of this paper is to demonstrate the need for digital transformation, increase awareness of it, and emphasize how developing countries like Bosnia and Herzegovina can benefit from its implementation.

The analysis of digital transformation in Bosnia and Herzegovina was conducted using nine international indices. These indices consider various factors such as supply and demand conditions, institutional environment, innovation, connectivity, human capital, integration of digital technologies, digital public services, research and development, regulation, knowledge, business environment, and readiness for the future. These dimensions help evaluate a country's digital maturity and provide insights for interventions and investments to help advance in digital readiness transformation. Understanding a country's digital readiness is crucial for creating a more inclusive digital future for all.

2 Digital Transformation

Digital transformation in the public sector involves new stakeholder engagement, service delivery frameworks, and connections. However, there is limited empirical evidence on how public administrations define digital transformation, their approach to projects, and the expected results from this process [1].

The literature uses terms like digitalization, digital management, and digital transformation interchangeably. Leaders should create an environment for empowerment and continuous improvement, while employees must have a unique vision of digital transformation. Leadership directly impacts digital maturity [2]. The success of digitalization in the public sector requires strong central leadership and preventive measures from local and regional actors [3].

Digital technology alone doesn't add value to organizations; organizational change is an evolving phenomenon, and its application in specific situations allows companies to discover new value creation methods, emphasizing the need for business model rethinking [4,5]. The effectiveness of e-government and digital management systems is significantly influenced by citizens' perception of the value they derive from these platforms [6].

The management of information systems in public administration is becoming increasingly complex. Improving communication between different structures within states and between them can be achieved by redirecting existing information flows, redesigning current applications, and developing new applications that efficiently use available information resources [7-9]. The quality of information technology systems in public administration significantly impacts decision-making systems [10]. New information programs and procedures are needed to support the process of improving public institution management, reducing corruption, and improving the business environment [11-13].

Technological changes, along with the economy, significantly impact society and public administration worldwide. Countries that fail to integrate digital technology across sectors will suffer. The digitalization of public services in developed countries signals the need for digital transformation to promote transparency, high-quality services, and combat corruption. The digital revolution has transformed modern economies, corporations, and public administration. State governments should continue modernizing public administration and services through the integration of information and communication technologies [14,15].

Digitalization has proven successful in reducing corruption in countries where digital reforms took place [16]. Digitalization of public organizations is essential for a smart community, increasing efficiency and transparency. Digitalization simplifies the operation of public institutions at internal, intra-institutional, and external levels, promoting transparency and openness for a democratic society [17-19].

Digitalization is transforming the workforce, increasing the need for professional work, and potentially causing job losses and wage inequality. This digital evolution affects not only the number of jobs created or lost but also the composition of available jobs [20]. The effects of AI are visible at the individual level, with a high risk of workers changing occupations or even losing employment. While cutting-edge digital technologies have little effect on the overall employment rate, they lead to a large movement of workers within occupations and industries [21]. The disruptive effect of AI on employment can take various forms, affecting different stages of evolution [22].

3 The Case of Bosnia and Herzegovina

The integrated assessment approach or the index method can be used to compare the degree of development of specific economic phenomena across nations. According to a comparative analysis of the theoretical frameworks for measuring the degree of digitalization [23-26], using international indices is the quickest and most efficient way to gather data on the digital transformation. Based on it, and with the intention of evaluating Bosnia and Herzegovina's current state of digital transformation objectively, the integrated evaluation approach, or the index method, was applied in this paper.

The nine most important metrics that are brought to the digital transformation are considered in the examination. An overview of Bosnia and Herzegovina's ranking in relation to the other ranked countries was provided after the methodology for each individual index was explained. Based on the indicators included in each individual index, it was noted which areas of the digital transformation Bosnia and Herzegovina performs the worst in, i.e., the most poorly compared to other leading countries. Table 1 provides an organized evaluation and presentation of the examined indices according to the level of digitalization [27].

The Digital Intelligence Index (DII) is the third edition of the Digital Evolution Index, providing insights into global digital development, key factors driving change, and the impact of digital trust and evolution on a country's digital competitiveness. It offers actionable insights for improving digital competitiveness, fostering trust in the digital economy, and promoting responsible data use [28].

The Digital Economy and Society Index (DESI) is a European Commission-developed composite index that measures EU countries' progress in promoting the digital economy and society. It consists of five main policies: connectivity, human capital, internet services, integration of digital technology, and digital public services [29]. In 2021, the index has been updated to focus on four main areas in the Digital Compass, replacing the previous five-dimensional structure. The rankings have been recalculated for all countries to reflect changes in the selection of indicators and corrections made to the data [30].

Cisco's Digital Readiness Index (DRI) is a comprehensive measure of a country's digital readiness, offering guidance on how to foster an inclusive digital economy. It includes factors beyond technology, such as basic needs, human capital development, business environment, and education. Access to technology and infrastructure is crucial, but basic needs like clean drinking water and education are also essential. DRI categorizes countries into three stages: Activate, Accelerate, and Strengthen. In Activate, countries focus on basic needs and human capital development, while in Acceleration, investments in business ease of doing business are essential. Regardless of the stage of digital readiness, human capital development is crucial for building a workforce capable of using and creating technology [31].

The Digital Adoption Index (DAI), a tool developed by the World Bank Group, measures digital adoption in three economic dimensions: people, government, and

business. It focuses on the supply side of digital adoption, with the total DAI calculated as the average of these dimensions [32].

Table 1: Digital transformation indices [27]

Index	Key drivers and indicators	Creator
DEI - <i>Digital Evolution Index</i>	Supply conditions, demand conditions, institutional environment and innovation and change.	University Tufts
DESI - <i>Digital Economy and Society Index</i>	Connectivity, human capital/digital skills, internet use, digital technology integration, digital public services, ICT research and development.	European Commission
DRI - <i>Digital Readiness Index</i>	Basic needs, human capital, ease of doing business, business and government investment, start-up environment, technology infrastructure, technology adoption.	Cisco
DAI - <i>Digital Adoption Index</i>	People, government, and business.	World Bank Group
GII - <i>Global Innovation Index</i>	80 indicators in categories: political environment, education, infrastructure, and business sophistication.	University Cornell, INSEAD, World Intellectual Property Organization
EDI - <i>Enabling Digitalization Index</i>	5 components: regulation, knowledge, connectivity, infrastructure, and size.	Euler Hermes
DiGiX - <i>Digitization Index</i>	6 dimensions: infrastructure, household adoption, business adoption, costs, regulation, and content.	BBVA Research
DCI - <i>Digital Competitiveness Index</i>	Knowledge, technology, and readiness for the future	World Competitiveness Center
ICTDI - <i>ICT Development Index</i>	ICT readiness, the intensity of ICT and the impact of ICT	International Telecommunication Union

The Global Innovation Index (GII) is an annual ranking of innovation ecosystems in 132 countries, tracking global trends and strengths and weaknesses. It consists of around 80 indicators, including political environment, education, infrastructure, and knowledge creation. The latest edition provides new data and analysis, enabling policymakers to compare performance in over 130 countries. The GII also offers insight into global innovation’s pulse, including the Covid-19 pandemic, showing resilience in investment in innovation. The index’s metrics can be used to monitor performance and compare development against countries within the same region or income group [33].

The Enabling Digitalization Index (EDI) evaluates countries’ support for digitalization and ranks them based on friendly digital regulation, institutional, logistical, and technical aspects. It measures the organizational environment and government support for technical innovation, aiming to help digital companies thrive and traditional

businesses take advantage of the digital dividend. The EDI scores countries from 0 to 100 based on regulation, knowledge, connectivity, infrastructure, and size [34].

The Digitization Index (DiGiX) measures a country's ability to fully utilize information and communication technologies for increased competitiveness and prosperity. It is a composite index of 21 sub-indicators for 99 countries worldwide, focusing on supply conditions, demand conditions, and institutional environment. DiGiX is structured around six main dimensions, each divided into individual indicators. In 2020, global improvements in the digital frontier were observed [35].

The Digital Competitiveness Index (DCI) ranks 63 global countries based on 52 criteria, including 32 quantitative data and 20 survey data. It provides comparisons based on population size, GDP per capita, and regional rankings from Europe-Middle East-Africa, Asia-Pacific, and the Americas. The index provides a detailed examination of specific aspects of digital transformation, allowing for comparisons between countries and assessing a country's technological framework. These rankings can support international investment decisions [36].

The ICT Development Index (ICTDI) was a composite index used to monitor and compare the development of information and communication technologies between countries. In 2017, a revised set of 14 indicators was adopted, but challenges in data collection and reporting arose. This led to half of the data having to be estimated for the 2018 ICTDI calculation. The index's methodological accuracy was not possible due to these shortcomings. Despite attempts to harmonize the index or develop a new one, consensus was not reached within expert groups, and the index will not be published until new agreements are reached.

Table 2 provides a summary of the top five nations' rankings on the indices of digital transformation, as well as the ranking or absence of Bosnia and Herzegovina. The table also displays Bosnia and Herzegovina's major vulnerabilities as seen throughout the ranking within each of the indices. We shall discuss in more detail which factors are impeding Bosnia and Herzegovina's digital transformation journey in the sections that follow.

Bosnia and Herzegovina ranks 77th out of 90 countries in the Digital Intelligence Index, amidst challenges in digitalization and driving potential. Despite infrastructure gaps, younger demographics are enthusiastic about the digital future, using social media and mobile payments. However, skepticism towards digitalization and technology, particularly from government institutions, is a significant issue. Rank is low in innovation (87/90), with a limited scope in digitalization. To boost innovation, the country needs to focus on talent development, collaboration between universities and industry, and the development of new digital products and services. Also rank is low in the institutional environment (82/90), with government policies playing a crucial role in supporting or hindering the business sector in creating and distributing digital technologies. A stable environment that encourages investment and protects consumers is essential for digitalization.

Bosnia and Herzegovina is not ranked in the Digital Economy and Society Index, with 27% of indicators still unavailable. However, it is moderately prepared with 73% of indicators aligned with the DESI methodology.

Bosnia and Herzegovina ranks 74th out of 146 countries in the Digital Readiness Index. The country has the worst results in two components: start-up environment and human capital. The start-up environment lacks venture capital and patent registration, which are crucial for creating wealth from digital technologies and job creation and brain drain increases especially in skilled labor force that has potential to support digital innovation.

Table 2: Ranking of Bosnia and Herzegovina

Index	Top 5 countries	B&H	Biggest weaknesses	Last ranking
DEI/DII	1. Singapore 2. USA 3. Hong Kong 4. Finland 5. Denmark	77/90	– Innovations – Institutional environment	2020
DESI	1. Finland 2. Denmark 3. Netherlands 4. Sweden 5. Ireland	Not ranked - missing data for 10 indicators	– Digital public services – Connection	2022
DRI	1. Singapore 2. Luxembourg 3. USA 4. Denmark 5. Switzerland	74/146	– Start-up environment – Human capital	2021
DAI	1. Singapore 2. Luxembourg 3. Austria 4. Korea 5. Malta	60/180	– The weakest adoption in the economy dimension - people	2016
GII	1. Switzerland 2. Sweden 3. USA 4. UK 5. Singapore	77/132	– Business sophistication – Institutions	2023
EDI	1. USA 2. Germany 3. Denmark 4. Netherlands 5. UK	Not ranked	Not ranked	2019
DiGiX	1. Denmark 2. USA 3. Singapore 4. Netherlands 5. Finland	Not ranked	Not ranked	2022
DCI	1. Denmark 2. USA 3. Sweden 4. Singapore 5. Switzerland	Not ranked	Not ranked	2022
ICTDI	1. Ireland 2. Korea 3. Switzerland 4. Denmark 5. UK	83/176	– Use of ICT	2017

Bosnia and Herzegovina ranks 60th in the Digital Adoption Index, with significant shortcomings in the economy related to people. Digital technologies can promote inclusion by increasing employment and earnings in the ICT sector and supporting jobs in sectors using ICT through new technology adoption. Bosnia and Herzegovina needs to increase job creation through existing businesses, entrepreneurship, and outsourcing, as well as increasing worker productivity through digital technologies. This can also benefit consumers by automating processes and generating economies of scale, leading to price reduction and new product and service creation.

The Global Innovation Index ranks Bosnia and Herzegovina in 77th place among 132 countries, 37th in Europe. The country falls in institutional, business, and regulatory environment. Improvements are needed in fostering entrepreneurial culture and nurturing trust in country institutions. To improve business sophistication, Bosnia and Herzegovina should focus on increasing cooperation between universities and businesses in research and development, developing clusters, and increasing knowledge absorption through intellectual property, high-tech imports, ICT services, and direct foreign investments.

The ICT Development Index ranks Bosnia and Herzegovina 83rd out of 176 countries, with 69.33% of individuals using the internet, 17.37 subscribers per hundred inhabitants to fixed broadband, and 37.35 subscribers per hundred inhabitants to mobile broadband connections, according to data from the index.

Bosnia and Herzegovina is not included in the Enabling Digitalization Index, Digitization Index, or the Digital Competitiveness Index.

4 Conclusion

Digital transformation is a crucial process for Bosnia and Herzegovina, as it aims to improve its digital competence and infrastructure. However, despite not being at the bottom of ranked countries, Bosnia and Herzegovina lags in this process compared to neighboring countries. This is due to the fact that many indices do not include Bosnia and Herzegovina in their calculations. To improve its digital transformation, Bosnia and Herzegovina needs to work on data collection methodologies and become part of indices that have not yet included Bosnia and Herzegovina in their rankings. The EDI index, which includes 115 countries, has not included Bosnia and Herzegovina, as well as the DiGiX index, which includes 99 countries. The DCI index, which includes some Balkan countries, does not include any of the Western Balkan countries in its ranking.

The digital transformation process in Bosnia and Herzegovina faces significant obstacles due to its economy's inability to become more competitive. The country's efficiency in public administration and innovation is low, with no adequate legal framework and coordination. Government policies disrupt the business sector, which is responsible for digital technology creation and distribution. Therefore, governments must create a suitable climate for investment and innovation in digital technologies, providing a stable environment that encourages investment and protects consumers. This will create favorable conditions for digitization and contribute to Bosnia and Herzegovina's digital transformation.

Bosnia and Herzegovina is lagging in business innovation, particularly in intangible assets, due to a lack of global brand value and new organizational models. To improve, the country should foster start-up business environments, invest in entrepreneurial capital, and protect innovations. This will create wealth from digital technologies and create jobs. Additionally, the country should increase knowledge absorption through intellectual property protection, high-tech imports, ICT services, and foreign investment inflows.

According to the literature, there are several prerequisites and obstacles to a successful transformation of government into a digital one. Technological constraints are just one of them. Numerous instances demonstrate how organizational, institutional, and legal obstacles frequently obstruct governments from adopting and utilizing new technologies. This is typically explained by the fact that new technologies are expected to jeopardize nearly every government policy, strategy, and organizational structure. These alterations, however, are large and complicated. The transformation component, which denotes a shift from the digitalization of public services to more major changes in the way government functions, is typically considered in the literature as the ultimate goal of digital governance development. To maintain this transition, institutional and regulatory factors must also be modified in addition to the organizational processes.

5 References

- [1] Eggers W, Bellman J. The journey to government's digital transformation. Deloitte 2015.
- [2] Xanthopoulou P, Dimitrios Karampelas I. The Impact of Leadership on Employees' Loyalty and on Organizational Success: Do Transformational and Transactional Leadership Ensure Organizational and Work Commitment? *International Journal of Sciences* 2020; 9:45–63. <https://doi.org/10.18483/ijsci.2389>.
- [3] Millard J. Is the bottle half full or half empty? *European Journal of ePractice* 2010:1–16.
- [4] Osterwalder A, Pigneur Y. *Business Model Generation*. John Wiley & Sons; 2010.
- [5] Morakanyane R, Grace A, O'Reilly P. Conceptualizing Digital Transformation in Business Organizations: A Systematic Review of Literature. *Digital Transformation – From Connecting Things to Transforming Our Lives* 2017. <https://doi.org/10.18690/978-961-286-043-1.30>.
- [6] Scott M, DeLone W, Golden W. Measuring eGovernment success: a public value approach. *European Journal of Information Systems* 2016; 25:187–208. <https://doi.org/10.1057/ejis.2015.11>.
- [7] Nica E, Stan CI, Luțan AG, Oașa R-Ș. Internet of Things-based Real-Time Production Logistics, Sustainable Industrial Value Creation, and Artificial Intelligence-driven Big Data Analytics in Cyber-Physical Smart Manufacturing Systems. *Economics, Management, and Financial Markets* 2021;16:52. <https://doi.org/10.22381/emfm16120215>.
- [8] Kovacova M, Lăzăroiu G. Sustainable Industrial Big Data, Automated Production Processes, and Cyber-Physical System-based Manufacturing in Smart Networked Factories. *Economics, Management, and Financial Markets* 2021;16:41. <https://doi.org/10.22381/emfm16320212>.
- [9] Novak A, Bennett D, Klietstik T. Product Decision-Making Information Systems, Real-Time Sensor Networks, and Artificial Intelligence-driven Big Data Analytics in Sustainable Industry 4.0. *Economics, Management, and Financial Markets* 2021;16:62. <https://doi.org/10.22381/emfm16220213>.
- [10] Bednářová L, Michalková S, Vandžura S. Public procurement in the conditions of the Slovak Republic concerning the participants in the procurement. *International Journal of Entrepreneurial Knowledge* 2021; 9:67–80. <https://doi.org/10.37335/ijek.v9i1.124>.
- [11] Ardielli E. Evaluation of eParticipation service's availability on Czech municipal websites. *International Journal of Entrepreneurial Knowledge* 2020; 8:19–33. <https://doi.org/10.37335/ijek.v8i2.99>.
- [12] Maris M. Municipal changes in Slovakia. The evidence from spatial data. *European*

- Journal of Geography 2020; 11:58-72. <https://doi.org/10.48088/ejg.m.mar.11.1.58.72>.
- [13] Góđány Z, Machová R, Mura L, Zsigmond T. Entrepreneurship Motivation in the 21st Century in Terms of Pull and Push Factors. TEM Journal 2021:334–42. <https://doi.org/10.18421/tem101-42>.
- [14] Dumitrica DD. Robin Mansell: Imagining the Internet. Communication, Innovation and Governance. Oxford: Oxford University Press. 2012. *MedieKultur: Journal of Media and Communication Research* 2015; 31:177–80. <https://doi.org/10.7146/mediekultur.v31i58.20476>.
- [15] Rymarczyk J. The impact of industrial revolution 4.0 on international trade. *Entrepreneurial Business and Economics Review* 2021; 9:105–17. <https://doi.org/10.15678/eber.2021.090107>.
- [16] Mouna A, Nedra B, Khaireddine M. International comparative evidence of e-government success and economic growth: technology adoption as an anti-corruption tool. *Transforming Government: People, Process and Policy* 2020; 14:713–36. <https://doi.org/10.1108/tg-03-2020-0040>.
- [17] Șandor SD. Measuring Public Sector Innovation. *Transylvanian Review of Administrative Sciences* 2018:125–37. <https://doi.org/10.24193/tras.54e.8>.
- [18] AfonasoVA MA, Panfilova EE, Galichkina MA, Ślusarczyk B. Digitalization in Economy and Innovation: The Effect on Social and Economic Processes. *Polish Journal of Management Studies* 2019; 19:22–32. <https://doi.org/10.17512/pjms.2019.19.2.02>.
- [19] Balzer R, Užík M, Glova J. Managing Growth Opportunities in the Digital Era – An Empiric Perspective of Value Creation. *Polish Journal of Management Studies* 2020; 21:87–100. <https://doi.org/10.17512/pjms.2020.21.2.07>.
- [20] Fossen FM, Sorgner A. The effects of digitalization on employment and entrepreneurship. Conference Proceeding Paper, IZA – Institute of Labor Economics n.d.
- [21] Arntz M, Gregory T, Zierahn U. Digitalization and the Future of Work: Macroeconomic Consequences. *SSRN Electronic Journal* 2019. <https://doi.org/10.2139/ssrn.3411981>.
- [22] Ping H, Yao ying G. Comprehensive view on the effect of artificial intelligence on employment. *Topics In Education, Culture and Social Development* 2018. <https://doi.org/10.26480/ismiemls.01.2018.32.35>.
- [23] Biegun K, Karwowski J. Macroeconomic imbalance procedure (MIP) scoreboard indicators and their predictive strength of multidimensional crises? *Equilibrium Quarterly Journal of Economics and Economic Policy* 2020; 15:11–28. <https://doi.org/10.24136/eq.2020.001>.
- [24] Roszko-Wójtowicz E, Grzelak MM. Macroeconomic stability and the level of competitiveness in EU member states: a comparative dynamic approach. *Oeconomia Copernicana* 2020; 11:657–88. <https://doi.org/10.24136/oc.2020.027>.
- [25] Zolkover A, Renkas J. Assessing The Level Of Macroeconomic Stability Of EU Countries. *SocioEconomic Challenges* 2020; 4:175–82. [https://doi.org/10.21272/sec.4\(4\).175-182.2020](https://doi.org/10.21272/sec.4(4).175-182.2020).
- [26] Yarovenko H, Bilan Y, Lyeonov S, Mentel G. Methodology for assessing the risk associated with information and knowledge loss management. *Journal of Business Economics and Management* 2021; 22:369–87. <https://doi.org/10.3846/jbem.2021.13925>.
- [27] Arnaut D, Bećirović D. Level of Digital Transformation in Bosnia and Herzegovina. *Proceedings of 5th International Scientific Conference on Digital Economy DIEC 2022* 2022; 5:87–106.

- [28] Chakravorti B, Chaturvedi RS, Filipovic C, Brewer G. Digital in the time of Covid. Trust in the digital economy and its evolution across 90 economies as the planet paused for a pandemic. The Fletcher School at Tufts University 2020.
- [29] Digital Economy and Society Index (DESI) 2021 - Thematic chapters. European Commission 2021.
- [30] Digital Economy and Society Index (DESI) 2021 - DESI methodological note. European Commission 2021.
- [31] Cisco Global Digital Readiness Index 2019 - White paper. Cisco Public 2020.
- [32] World Development Report, 2016: digital dividends. Choice Reviews Online 2016; 53:53–4889. <https://doi.org/10.5860/choice.196952>.
- [33] Global Innovation Index 2021: Tracking Innovation through the COVID-19 Crisis. World Intellectual Property Organization 2021.
- [34] Enabling Digitalization Index: Beyond potential. Euler Hermes 2019.
- [35] Cámara N. DiGiX 2020 Update: A Multidimensional Index of Digitization. BBVA Research 2020.
- [36] IMD World Digital Competitiveness Ranking 2022. IMD World Competitiveness Center 2022.

Clustering of Employee Absence Data: A Case Study

Peter Zupančič

Faculty of Information Studies

Ljubljanska cesta 31a, 8000 Novo mesto, Slovenia

`peter.zupancic@fis.unm.si`

Panče Panov

Jožef Stefan Institute

Jamova cesta 39, 1000 Ljubljana, Slovenia

`pance.panov@ijs.si`

Abstract. *This paper provides an initial exploration of the emerging field of workforce analytics, focusing on the analysis of employee absence data. Using data from the MojeUre.com system, which collects and aggregates employee data across different companies, regions, and work types, we aim to uncover the hidden patterns and details associated with absences. In this paper, we employ cluster analysis and compare three clustering techniques: k-means clustering, agglomerative clustering, and biclustering. The goal is to identify the most appropriate approach for segmenting absence data into unique and meaningful clusters. The comparative analysis in this paper provides valuable initial insights for organizations seeking to leverage the power of data analytics for workforce management and strategic decision making.*

Keywords. people analytics, absence data, machine learning, clustering, historical data

1 Introduction

In the rapidly evolving landscape of the modern workplace, organizations are constantly looking for ways to leverage data to gain competitive advantage. Amidst the vast array of available data sources, a niche area has emerged that is gaining traction: "people analytics" [1]. Essentially, people analytics is about leveraging employee-related data to derive actionable insights that can inform strategic decision-making and improve business performance.

At the heart of people analytics is the principle that employees, often referred to as an organization's most valuable asset, produce a wealth of data that, when accurately collected and analyzed, can offer profound insights into their behavior, productivity and overall well-being [2]. One specific facet of this data that has been underutilized but has immense potential is employee absence data.

Absenteeism, whether due to illness, personal reasons or other factors, not only impacts the individual employee, but affects the entire organization [3]. Frequent absences, for example, can hinder teamwork, delay the completion of projects and lead to higher

operating costs. Consequently, understanding the patterns, predictors and impact of employee absenteeism is paramount.

While there are several analytical approaches to analyze these data, this paper will focus on descriptive analysis and clustering techniques. Descriptive analysis, as the name implies, provides a comprehensive overview of the data by summarizing the most important aspects, while clustering aims to divide the data into different groups based on certain characteristics or patterns [4]. Such techniques can help organizations identify specific clusters or patterns of absenteeism, determine the underlying causes, and then take targeted action.

The advantages of such an analytical approach are manifold. First, it can help organizations prevent potential problems by identifying employees or teams at risk. Second, by identifying the causes of frequent absenteeism, organizations can implement personalized wellness programs or provide additional resources to promote a healthier work environment. Finally, understanding absence patterns can help with workforce planning, ensuring projects and teams are staffed appropriately and remaining employees are not overly burdened.

Clustering in Human Resources Management (HRM) utilizes data analysis techniques to group employees based on shared attributes such as job roles, skills, or performance metrics. This segmentation empowers HR professionals to tailor talent management strategies, optimize recruitment efforts, and address specific needs for retention, skill development, and diversity and inclusion initiatives. By categorizing employees into clusters, organizations can make data-driven decisions, fostering more effective HR practices and a deeper understanding of their workforce dynamics.

In this paper, we focus on a case study of clustering employee absence data from the MojeUre system. In our first study, we use one year of employee absence data together with demographics attributes and absence aggregate attributes to compare the results of different clustering algorithms, such as k-means clustering, agglomerative clustering, and biclustering.

2 The task of clustering employee absence data

Grouping workers based on patterns in their absenteeism is at the core of clustering worker absenteeism data. The goal is to uncover latent patterns and trends in the workforce to gain a more accurate understanding of absenteeism. When employees who exhibit similar patterns of absenteeism are grouped together, organizations can identify and address certain behaviors or challenges that are specific to certain groups.

Cluster analysis, a cornerstone of machine learning, focuses on segmenting unlabeled data sets into distinct groups or clusters, each containing similar data points. The premise is to identify inherent similarities within the data and divide it into groups to ensure that the data within one group is more similar compared to those in other groups. Key attributes such as shape, size, and behavior are evaluated to identify these patterns. Given its unsupervised nature, the algorithm autonomously finds these groupings on its own without prior labeling. After clustering, each group is given a unique cluster ID, which can greatly simplify the processing of large data sets [5].

Accurate records of employee absences, whether due to illness or vacation, are essential to a company's efficiency. Such data helps understand workforce availability and supports planning. Clustering absence data – for example, grouping employees from similar regions who exhibit analogous absence patterns – enables a company to allocate re-

sources wisely, effectively address absence-related issues, and ultimately make informed decisions that optimize both costs and workforce management.

3 Dataset description

Our dataset was generated from a timekeeping system called MojeUre, which is used to record each employee’s attendance and absence times, as well as other human resource management data. The period we consider in our analysis is one year, exactly the year 2019.

After refining our initial dataset, we obtained a streamlined set with 3637 unique instances and 388 distinct attributes. Each instance or row represents a specific employee. The attributes within the dataset are diverse in nature. While some attributes are pre-defined in the database, others are aggregated features derived from various attributes.

A key criterion in our data cleaning process was the VacationLeaveTotalDays score for each employee. If an employee had less than one vacation day within the specified time period, this could indicate several scenarios. The employee could no longer be employed by the company, the company could have stopped using the recording system, or the employee could be a student for whom vacation days are not normally recorded because they are typically compensated on an hourly basis.

Figure 1 shows the structure of the employee data used in the case study. One data instance represents one employee. In general, our dataset provides a rich, multi-layered perspective on each employee that includes demographic details, aggregated absence metrics, and a granular one-year absence profile.

Employee demographic data (5 attributes)	Aggregated absence data (18 attributes)	One year absence profile data (per day – 0/1) 365 attributes
--	---	--

Figure 1: The structure of employee data used in the case study.

The demographic segment deals with the individual and occupational characteristics of each employee. It includes the average number of hours an employee works per day, known as the "WorkHour". In addition, we categorize employees by the type of company in which they’re employed, with up to 31 different company types available under "CompanyType". The dataset also highlights each employee’s length of service under "EmploymentYears" and distinguishes their employment status through "JobType", which includes categories such as full-time and part-time. Finally, the company region is identified, with 12 different regions listed under "Region".

The aggregated absence data provide information on the leave habits of individual employees. They provide an overview of the total number of vacation and sick days taken by an employee in a year. The data further breaks down absences into short-term (3 days) and long-term (5 days) categories for vacation and sick leave. In addition, a seasonal analysis shows the number of vacation and sick days taken in winter, spring, summer and fall. This seasonal breakdown is supplemented by data on vacation days taken during school holidays in each of the four seasons.

Complementing this, the one-year absence profile provides a comprehensive, day-by-day overview of an employee’s attendance over the course of a year. The 365 attributes of each day indicate whether an employee was present or absent, marked by a binary '0' for presence and '1' for absence.

4 Methodology and experiments

4.1 Methodology description

Our analysis followed a structured approach to uncover patterns and clusters in our dataset. This methodology consisted of several steps, beginning with dimensionality reduction using principal component analysis (PCA) [6], followed by the application of clustering algorithms and visualization techniques.

Before engaging in clustering, we performed PCA on specific subsets of our dataset, namely the one-year absence profile and the aggregate absence metrics. PCA was intended to transform these high-dimensional data into a new coordinate system in which the greatest variance in any projection of the data was on the first few coordinates. This effectively reduced the dimensionality of our data, allowing us to focus on the most important patterns.

After transforming our data by PCA, we applied the k-means clustering algorithm [7]. Since this algorithm is sensitive to the choice of 'k' (number of clusters), we experimented with different k values to determine the optimal number of clusters for our data. In doing so, we ran through a range of k values and evaluated the results in terms of coherence and separation.

To visualize the results of our k-means clustering, we plotted the data points on a two-dimensional plane defined by the first two principal components of PCA. This provided a clear visual representation of how the data points were clustered. To enrich our visualization, we overlaid the demographic information about the employees. This allowed us to see possible correlations or patterns between demographic characteristics and the clusters formed.

Following our k-means clustering, we adopted a hierarchical clustering approach, specifically agglomerative clustering [8]. In this bottom-up strategy, each data point is considered as a separate cluster and clusters are paired sequentially based on their proximity to each other. We tested different linkage criteria, such as single, complete, average, and Ward's method, to investigate how these criteria affect the formation of hierarchical clusters.

Finally, we applied biclustering specifically to the one-year absence profiles. Unlike traditional clustering, which groups data in a single direction (rows or columns), biclustering clusters rows and columns simultaneously [9]. This technique was particularly well suited for our one-year absence profiles because it allowed us to find subgroups of employees with similar absence patterns during specific time periods.

4.2 Experimental setup

For our clustering study, we used the Python programming language and accessed the scikit-learn¹, scipy², and numpy³ libraries. All of our tests ran on the Anaconda/Spyder platform⁴. The following sections describe the setup for each algorithm we used in our experiments.

¹<https://scikit-learn.org/>

²<https://scipy.org/>

³<https://numpy.org/>

⁴<https://anaconda.org/anaconda/spyder>

4.2.1 Principal Components Analysis

For our analysis, we employed Principal Component Analysis (PCA) as a crucial preparatory step. PCA serves to reduce the complexity of our high-dimensional dataset by transforming it into a more manageable size while preserving its primary patterns and structures. Specifically, we extracted the first two principal components from our dataset, which are the most informative axes capturing the maximum variance. These components were then used as the foundational axes for visualizing the results of our k-means clustering, providing a clear, two-dimensional representation of the cluster groupings and enabling a more intuitive understanding of the clustering outcomes.

4.2.2 k-Means clustering

In our study, using the k-means clustering algorithm, we investigated in detail the effects of varying the number of clusters, parameterized by k . In particular, we experimented with values of k ranging from 2 to 5. Our dataset consisted mainly of aggregated absence data supplemented by a detailed one-year absence profile. To improve the interpretability of our results, clustering assignments were visualized in the context of the previously identified principal components. To enrich the visualization and provide deeper context, additional employee demographic data – such as Work hour, Company type, Employment years, and Job types also integrated into the visual representation. This approach ensured that our visual results were not only clear but also provided meaningful and actionable insights.

The silhouette score [10] is a metric used to calculate the goodness of a clustering algorithm. It measures how close each point in a cluster is to points in neighboring clusters. Its values range from -1 to 1. A high value indicates that the object is a good match to its own cluster and a poor match to the neighboring clusters. If most objects have a high value, the cluster configuration is appropriate. If many items have a low or negative value, then the cluster configuration may have too many or too few clusters. In the context of k-means clustering, the silhouette score provides an insightful tool for determining the optimal number of clusters k by comparing scores for different k values and selecting the configuration that yields the highest silhouette score.

4.2.3 Agglomerative clustering

Before proceeding with agglomerative clustering, we standardized all attributes to ensure uniformity of scale across the dataset. Our data consisted of both aggregated absence records and a detailed absence profile for one year. In the clustering method, different linkage methods were explored. The Ward method aims to minimize the variance of the distances between the clusters to be merged, making it particularly effective for clusters of similar size. The Complete linkage, on the other hand, defines the distance between two clusters based on the maximum pairwise distance between their elements, which often results in compact clusters. The Average linkage method calculates the average of all pairwise distances between elements in the clusters, resulting in more balanced cluster sizes. Finally, the Single linkage method, which determines cluster distance based on the smallest pairwise distance between elements, can be sensitive to outliers and lead to elongated, ‘chain-like’ clusters.

To illustrate our results, we used a dendrogram that clearly represented the cluster hierarchy. This dendrogram highlighted the number of examples in the most granular

clusters, provided insight into the distribution and density of our data points, and allowed us to understand the decisions made at each iterative merging step.

4.2.4 Biclustering analysis

Spectral biclustering works by simultaneously clustering rows and columns of a data matrix, essentially grouping rows and columns with similar properties. Using the eigenvalues of matrices derived from the original data, biclusters of different structures are captured by creating coherent partitions in both dimensions, which in the context of our data refers to the simultaneous clustering of employees (rows) and absence (columns) patterns. For our experiment, the analysis was conducted using data from a one-year absence profile, with the number of clusters serving as a variable parameter that was adjusted to observe different outcomes. A key step in our process was to pre-select employees, using demographic data, specifically the attributes 'CompanyType' and 'Region', to narrow down specific areas or regions so that the analysis could even focus on employees from a single company. The resulting clusters were visualized using a heat map, a graphical representation that uses color gradients to intuitively represent the clusters, with similar colors denoting similar absence patterns, providing a clear and immediate visual guide to interpreting the clustered data and understanding the emerging patterns and trends.

5 Results and discussion

5.1 k-Means clustering

In k-means clustering we generate different graphs showing how data can be grouped into different clusters, kind of like categories. We decide how many clusters there should be using variable called K. Each graph helps us see how well the data fits into these clusters.

Figure2 shows the results of k-means clustering with $k = 2$, visualized using the two principal components. The red dots indicate the centers of these clusters. In addition, the clusters are enriched with data on the job types of the employees. Noticeably, one cluster consists predominantly of workers with permanent contracts. In contrast, the second cluster includes not only employees with student status, but also external collaborators. This suggests possible differences in the absence profiles of these groups.

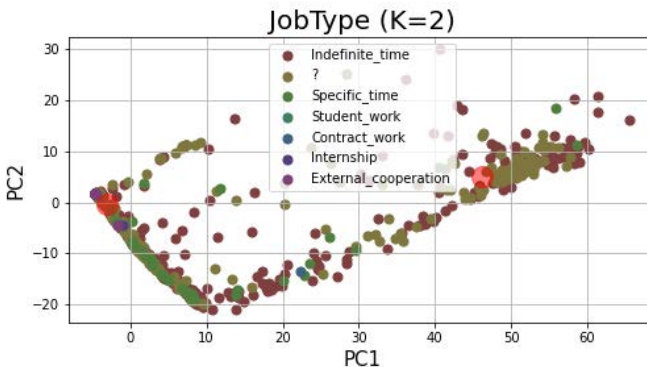


Figure 2: Example of k-means clustering with information about employee JobType as an overlay.

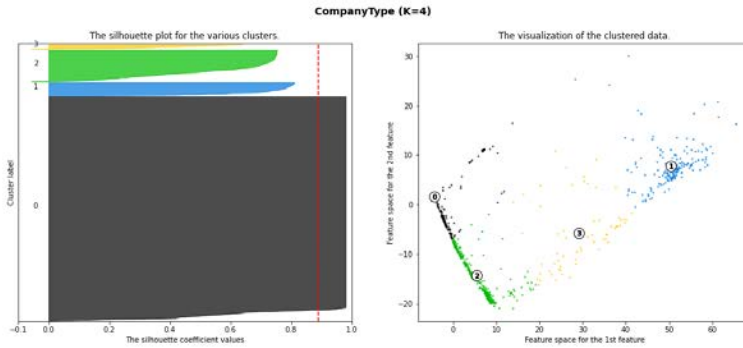


Figure 3: K-means clustering illustration with $k = 4$. On the right, the figure displays the cluster assignments, while the left visualizes the silhouette coefficients for each cluster.

Figure 3 shows the results of k-means clustering for $k = 4$, using the first two principal components for visualization. On the left side, the plot of silhouette coefficient values provides insight into the clustering quality. The right side provides a colored representation of the clusters with clearly marked centers. Cluster 0 stands out as the largest and includes the majority of employees. It's noteworthy that the silhouette coefficient values for all clusters are very high, indicating that the clusters are well defined and appropriately configured. It's also noted that Cluster 3 has data points that are widely dispersed within the group, indicating that they're very different from each other. Cluster 2, on the other hand, has data points that are more similar, meaning that they aren't as spread out.

In simple terms, we use different charts to understand how the data groups and how similar or different things are within those clusters. This helps us understand the data and make decisions based on our findings.

5.2 Agglomerative clustering

In Figure 4, we present an example result of using average linkage. This method calculates the distance between clusters by taking the average of all pairwise distances between data points in one cluster and data points in the other cluster [11]. Numerical taxonomy. The principles and practice of numerical classification.). In our agglomerative clustering visualization, we employed truncation to make the dendrogram more readable, especially given the large size of our original dataset. The parameter k plays a crucial role here. Two different scenarios are depicted: one with $k=2$ and another with $k=3$. In both scenarios, the last k non-singleton clusters formed are the only non-leaf nodes showcased in the dendrogram, simplifying its structure. This means, in the first example where we delve into two clusters and in the second where we explore three clusters, we are only displaying the last k significant merges in the dendrogram. In both graphs, the final cluster contains most of the data, suggesting a high similarity among the majority of employees in that cluster.

In Figure 5, we display dendrograms constructed using the ward linkage method. We explore two truncation levels: $K = 2$ and $K = 3$. Notably, the dendrogram with $K = 3$ offers a more detailed tree structure compared to its $K = 2$ counterpart. This granularity allows for a more nuanced clustering of the data. Intriguingly, in both truncations, the first and second clusters comprise the smallest subsets of employees, indicating these groups

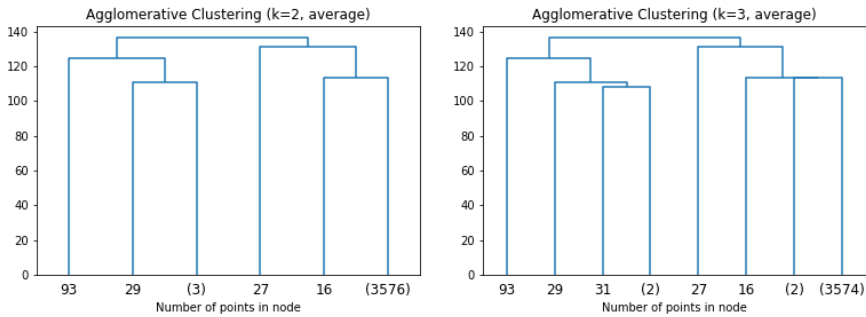


Figure 4: Agglomerative clustering using average linkage

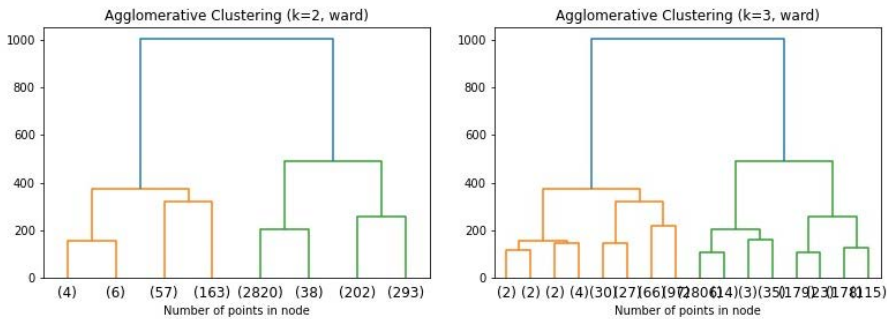


Figure 5: Agglomerative clustering using ward linkage

possess distinct attributes. Conversely, the fifth cluster in both instances encompasses the majority of employees, indicating shared or similar characteristics among them.

5.3 Biclustering analysis

In Figure 6, we hone in on the one-year absence profiles of employees from the 'Drava' region, identified by the "Region" attribute. By employing biclustering with $k = 2$ and $k = 3$, we aim to segment the data into two and three clusters respectively. This stratification may shed light on the distribution patterns of various companies within the Drava region over distinct employee categories and time frames. For example, using $k = 2$ could differentiate between major and minor company groups in the area, whereas $k = 3$ might further discern specific subsegments within these groups.

In Figure 7, we focus on companies characterized by the type "production." Through biclustering with parameters $k = 4$ and $k = 5$, our aim is to discern four and five clusters, respectively. These clusters aim to provide insights into the workforce trends of production-based companies over varying numbers of days.

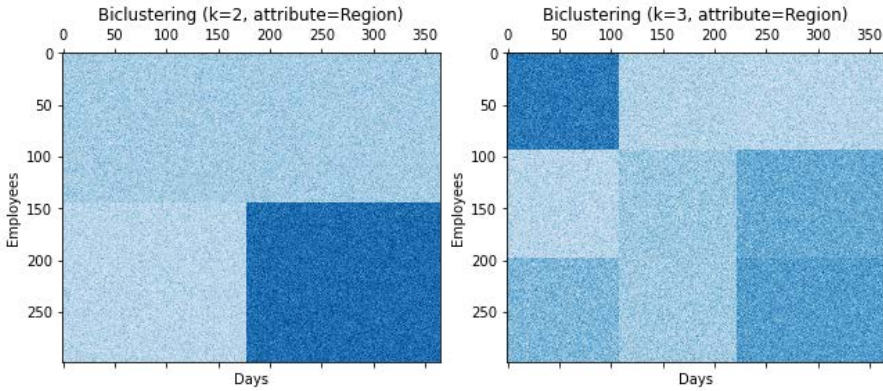


Figure 6: Results of biclustering on preselected employees from the 'Drava' region.

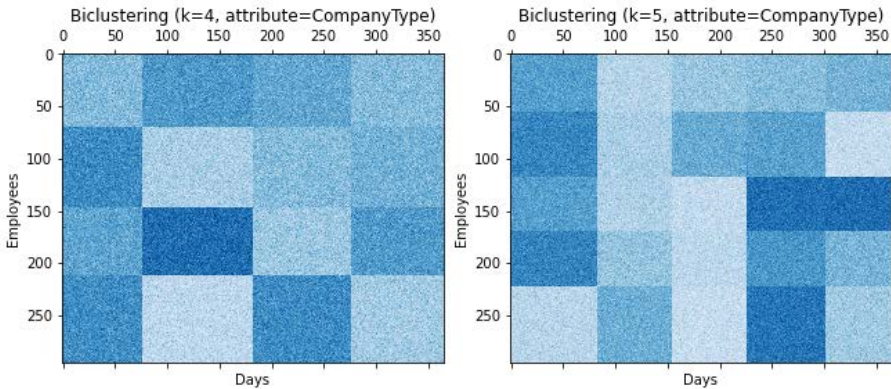


Figure 7: Results of biclustering on preselected employees from 'production' company type.

6 Conclusions

In our initial study, we wanted to understand how employee data can be grouped into different clusters based on absence profiles and demographic information. We used different clustering methods to do this, looking at 2019 data from the MojeUre system. Our constructed dataset includes employee profiles with information about the company and whether they were present or absent each day as well as aggregated information about the absence.

In our study, we employed the k-means clustering algorithm to explore the impact of varying the number of clusters (k) on our dataset. We examined k values ranging from 2 to 5, focusing on an employee absence dataset enriched with detailed absence profiles. To enhance the interpretability of our results, we integrated clustering assignments with principal components and demographic data, such as work hours and company type. We also evaluated clustering quality using the silhouette score, guiding our selection of the optimal number of clusters. Additionally, hierarchical clustering with different linkage methods provided insights into the dataset's inherent structure, while biclustering aims

to find hidden patterns in binary attendance data, all contributing to a comprehensive analysis of our dataset.

Building on our current study, there are several avenues for further research and practical applications. Future work could involve exploring other clustering algorithms or introducing deep learning techniques to automatically recognize more intricate patterns in the data, especially as datasets grow in complexity and volume. Integrating temporal analyses could allow us to better understand seasonal or cyclical patterns in employee attendance.

References

- [1] Leonardi, P., & Contractor, N. (2018). Better people analytics. *Harvard Business Review*, 96(6), 70-81.
- [2] Huselid, M. A. (2018). The science and practice of workforce analytics: Introduction to the HRM special issue. *Human Resource Management*, 57(3), 679-684.
- [3] Sagie, A. (1998). Employee absenteeism, organizational commitment, and job satisfaction: Another look. *Journal of Vocational Behavior*, 52(2), 156-171.
- [4] De Oliveira, J. V., & Pedrycz, W. (Eds.). (2007). *Advances in Fuzzy Clustering and Its Applications*. John Wiley & Sons.
- [5] Javatpoint. (2023). Clustering in Machine Learning - Javatpoint. Retrieved from <https://www.javatpoint.com/clustering-in-machine-learning>
- [6] Kurita, T. (2019). Principal component analysis (PCA). In *Computer Vision: A Reference Guide* (pp. 1-4).
- [7] Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1), 100-108.
- [8] Mǎllner, D. (2011). Modern hierarchical, agglomerative clustering algorithms. *arXiv preprint arXiv:1109.2378*.
- [9] Busygin, S., Prokopyev, O., & Pardalos, P. M. (2008). Biclustering in data mining. *Computers & Operations Research*, 35(9), 2964-2987.
- [10] Peter J. Rousseeuw. *Silhouettes: A graphical aid to the interpretation and validation of cluster analysis*. *Journal of Computational and Applied Mathematics*, 20:53–65, 1987. <https://www.sciencedirect.com/science/article/pii/0377042787901257> [Online; accessed 03.10.2023].
- [11] Ackermann, M. R., Blǎmer, J., Kuntze, D., & Sohler, C. (2014). Analysis of agglomerative clustering. *Algorithmica*, 69, 184-215.

A Multi-Factorial ANOVA Framework for STL File Quality Optimization via Point Cloud Analysis

International Conference on Information Technologies and Information Society (ITIS)

Bojan Pažek, Slavko Arh
Rudlfovo – Science and Technology Center Novo mesto
Podbreznik 15, 8000 Novo mesto, Slovenia
{bojan.pazek, slavko.arh}@rudolfovo.eu

Abstract: *Rapid advances in 3D scanning technologies have forced the development of strong statistical data analysis methodologies, notably in the context of frame selection and Stereolithography (STL - Standard Tessellation Language) file production. As a prelude to experimental validation, this article proposes a theoretical framework based on Multi-Factorial Analysis of Variance (ANOVA). The framework is intended to provide a thorough methodology for evaluating the impact of various independent variables on the quality of generated STL files. Sample Size, which refers to the various sizes of samples collected from the dataset, and Sampling Technique, which includes approaches such as random, stratified, or systematic sampling, are two of the important independent variables we focus on. Many other factors could be considered, including Frame Resolution, Scanning Speed, and Point Cloud Density. By establishing this framework, we lay the groundwork for future empirical studies that will rigorously test the proposed hypotheses and models. The paper acts as a thorough guide for computer science, engineering, and applied mathematics researchers and practitioners, providing insights into the intricate links between frame selection strategies and STL file quality.*

Key Words: *Multi-Factorial Analysis of Variance (ANOVA), Stereolithography (STL), Point Cloud Segmentation, Stratified Sampling, Hypothesis Testing, Quality Metrics in 3D Scanning.*

1 Introduction

Recently developed 3D scanning technologies have shown remarkable promise in accurately capturing the intricate geometries of real-world objects. These technologies employ various active non-contact scanning methods such as time of flight, or structured light, to generate a 3D point cloud, a digital representation of an object's surface coordinates [1]. Technology has found applications in diverse fields, including reverse engineering, where it creates a digital duplicate of reality [2]. The accuracy of these scanning processes may go well beyond 90 micrometers but is usually strongly setting-dependent, encompassing factors like scanning conditions and part surface quality [5]. Despite these advancements, a critical challenge remains - optimizing the selection of frames for the .STL file generation.

Various methodologies have been proposed for 3D scanning and point cloud analysis. For instance, Hegedus-Kuti et al. focused on the identification of welding defects through point cloud alignment [4]. Wang et al. developed a structured light 3D-

scanner to ensure quality in additive manufacturing [5]. There is, however, an opening in the literature for a comprehensive statistical framework that can guide the selection of frames for STL file generation, a process that has a direct impact on the quality and efficiency of 3D models.

The primary objective of this article is to introduce a theoretical framework based on Multi-Factorial Analysis of Variance (ANOVA) for optimizing frame selection in STL file generation. The framework aims to provide a robust methodology for evaluating the impact of various independent variables, such as Sample Size, Sampling Technique, Frame Resolution, Scanning Speed, and Point Cloud Density, on the quality of generated STL files. The novelties of this article include:

- ▲ **Multi-Dimensional Optimization:** In contrast to current methodologies that focus on a limited number of factors (one or two), our proposed framework is designed to incorporate several independent variables, hence providing a comprehensive and integrated approach.
- ▲ **Empirical Foundation:** While the article is theoretical, it lays the groundwork for future empirical studies, thereby bridging the gap between theory and practice.
- ▲ **Interdisciplinary Approach:** The framework is designed to be applicable across various disciplines, including computer science, mechanical engineering, and applied mathematics, making it a versatile tool for researchers and practitioners alike.

By establishing this framework, we aim to fill a critical gap in literature and provide a comprehensive guide for future research in this burgeoning field.

For the empirical foundation, we utilized the Calibry 3D scanner, capable of capturing medium to large objects ranging from 20 cm to 10 m in length. The scanner achieves up to 0.6 mm resolution and has a built-in texture camera, allowing for the collection of up to 3 million points per second. This ensures fast and accurate results with a precision of up to 0.1 mm.

The remainder of this article is organized as follows: Section 2 provides a detailed description of the Multi-Factorial ANOVA framework; Section 3 discusses the independent variables in depth; and finally, Section 4 outlines potential applications and limitations concluding the article with final remarks and future research directions.

2 Detailed Description of the Multi-Factorial ANOVA Framework

Analysis of Variance (ANOVA) is a statistical method used to analyze the differences among group means [6]. In its simplest form, one-way ANOVA provides a statistical test of whether two or more population means are equal. However, in many real-world scenarios, especially in the context of 3D scanning and STL file generation, multiple factors simultaneously influence the outcome [5]. This necessitates the use of Multi-Factorial ANOVA, which allows for the analysis of multiple independent variables at once [1].

The Multi-Factorial ANOVA model serves as a robust statistical framework for analyzing the complex relationships between multiple independent variables and a dependent variable. In the context of this research, the model is particularly tailored to investigate how various factors such as Sample Size, Sampling Technique, and Frame Resolution affect the quality of generated STL files. The mathematical representation of this model is a natural extension of the one-way ANOVA model and is given by:

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijkl}$$

Where:

- Y_{ijkl} is the dependent variable representing the quality of the STL file. It is a function of multiple factors and their interactions. In our study, this mapping is quantified using metrics such as surface smoothness, fidelity to the original object, or even computational efficiency in the STL file generation process.
- μ is the overall mean quality of the STL files across all levels of all factors. It serves as a baseline against which the effects of the individual factors and their interactions are compared.
- α_i quantifies the deviation of the mean quality of the STL file when the i^{th} level of factor A (e.g., Sample Size) is applied, from the overall mean μ .
- β_j is the effect of the j^{th} level of factor B (e.g., Sampling Technique).
- γ_k is the effect of the k^{th} level of factor C (e.g., Frame Resolution).
- $(\alpha\beta)_{ij}$, $(\alpha\gamma)_{ik}$, $(\beta\gamma)_{jk}$ are the interaction terms that capture the combined effects of pairs of factors. Interaction terms are crucial for understanding how the effect of one factor depends on the level of another factor. For example, $(\alpha\beta)_{ij}$ would capture how the effect of Sample Size α_i on STL file quality changes depending on the Sampling Technique β_j used.
- $(\alpha\beta\gamma)_{ijk}$ is the term which captures the three-way interaction between factors A, B, and C.
- ϵ_{ijkl} is the random error term, accounting for the variability in STL file quality that cannot be explained by the factors and their interactions. It is assumed to be normally distributed with a mean of zero and constant variance, $\epsilon_{ijkl} \sim N(0, \sigma^2)$.

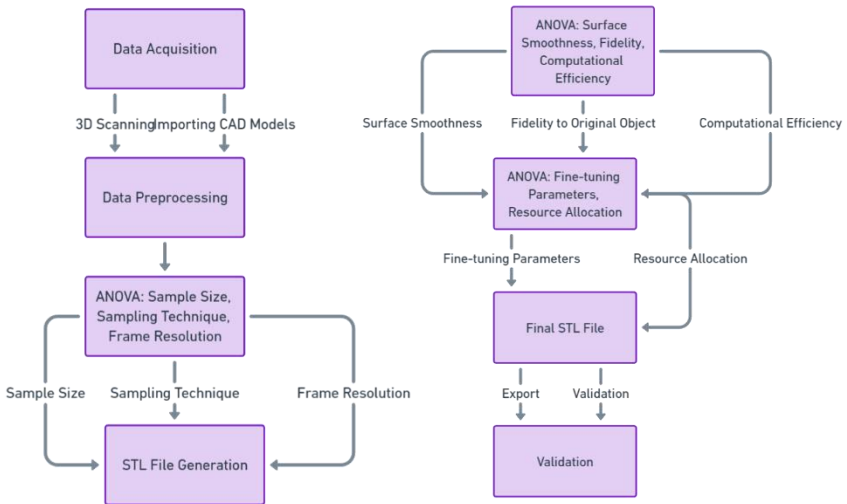


Figure 1 ANOVA framework

This precise formulation establishes a solid mathematical foundation for the Multi-Factorial ANOVA model, which is designed for the complicated and diverse process of STL file production. First, it allows for a thorough analysis by considering many aspects and their interconnections, resulting in a comprehensive understanding of the STL file generating process. Second, the model aids optimization by assisting us in understanding the main and interaction impacts, allowing us to fine-tune each aspect to

produce the greatest quality STL files. Third, the model helps with resource allocation by detecting which parameters have the most influence on STL file quality, guiding decisions in both computational and experimental scenarios. Finally, once validated, the model may be used to forecast STL file quality based on specified levels of the independent variables.

The Multi-Factorial ANOVA model's validity is dependent on several critical assumptions that must be met for the results to be reliable and generalizable. Violations of these assumptions can result in biased or misleading results. Each of these assumptions is discussed in depth below [7]:

1. **Independence of Observations.**
2. **Normality.**
3. **Homogeneity of Variances.**

Each observation is assumed to be independent of the others. This means that in the context of 3D scanning and STL file generation, each frame or point cloud data point must be captured independently of the others. Violation of this assumption can inject bias into the ANOVA results, resulting in inaccurate conclusions. The mathematical way to express the assumption of independence is given by [8]:

$$Cov(Y_{ijkl}, Y_{i'j'k'l'}) = 0$$

The covariance measures how much two variables change together. If the covariance is zero, it implies that the two variables are independent of each other. Furthermore, the dependent variable should be approximately normally distributed for each combination of the levels of the independent variables. This is crucial for the validity of the F-statistic used in ANOVA. For each combination of i, j, k , the distribution of Y_{ijkl} should be normal with mean μ_{ijk} and variance σ^2 , i.e.:

$$Y_{ijkl} \sim N(\mu_{ijk}, \sigma^2).$$

The assumption of homogeneity of variances (also known as homoscedasticity) posits that the variances of the dependent variable should be equal across all levels and combinations of levels of the independent variables. This means that the variance of Y_{ijkl} should be constant for all combinations of i, j, k :

$$Var(Y_{ijkl}) = \sigma^2.$$

Failure to meet these assumptions may necessitate data transformations or the use of non-parametric statistical methods as alternatives to ANOVA.

Next, we are going to discuss hypothesis testing and statistical analysis. The null and alternative hypotheses for each factor and their interactions are formulated as follows:

- **Null Hypothesis (H_0):** The means of the different levels of each factor are equal. Mathematically, for a given factor A , this can be expressed as:

$$H_0: \mu_{A_1} = \mu_{A_2} = \dots = \mu_{A_i}$$

- **Alternative Hypothesis (H_a):** At least one mean of the different levels of each factor is different. Mathematically, this can be expressed as:

$$H_a: \text{At least one } \mu_{A_i} \text{ is different}$$

In addition, the F-statistic is an important component in ANOVA for hypothesis testing. It is calculated as follows for each main effect and interaction effect:

$$F = \frac{\text{Mean Square (Effect)}}{\text{Mean Square (Error)}}$$

The F-statistic is used to test the null hypothesis for each main effect and interaction effect. Once the F-statistic is calculated, the corresponding p-value is obtained to make a decision regarding the null hypothesis. The decision rule is as follows (Here, α is the level of significance, commonly set at 0.05) [9]:

- If $p \leq \alpha$, reject the null hypothesis (H_0).
- If $p > \alpha$, fail to reject the null hypothesis (H_0).

Upon obtaining the F-statistics and their corresponding p-values, we proceed to interpret the results. If the p-value is less than the chosen alpha level, we reject the null hypothesis for that effect. If the null hypothesis is rejected for a particular factor or interaction, post-hoc tests such as Tukey's HSD (Honestly Significant Difference) or Bonferroni can be employed to identify which specific levels of the factor are significantly different from each other [10].

A post-hoc test's output typically contains corrected p-values for each pairwise comparison. Similarly, to how the initial ANOVA was interpreted, if the adjusted p-value is less than, the difference between those specific groups is judged statistically significant. We can identify which specific levels of the parameters (e.g., different Sample Sizes or Sampling Techniques) have a substantial impact on the quality of the STL file by performing these post-hoc tests. This knowledge is priceless for practitioners looking to improve 3D scanning methods. This in-depth interpretation and post-hoc analysis allow us to reach nuanced conclusions regarding the factors influencing STL file quality, delivering useful insights for both researchers and practitioners.

3 In-Depth Discussion of Independent Variables

In this section, we will delve into the specific independent variables that are considered in our Multi-Factorial ANOVA framework. We intend to provide a comprehensive understanding of each variable, its possible impact on the quality of STL files, and its significance in the context of 3D scanning and point cloud analysis.

3.1 Sample Size

Sample Size refers to the number of frames selected from the original dataset of 1323 frames for the purpose of generating an STL file. The size of the sample has a direct impact on the computational complexity and the quality of the resulting STL file. A larger sample size generally provides a more accurate representation of the object being scanned but at the cost of increased computational time and resources.

In the context of our Multi-Factorial ANOVA model, the sample size n is a crucial independent variable that we aim to optimize. It is selected from a dataset containing $N = 1323$ frames. The sample size is not just a mere count but serves as a ratio or fraction of the total dataset, which we denote as f :

$$f = \frac{n}{N}$$

Here, f is the Sample Fraction, and it serves as a normalized representation of the sample size. This normalization is essential for two reasons:

1. **Comparability:** By converting the sample size into a fraction of the total, we can more easily compare the effects of different sample sizes across various datasets, even if the total number of frames N varies.
2. **Dimensionless Quantity:** Making f dimensionless aids in the mathematical modeling and interpretation of results, as it removes the units associated with the count, making it a pure number that can be universally understood.

In the Multi-Factorial ANOVA model, f can be included as a continuous variable or discretized into categories (e.g., 'Small', 'Medium', 'Large') depending on the research question. Discretizing the sample size into categories like 'Small', 'Medium', and 'Large' can offer several advantages, particularly when dealing with complex systems where computational resources are a constraint. This approach simplifies the analysis and makes the results more interpretable. By rigorously defining and incorporating the sample size

in this manner, we aim to provide a robust mathematical framework for its analysis, thereby enabling more precise and actionable insights into its effects on STL file quality.

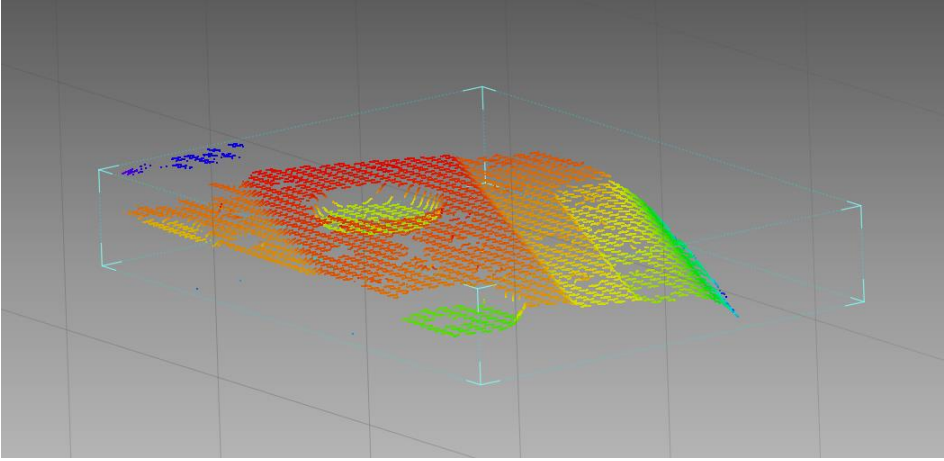


Figure 2 Frame_1.ply (Point Shader by Height) in 1323 dataset

3.2 Sampling Technique

Sampling Technique refers to the method employed to select frames from the original dataset of 1323 frames. The choice of sampling technique can significantly influence the representativeness of the sample and, consequently, the quality of the resulting STL file. Common sampling techniques include random sampling, stratified sampling, and systematic sampling.

Let S be the set of all frames, and \tilde{S} be the subset selected through a given sampling technique. The sampling technique can be mathematically represented as a function:

$$f: S \rightarrow \tilde{S} \subseteq S$$

This function maps the total set of frames S to the selected subset \tilde{S} , based on the rules defined by the sampling technique.

The choice of sampling technique can affect several quality metrics of the STL file like representativeness, complexity, and bias. For example, some techniques, like systematic sampling, are computationally less intensive but might miss out on capturing specific features. It is noteworthy that random sampling minimizes the risk of bias, leading to a more accurate STL file, but at the same time could potentially miss those overlapping frames that are of our interest. Furthermore, techniques like stratified sampling ensure that all characteristics of the object are adequately represented.

An important concept to consider is that of 'Optimal Sampling,' which aims to find the best trade-off between computational efficiency and STL file quality. This can be quantitatively assessed using a cost function $C(\tilde{S})$ that considers various quality metrics and computational time. Mathematically, the optimal sampling technique f^* is defined as:

$$f^*(S) = \underset{f \in F}{\operatorname{arg\,min}} C(f(S)).$$

Here, $\operatorname{arg\,min} C(f(S))$ stands for 'argument of the minimum,' identifying the function f in the set F of all possible sampling techniques that minimizes the cost function $C(f(S))$. Unlike min, which gives you the smallest output value itself, arg min tells you what input

produced that smallest output. This concept adds another layer of depth to our study and will be a focal point in our Multi-Factorial ANOVA model.

The sampling technique chosen should be in line with the study objectives. For example, if the goal is to capture specific properties of the object, stratified sampling may be preferable. If the goal is to decrease processing time, systematic sampling may be a preferable option. The efficiency of a sampling approach is frequently dependent on other factors such as Sample Size and Frame Resolution. A bigger Sample Size, for example, may overcome the drawbacks of a less representative sampling technique. These interactions are important and will be taken into consideration in the Multi-Factorial ANOVA model.

3.3 Additional Independent Variables for Comprehensive Analysis

While the primary focus of this study is on Sample Size and Sampling Technique, it is crucial to acknowledge other independent variables that could be integrated into future research or more intricate models. These variables have the potential to interact with the primary variables, thereby influencing the quality of the generated STL files.

Firstly, Frame Resolution is a significant factor; it pertains to the number of points in each frame and affects the granularity of the object's representation. Higher resolution frames offer more detail but demand greater computational resources. Secondly, Scanning Speed is another variable of interest. It refers to the rate at which the 3D scanner captures frames. While faster scanning speeds can expedite the scanning process, they may compromise the quality of the captured frames.

Point Cloud Density is also noteworthy. It signifies the number of points per unit volume in the 3D space of the object. A higher density can enhance the accuracy and detail of the STL file, albeit at the cost of increased computational complexity.

Given that our work involves optical scanning, Light Source Intensity is a variable that cannot be overlooked. The intensity can significantly affect the quality of the captured frames, and optimal intensities may vary depending on the material and color of the object being scanned. Material Properties, such as reflectivity and texture, can also influence the quality of the point cloud and, consequently, the STL file. Environmental Factors like ambient light, temperature, and air quality can impact the scanning process and should be either controlled or accounted for in the experimental design.

Lastly, Computational Hardware, including CPU speed and RAM size, can affect the time required for STL file generation and may also influence the feasible range for other variables like Sample Size. These additional variables offer promising avenues for future research. They can be integrated into more complex models to provide a comprehensive understanding of their impact and interactions with the primary variables of Sample Size and Sampling Technique.

4 Applications, Limitations, and Conclusion

The goal of this final section is to examine the advantages and drawbacks of the Multi-Factorial ANOVA approach utilized in this study. It will also include some closing remarks and suggestions for further research. Our goals are to make clear how the framework may be used in the real world, to highlight its limitations, and to point the way toward future research that can expand on these findings.

4.1 Potential Applications and Limitations

The approach described in this article has potential for use in many contexts. The field of reverse engineering benefits greatly from its use. The framework may then be used to fine-tune the scanning procedure, guaranteeing the highest quality STL files possible. This is of critical importance in sectors like aerospace and automobile design, where

accuracy and attention to detail are essential. Usefulness in manufacturing quality control comes in close second. The framework can be used as a reliable instrument for evaluating the quality of components and finished goods by means of non-destructive testing. Manufacturers can get highly accurate point clouds by tweaking the scanning parameters, which can subsequently be utilized for comparison against preset quality criteria.

The preservation of cultural artifacts is another interesting use example. The framework can be used to create digital representations of historic buildings and other cultural treasures, ensuring their survival for future generations. By optimizing the scanning procedure, even the smallest of details may be recorded, which is very useful for studying the past and teaching future generations. Last but not least, the system has practical applications in robotics and automation. Here, it can be put to use improving the scanning procedure for jobs requiring recognition and manipulation of physical objects. Optimizing the point cloud, for example, can improve the efficiency of object picking and placement in automated warehousing solutions, which in turn can improve the efficiency of the entire supply chain.

However, the constraints of this framework must be taken into account. The computational complexity of maximizing a large number of independent variables at once is an important barrier. This can be especially difficult in real-time contexts where prompt action is required. Overfitting is another issue that might arise when training a model on a small dataset. It is possible that this would result in a model that does very well on the training data but poorly on novel, unseen data. The paradigm also makes the assumption that all independent variables are of equal significance, which may not always be the case. Finally, environmental elements such as ambient lighting, temperature, and air quality can influence the scanning process and should be controlled or accounted for in the experimental design, but these are not currently taken into consideration by the framework.

4.2 Potential Applications and Limitations

The Multi-Factorial ANOVA framework described here provides a powerful way for enhancing the integrity of STL files made from 3D scans. The framework gives a detailed knowledge of the elements that affect STL file quality by taking into account several independent variables. There are, as with any theoretical framework, many potential directions for future study and expansion. Adding new variables to the mix is one way to broaden the scope of the framework. The substance being scanned, environmental considerations, and hardware requirements could all be investigated in further research. Incorporating machine learning methods to estimate STL file quality based on the independent variables is another potential direction. It is possible that this dynamic method, with its ability to make adjustments in real time, may enhance the scanning process's efficiency and accuracy.

Empirical validation is another critical aspect that needs attention. The goal of future research should be to verify the theoretical framework in practice. To do this, we would need to construct experiments to test the framework in a variety of settings, and then analyze the data to either validate or modify our approach. The creation of cutting-edge optimization algorithms may also prove useful. Finding the best values for the independent variables could be expedited using methods like genetic algorithms or gradient descent.

Our hope is that the framework we have established in this article will inspire others to continue exploring this fascinating area at the intersection of computer science, engineering, and applied mathematics. The ultimate purpose of the framework is to enhance the quality and dependability of 3D scanning technologies and their many

applications, and the structure it provides might be useful for both researchers and practitioners.

5 References

- [1] Oh, S.; Suh, D. Mannequin fabrication methodology using 3D-scanning, modeling and printing. *Int. J. Cloth. Sci. Technol.* 2021, 33, 683–695.
- [2] Alcácer, V.; Cruz-Machado, V. Scanning the industry 4.0: A literature review on technologies for manufacturing systems. *Eng. Sci. Technol. Int. J.* 2019, 22, 899–919.
- [3] Collins, P.C.; Martill, D.M.; Smyth, R.S.; Byrne, R.; Simms, M.J. Chaperoning digital dinosaurs. *Proc. Geol. Assoc.* 2021, 132, 780–783.
- [4] Hegedus-Kuti, J.; Szolosi, J.; Varga, D.; Abonyi, J.; Ando M.; Ruppert, T. 3D Scanner-Based Identification of Welding Defects - Clustering the Results of Point Cloud Alignment. *Sensors* 2023, 23, 2503. <https://doi.org/10.3390/s23052503>.
- [5] Wang, R.; Law, A.C.; Garcia, D.; Yang, S.; Kong, Z. Development of structured light 3D-scanner with high spatial resolution and its applications for additive manufacturing quality assurance. *Int. J. Adv. Manuf. Technol.* 2021, 117, 845–862.
- [6] Box, G. E. P. (1954). "Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variance Problems, I. Effect of Inequality of Variance in the One-Way Classification". *The Annals of Mathematical Statistics.* 25 (2): 290. doi:10.1214/aoms/1177728786.
- [7] Freedman, David A.; Pisani, Robert; Purves, Roger (2007) *Statistics*, 4th edition. W.W. Norton & Company ISBN 978-0-393-92972-0.
- [8] Douglas C. Montgomery *Design and analysis of experiments*. Eighth edition. Publisher: Montgomery, Douglas C. ISBN: 978-1-118-14692-7.
- [9] Cox, D.R.; Reid, N. *The Theory of the Design of Experiments*. First edition. Publisher: Chapman and Hall/CRC. ISBN-13: 978-1584881954.
- [10] Dunn, O. J. (1961). Multiple Comparisons Among Means. *Journal of the American Statistical Association*, 56(293), 52–64.

Knowledge base assisting PCB Design tool selection and combination in Online CADCOM platform

Lavdim Menxhiqi, Galia Marinova

UBT, Pristina Kosovo Technical University of Sofia, Bulgaria
lavdim.menxhiqi@ubt-uni.net, gim@tu-sofia.bg

Abstract: *Recently a high number of printed circuit board (PCB) design tools with different design capacity and dynamic status (free, updated, accessible, etc.) are available on the net. The study in the paper aims to recommend to the designers the best tool for a concrete PCB design. It presents the development and implementation of a new feature in the Online-CADCOM platform which already offers filter design tool selection. A Knowledge base supporting PCB design tool selection is added. A set of criteria with different weigh coefficients is defined and a multi-criteria decision model is integrated in the Online-CADCOM expert tool for automatic selection of the PCB design tools. The Decision Matrix methodology is used in the selection process and MAUT method is used for option ranking. A set of 13 PCB free design tools are analysed and evaluated. These tools are tested for the PCB design of a predefined generator circuit for sinusoidal signal with quartz stabilization. The set of these tools completing the task is considered in terms of the set of predefined criteria in the Knowledge base. Through the multi-criteria decision methodology implemented in Online-CADCOM and based on the Decision Matrix technique, the PCB design tools are evaluated for solving a concrete task defined by the users and allowing them to choose the most suitable tool for the job based on their needs and preferences. Illustrative examples of the decision making are presented.*

Key Words: *knowledge base, tool selection, computer-aided design tools, PCB design tools, Online-CADCOM platform, multicriteria decision analysis, decision matrix and MAUT model, criteria classification*

1 Introduction

Designers face numerous challenges when selecting an optimal printed circuit board (PCB) design tool from countless resources available online. The objective of the study described in the paper is to simplify this process through introducing a new feature within the Online-CADCOM platform – a Knowledge Base designed specifically for supporting informed decision-making processes [1].

A multi-criteria model integration that consists of 3 categories determined via weights calculation enabling assessment of distinct freely accessible PCB design tools is implemented. By combining Decision Matrix approach with Multi-Attribute Utility Theory (MAUT), are evaluated thirteen software platforms and created characterization passports for 8 easily accessible alternatives based on our analysis [1] [2].

The process begins with a testing procedure, replicating the creation of a benchmark circuit. It then proceeded to illustrate PCB design tool selection, enhancing the comprehensiveness and efficiency of the selection [3]. Illustrative examples, ultimately making informed decisions when choosing a PCB design tool are proposed.

2 Analysis of PCB Design Tools and characterization passports

This section sets the stage by initially examining thirteen freely accessible Printed Circuit Board (PCB) design tools available on the internet [6][7][8]. The tools considered in this study were: DesignSpark PCB [9], EasyEDA [10], CircuitMaker [11], EAGLE PCB [12], KiCad [13], LibrePCB [14], ExpressSCH & ExpressPCB [15], TinyCAD [16], Fritzing [17], Osmond PCB [18], gEDA PCB [19], PCBWeb Designer [20] and ZenitPCB [21]. The selection process considered popularity and accessibility as primary factors when choosing PCB design tools for testing. Findings revealed that not all tools met our criteria as only 8 out of 13 tested were operable on Windows platforms and freely accessible. These tools were DesignSpark PCB [9], EasyEDA [10], CircuitMaker [11], EAGLE PCB [12], KiCad [13], LibrePCB [14], ExpressSCH & ExpressPCB [15], TinyCAD [16].

The reasons for dropping the initial 5 PCB Tools are: Fritzing Compatibility isn't freely accessible, Osmond PCB and gEDA PCB don't provide Windows, PCBWeb Designer is a Converter and calculator, there is not a link available for ZenitPCB.

Following the initial selection, a more detailed analysis was conducted to establish characterization passports for each of the 8 tools [3]. Technical expertise on PCB design was necessary alongside a systematic approach balancing practicality so that all crucial aspects of every tool were extensively covered before moving on to the next one.

3 Criteria for PCB Design Tool Selection

The following stage in the study involved the creation of a criteria selection mechanism for the PCB design tools. These selection criteria allow to evaluate and compare crucial functionalities of these design tools in great detail and they were chosen keeping in mind their suitability for diverse users and coherence with our passport findings. These criteria were classified into 3 overarching groups, Mandatory; Highly Desirable; and Desired, each carrying its own significance with respective weighting coefficients of 1.00, 0.50, and 0.33 respectively. Tables 1, 2, and 3 lay out these criteria and their corresponding categories. Table 1 features the list of essential criteria and options pertinent to PCB Design tools. Table 2 illustrates the desirable criteria assigned with a weight coefficient of 0.5, and Table 3 displays another set of desirable criteria, this time assigned with a weight coefficient of 0.33 [5].

This research uses the Multi-Criteria Decision Analysis (MCDA) approach as detailed in paper [2]. Two steps algorithm is implemented. First, mandatory criteria are defined. Then, viable options (VO) that satisfy individual mandatory criteria are set on using the formula in [2]. Desirable criteria are defined and weight coefficients w_j (1/2, 1/3, 1/6) are determined. Viable options O_k are ranked conform to the utility function $F(O_k)$ [2].

Table 1 presents the mandatory criteria and options for PCB design tools, while Tables 2 and 3 showcase the desirable criteria with their respective weight coefficients.

4 Assessment of PCB Design Tools

As part of our assessment of 8 freely available PCB design tools, the analysis is centered on their capacity in the design process of a predefined quartz-stabilized generator circuit. Out of the initial selection, 5 tools (DesignSpark PCB, EasyEDA, CircuitMaker, EAGLE and KiCad) were able to complete the task satisfactorily. These tools demonstrated the capabilities of circuit capture, automatic transfer to PCB design, and 3D view generation, providing a comprehensive solution for the given task.

Table 1: Essential criteria and options

Criteria	Options	Tools providing the options	Weight Coefficient
Board Size	Up to 80cm ² (10cm x 8cm or 12.4 cm x 6.4 cm)	All 5 tools	1.0
	Up to 1m x 1m	DesignSpark PCB, EasyEDA, KiCad, CircuitMaker	
	4m x 4m	KiCad, CircuitMaker	
	Unspecified from company	CircuitMaker	
Number of layers	Up to 16	All 5 tools	1.0
Number of footprints in the libraries	Up to 8000	All 5 tools	1.0
	Up to 250,000	Design Spark PCB, Circuit Maker, EAGLE PCB, Easy EDA	
	Up to 300,000	Circuit Maker, EAGLE PCB, Easy EDA	
	Up to 750,000	EAGLE PCB, Easy EDA	
	Up to 1000000	Easy EDA	

Table 2: Desired criteria and options, with weight coefficient 0.5

Criteria	Options	Tools providing the options	Weight coefficient
Design Verification	Design Verification	All 5 tools	0.5
Auto-Routing	Auto-Routing	EAGLE, DesignSpark PCB, EasyEDA, CircuitMaker	0.5
Import/Export	Import DXF	All 5 tools	0.5
	Import STEP	EAGLE, KiCad	
	Import Design Rules	All 5 tools	
	Import Custom Libraries	All 5 tools	
	Export Gerber Files	All 5 tools	
	Export Drill Files	All 5 tools	
	Export BOM	All 5 tools	
Export Netlist Files	All 5 tools		
Simulation	SPICE simulation	EAGLE, EasyEDA, KiCad, CircuitMaker	0.5

Table 3 Desired criteria and options for PCB Design Tools, with weight coefficient 0.33

Criteria	Options	Tools providing the options	Criteria
3D View	3D View	All 5 tools	0.33
Technical Support	Direct Email Support	EAGLE, CircuitMaker	0.33
	Contact Form on Website	EAGLE, EasyEDA	
	User Forums	All 5 tools	
	Online Documentation	All 5 tools	
	Tutorials	All 5 tools	
Cloud Integration	Cloud Integration	EasyEDA, CircuitMaker	0.33
Cross-Platform Support	Windows	EAGLE, EasyEDA, KiCad, CircuitMaker	0.33

This set of 5 tools became the subject of further analysis. The primary goal of this study is to first identify tools that can effectively complete a given task. Subsequently, the study aims to rank these capable tools based on a set of predefined criteria, distinguishing the best tools among them.

Hence, these 5 tools were assessed against the predefined set of criteria in the Knowledge Base. This criterion set has been carefully curated to cover essential capabilities and desired features of the PCB design tool, ensuring a comprehensive evaluation process.

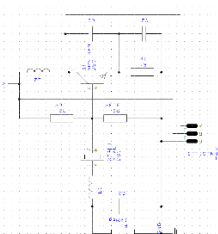
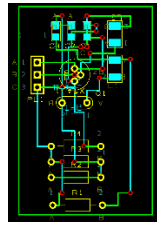

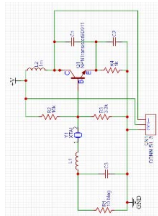
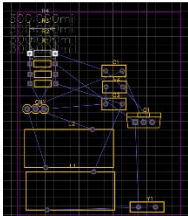
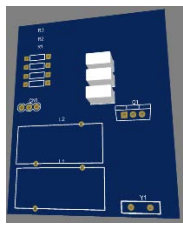
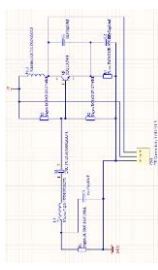
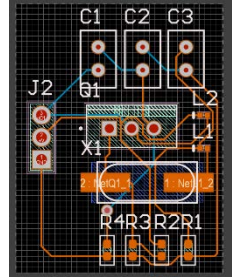
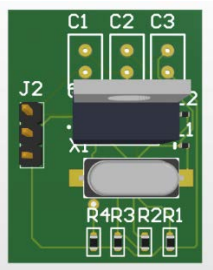
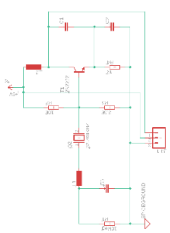
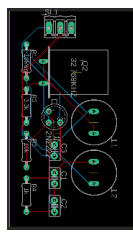

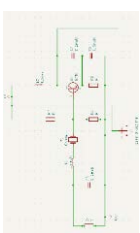
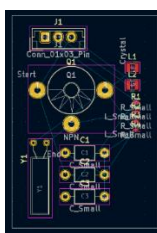
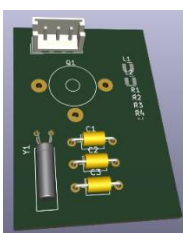
In the subsequent sections, is discussed how each of these 5 tools performed against selection process using defined criteria. This analysis allows to offer informed recommendations to designers, ensuring the most suitable tool is chosen for their specific PCB design needs.

The reasons for dropping 3 out of 8 PCB Design Tools after detailed analysis are: LibrePCB doesn't offer Auto Transfer to PCB Design and Due to errors from circuit the PCB Design not created. ExpressSCH & ExpressPCB doesn't offer Auto Transfer to PCB Design and 3D View and it doesn't provide all the components in manual PCB Design, TinyCAD doesn't offer Circuit Capture and there are Difficulties in finding/not containing the components.

5 Illustration of PCB Design Tool selection in ONLINE-CADCOM

In the illustration is highlighted the process of selecting appropriate tools for the creation of a PCB design. Tools capable of facilitating the creation of a PCB design that supported a board size of up to 1m x 1m, could accommodate up to 32 layers and contains up to 250000 footprints in the libraries. Furthermore, the presence of auto-routing functionality is highly sought after, as it could significantly streamline the design process, allowing for more efficient handling of the large board and multiple layers. Additionally, other desirable features were simulation capabilities, 3D visualization, direct email support, and cloud integration. The simulation feature could potentially allow for preemptive troubleshooting of the design, while 3D visualization would provide a comprehensive perspective of the end product. Direct email support would ensure we had prompt technical assistance when needed, and cloud integration would facilitate easier collaboration and data management [5].

Table 4 PCB Design Tools (Circuit, PCB Design and 3D View)

Tool	Circuit	PCB Design	3D View
<p>Design Spark PCB</p>			
<p>EasyEDA</p>			
<p>Circuit Maker</p>			
<p>EAGLE</p>			
<p>KiCad 7.0</p>			

In Table 5 the mandatory criteria (Board Size: Up to 1m x 1m, Number of layers: Up to 32 and Number of footprints in the libraries: Up to 250000) from Table 1 are applied for the 5 tools for PCB design. In Table 6 the desired criteria from are applied for the 2 tools that fulfill the mandatory criteria [5].

Table 5 Implementation of mandatory criteria

<i>Mandatory criteria</i>	<i>Board Size: Up to 1m x 1m</i>	<i>Number of layers: Up to 32</i>	<i>Number of footprints in the libraries: Up to 250000</i>	<i>VO</i>
DesignSpark PCB	1	0	1	0
EasyEDA	1	1	1	1
CircuitMaker	1	1	1	1
EAGLE	0	1	1	0
KiCad	1	1	0	0

The Viable options (VO) for this case are calculated with formula (Eq. 2). The VOs are the 2 following tools: EasyEDA and CircuitMaker [5].

Table 6 Implementation of desirable criteria

<i>Tool for PCB design</i>	<i>EasyEDA</i>	<i>CircuitMaker</i>	<i>Weight coefficient</i>
Auto-Routing	1	1	0.5
Simulation	1	1	0.5
3D View	1	1	0.33
Direct Email Support	0	1	0.33
Cloud Integration	1	1	0.33
<i>F(Ok)</i>	1.66	1.99	

F(Ok) are calculated with formula (Eq. 1). The rank of the viable options is [5]:

1. The tool most well-suited for addressing the task is: CircuitMaker (F(Ok) = 1.99).
2. As a second choice is EasyEDA (F(Ok) = 1.66).

6 Conclusion

This paper presented a systematic approach utilizing a feature in the Online-CADCOM platform. 13 PCB design tools were examined, detailed the characteristics of 8 in passports, and evaluated the capabilities of 5 of them in designing a specific circuit. A set of weighted selection criteria was formulated, divided into 3 categories to reflect their relative importance.

These criteria were integral to multi-criteria decision model, which relied on Decision Matrix methodology to filter options and rank them using MAUT.

The practical application of this methodology led to the selection of the 2 most viable PCB design tools from the original 5, namely EasyEDA and CircuitMaker, which were ranked based on their utility scores. CircuitMaker emerged as the most suitable tool for our specific design task, closely followed by EasyEDA.

The study confirms the utility of the approach in choosing the right PCB design tool tailored to a designer's specific needs and preferences. However, as the state of available tools is dynamic, this process would need periodic updates to the knowledge base, ensuring that the recommendations stay current and relevant.

References

- [1] G. Marinova, V. Guliashki, O. Chikov, "Concept of Online Assisted Platform for Technologies and Management in Communications – OPTIMEK", Proc. of ICCSIST 2014, 7-9 November 2014, Durres, Albania, 2014, pp.55-62
- [2] B. Rodic, G. Marinova, O. Chikov, "Algorithms and Decision Making Methods for Filter Design Tool Selection for a Given Specification in Online-CADCOM Platform", Proc. of ERK'2017, Portoroz, Slovenia, 25-26.09.2017, pp. 247-25
- [3] G. Marinova, O. Chikov, "Methodology for tools integration in the Online assisted Platform for Computer-aided design in communications", Proc. ICEST'2015, 24-26 June 2015, Sofia, pp.31-36,
- [4] G. Marinova, O. Chikov, "E-content Development and Task Solution Using the Content Management System of Online-CADCOM", Proc. Of ICEST 2016, Ohrid, Macedonia, June 28-30, 2016, pp.213-216
- [5] G. Marinova, O. Chikov, B. Rodic, "E-Content and Tool Selection in the Cloud-based Online-CADCOM Platform for Computer-Aided Design in Communications", Proc. of CONTEL'2019, Graz, Austria, 3-5 July 2019, IEEE 2019, pp.1-4,
- [6] Matt Krysiak, "Top PCB Design Software Tools for Electronics Engineers: 46 Must-Have Tools to Streamline PCB Design" www.pannam.com.
<https://www.pannam.com/blog/best-pcb-design-software-tools/> (accessed April 13, 2023).
- [7] Abby Hao, "10 Best PCB Design Software Tools In 2021" [linkedin.com](https://www.linkedin.com/pulse/10-best-pcb-design-software-tools-2021-abby-hao/?trk=pulse-article_more-articles_related-content-card).
https://www.linkedin.com/pulse/10-best-pcb-design-software-tools-2021-abby-hao/?trk=pulse-article_more-articles_related-content-card (accessed April 14, 2023).
- [8] G2, "10 Best PCB Design Software Tools In 2021" www.g2.com.
<https://www.g2.com/categories/pcb-design> (accessed April 14, 2023).
- [9] "DesignSpark PCB - Version 11.0 Release Notes" www.rs-online.com.
<https://www.rs-online.com/designspark/designspark-pcb-version-11-0-release-notes> (accessed Jun. 5, 2023).
- [10] EasyEDA, " EasyEDA Std Edition Client Download " easyeda.com.
<https://easyeda.com/page/download> (accessed Jun 5, 2023).
- [11] Altium, " Download CircuitMaker " www.altium.com.
<https://www.altium.com/circuitmaker/download/b> (accessed March 7, 2023).
- [12] Autodesk, " EAGLE / Fusion 360 " www.autodesk.in.
<https://www.autodesk.in/products/eagle/> (accessed March 8, 2023).
- [13] Kicad, " Download Kicad " www.kicad.org. <https://www.kicad.org/download/> (accessed March 09, 2023).
- [14] LibrePCB, " Download LibrePCB " www.librepcb.org.
<https://librepcb.org/download/> (accessed March 15, 2023).

- [15] ExpressPCB, "Download ExpressPCB," [expresspcb.com](https://www.expresspcb.com/).
<https://www.expresspcb.com/> (accessed March 18, 2023).
- [16] TinyCAD, "Download TinyCAD" [tinycad.net](https://www.tinycad.net/). <https://www.tinycad.net/Download>
(accessed March 25, 2023).
- [17] Fritzing, "Download Fritzing" fritzing.org. <https://fritzing.org/download/> (accessed
March 25, 2023).
- [18] Osmond, "Download Osmond" [osmondpcb.com](https://www.osmondpcb.com/).
<https://www.osmondpcb.com/download.html> (accessed March 26, 2023).
- [19] PCB - gEDA project, "Download gEDA PCB" [geda-project.org](http://pcb.geda-project.org). [http://pcb.geda-](http://pcb.geda-project.org)
[project.org](http://pcb.geda-project.org) (accessed March 27, 2023).
- [20] PCBWeb Designer, "PCBWeb Designer" [pcbweb.com](https://www.pcbweb.com/). <https://www.pcbweb.com/>
(accessed March 28, 2023).
- [21] Zenit PCB, "Zenit PCB" [zenitpcb.com](https://www.zenitpcb.com/). <https://www.zenitpcb.com/> (accessed March
29, 2023).