Proceedings of the 3rd Conference on Intelligent Systems and Information Technologies

Artificial Intelligence Impact
on Physical Sciences and Medicine

Editors:
Anna Wawrzynczak
Piotr Switalski

Proceedings of the 3rd Conference on Intelligent Systems and Information Technologies

Artificial Intelligence Impact on Physical Sciences and Medicine

Cover design: Piotr Świtalski

ISBN: 978-83-68754-07-0
e-ISBN: 978-83-68754-08-7

Wydawnictwo Naukowe
Uniwersytetu w Siedlcach
08-110 Siedlce, ul. Żytnia 17/19, tel. 25 643 15 20
wydawnictwo-naukowe.uws.edu.pl

Typesetting: Piotr Świtalski

Druk i oprawa: Volumina.pl; arkuszy wydawniczych: 10,46; arkuszy drukarskich: 5,03

Preface

It is with great honour and scientific satisfaction that we present the Proceedings of the 3rd International Scientific Conference on Intelligent Systems and Information Technologies (ISIT 2025). As a well-established annual event organised by the Institute of Computer Science, the third edition of ISIT has further consolidated its position as a recognised forum for the exchange of advanced research results in the rapidly evolving field of intelligent systems. The conference was held in a hybrid format from 24 to 26 September 2025 and was hosted by the Institute of Computer Science at the University of Siedlce, Poland.

The central theme of this year's conference, *"Artificial Intelligence Impact on Physical Sciences and Medicine,"* positioned ISIT 2025 at the forefront of contemporary scientific discourse, highlighting the growing role of artificial intelligence as a fundamental enabler of innovation across theoretical and applied research. The conference provided an international platform for presenting and discussing state-of-the-art methodologies, algorithms, and computational frameworks that advance precision, scalability, and interpretability in complex physical systems, energy-related applications, and medical diagnostics and decision support.

ISIT 2025 brought together a diverse and distinguished group of researchers, practitioners, and contributing authors from numerous countries, including India, Spain, the Czech Republic, Slovakia, and Morocco. This strong international participation reflects the global relevance of the conference topics. It underscores ISIT's role in fostering interdisciplinary and cross-border collaboration in artificial intelligence, data-driven modelling, and intelligent information technologies.

A prominent component of the scientific programme was a series of four keynote lectures delivered by internationally recognised experts, addressing both foundational and applied aspects of artificial intelligence:

- Prof. dr hab. Zbigniew Józef Michalewicz (University of Adelaide, Australia; Chief Scientific Officer, Complexica): *"AI-based Business Applications for Transforming Data into Decisions"*
- Prof. dr hab. inż. Ryszard Tadeusiewicz (AGH University of Science and Technology, Poland): *"Areas and Forms of Artificial Intelligence Applications in Medicine"*
- Prof. dr hab. inż. Włodzisław Duch (Nicolaus Copernicus University, Poland): *"Physics-Informed Artificial Intelligence"*
- Prof. dr hab. Marcin Paprzycki (Systems Research Institute, Polish Academy of Sciences, Poland): *"Extended Green Cloud Simulator – Leveraging Multi-Agent Systems for Green Cloud Simulations"*

These lectures provided valuable insights into current research trends, emphasising the integration of domain knowledge with machine learning, the development of trustworthy and explainable AI models, and the application of intelligent systems to real-world, large-scale problems.

To complement the theoretical contributions, ISIT 2025 also offered a full-day, hands-on tutorial entitled *"Leveraging Large Language Models for Private Data."* Organised within the framework of the EuroCC project and conducted by Dr. Krzysztof Nawrocki and Dr. Rafał Możdżonek from the National Centre for Nuclear Research (NCBJ), the tutorial focused on practical aspects of deploying large language models in secure and privacy-preserving environments, making use of the Athena Cluster infrastructure.

The peer-reviewed papers collected in these Proceedings present original research results and novel applications reflecting the broad thematic scope of the conference. The contributions address a wide range of topics, including artificial intelligence methods in the physical and energy sciences, the theoretical foundations and security aspects of AI, medical and healthcare applications, intelligent solutions for people with special needs, and societal and social science perspectives on AI. Together, these works constitute a comprehensive record of the scientific achievements and discussions at ISIT 2025.

We express our sincere gratitude to the members of the Scientific Committee for ensuring high academic standards, as well as to the reviewers whose thorough and constructive evaluations were essential in maintaining the quality and scientific integrity of the accepted papers.

We also gratefully acknowledge the support of our sponsors, Evobot 2.0 and Red Ocean sp. z o. o., whose contributions were instrumental in the successful organisation of the conference.

This volume will serve as a valuable reference for researchers and practitioners, stimulate further scientific inquiry, and contribute to the continued development of intelligent systems and information technologies at the intersection of physical sciences and medicine.

Dr Anna Wawrzyńczak-Szaban
Chairwoman of the Scientific Committee and Organising Committee
3rd Conference on Intelligent Systems and Information Technologies

Conference video recordings

The video recordings from the 3rd Conference on Intelligent Systems and Information Technologies (ISIT 2025) are now available on the official YouTube channel. All interested readers are invited to access the recordings via the following link:
`https://www.youtube.com/@CISIT-InformatykaSiedlce`.

Contents

Part I

Invited lectures

Prof. Dr Habil. Eng. Ryszard Tadeusiewicz
ORCID: 0000-0001-9675-5819

AGH University, Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie
Al. Mickiewicza 30, 30-059 Kraków, Poland

rtad@agh.edu.pl

Areas and Forms of Artificial Intelligence Applications in Medicine

Abstract. Currently, I try to apply artificial intelligence everywhere and to everything, so it is not surprising that many researchers and practitioners are also attempting to use it in medical applications. When discussing the areas and forms of these applications, it is worth mentioning that applications in medicine were among the first during a period when artificial intelligence was not as widely known or – let's not be afraid of the word – trendy as it is today. One of the first expert systems that was created and used was called MYCIN. It was developed at Stanford University in the 1970s and was used to support diagnostics and antibiotic therapy for blood infections. However, in the keynote lecture inaugurating the proceedings of the Intelligent Information Systems conference with the subtitle Artificial Intelligence Impact on Physical Sciences and Medicine, the focus will not be on the history of artificial intelligence in medicine, but on the areas and forms of its current applications. The first area of these applications is undoubtedly supporting the work of devices that acquire patient data. The equipment currently used collects information about the morphological and functional changes that the disease has caused in the patient's body, but the high quality of this information is ensured not only by the excellence of the information acquisition systems but also by the possibility of intelligent processing of this information before it is presented to the doctor. The role of artificial intelligence in this area is already very high and continues to grow, because intelligently collected and presented symptoms are a condition for the effective work of doctors. The second area in which artificial intelligence contributes to successes in medicine is supporting the diagnostic process and therapy planning. The role of artificial intelligence here is very significant because, formally, the doctor making the appropriate decisions must take into account the entirety of medical knowledge, not just that which he or she acquired during studies and from their own practice. The second area in which artificial intelligence contributes to successes in medicine is supporting the diagnostic process and therapy planning. The role of artificial intelligence here is very important because, formally, a doctor making the relevant decisions must take into account the entirety of medical knowledge, not only what they have learned from their studies and their own medical practice, but according to the principle of knowledge-based medicine, they should consider the latest research results and practical reports published in millions of scientific articles around the world. The use of artificial intelligence methods and techniques is very helpful here, although it must be remembered that even the most intelligent computer cannot make a diagnosis or recommend therapy, as these activities are reserved for a human (doctor), who is fully responsible for them. The third area of applications of artificial intelligence in medicine is controlling the operation of medical equipment used (under the supervision of a doctor) in therapy and rehabilitation. From intensive care unit (ICU) stations equipped with a multitude of electronic devices, through modern operating rooms (anesthesiology, surgical robots, in situ pathological examinations), to rehabilitation devices – the assistance of artificial intelligence is extremely useful everywhere. The fourth important area of applications of artificial intelligence in healthcare is management. Healthcare facilities (hospitals, but also clinics) today are workplaces for a huge number of people, as well as places where large amounts of money are spent, important information is collected, and significant organizational decisions are made. All of this can be supported by artificial intelligence systems that have been tested in numerous business applications.

Keywords: Biomedical Engineering, Artificial Intelligence, Neural Networks, Image Understanding, Image Recognition, Speech processing.

The presented text below is a summary of my introductory lecture, which will open the sessions of the 3rd Conference on Intelligent Systems and Information Technologies – Artificial Intelligence Impact on Physical Sciences and Medicine (24-26th September, 2025). This lecture relates to only one of the five problem areas of the Conference, namely the AI in Medicine and Healthcare area. However, limiting the topic of the introductory lecture to a single area was necessary, as it seems impossible to present a discussion that would non-trivially cover all the planned areas of discussion, due to their vast scope – from Core Methods and Theoretical Foundations through AI in Physical Sciences, AI in Social Sciences and Society, my area AI in Medicine and Healthcare, to the particularly current area of Intelligent Systems in Practice.

Narrowing the discussion to artificial intelligence in medicine, I am aware that I am touching upon an area that is in a sense "frontline." The theory of artificial intelligence (if it ever comes into being), its applications in physics (including technical sciences), as well as the social and economic problems associated with the development of artificial intelligence are undoubtedly important. However, it is in the area of artificial intelligence in medicine that the struggle for people's health and life takes place, and unfortunately, no one can know when this struggle will become a fight for their own life or the life of their loved ones. Therefore, I believe that all participants of this interesting and inspiring Conference will listen to this introductory lecture with interest, regardless of which area of artificial intelligence is the subject of their own research and scientific reflection.

I dared to agree to prepare and deliver this introductory lecture because I have been dealing with artificial intelligence in medicine for a long time, in fact since the period when the term "artificial intelligence" was not yet used in Poland at all. Indeed – as everyone knows – it was adopted in 1956 during the Dartmouth conference, but in Polish computer science of the 1970s it was practically not used. Nevertheless, when at the conference "Electronics and Automation in the Service of Medicine" (Katowice 1974) I presented the papers "Determining the Type of Required Therapy in Breast Cancer Cases by Image Recognition Using a Digital Machine" and "Application of a Digital Machine in In Vivo Virological Research" (the texts are now available on the Internet) – these were already papers de facto concerning artificial intelligence in medicine. One could somewhat jokingly point out at this point that it was a time when the term 'computer' was not yet used at all in Polish technical and popular literature, and people wrote 'digital machine' instead, as that terminology was in use at that time.

However, let's leave this prehistory behind and focus on the here and now.
In recent years, on October 25-25, 2021, I had the opportunity to speak about artificial

Figure 1: Cover and the first page of the brochure published in connection with the Economic Forum

intelligence in medicine during the 3rd Congress "Health of Poles". The materials are available online. In the same year, 2021, a brochure appeared in connection with the Economic Forum in Karpacz, published by Microsoft and the AI-Law-Tech Institute, with the subtitle "The Potential of Artificial Intelligence in the Healthcare Sector," for which, at the invitation of the editors, I wrote the Introduction and three substantive chapters (Fig. 1).

Considering the aforementioned facts, I allow myself to believe that I can also say something important and relevant about artificial intelligence in medicine to the participants of the "Intelligent Systems and Information Technologies" Conference (ISIT 2025). I would like to

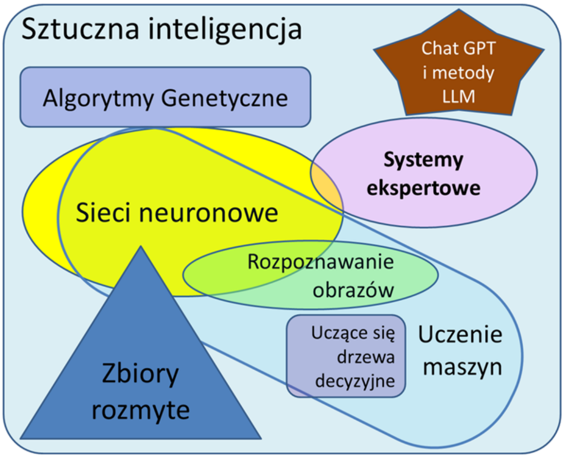


Figure 2: Map of artificial intelligence methods. Labels on the figure in Polish correspond to (from top, row by row): Sztuczna inteligencja – Artificial Intelligence, Algorytmy Genetyczne – Genetic Algorithms, Chat GPT i metody LLM – Chat GPT and LLM Methods, Systemy Ekspertowe – Expert Systems, Sieci neuronowe – Neural Networks, Rozpoznawanie obrazów – Image Recognition, Zbiory rozmyte – Fuzzy Sets, Uczące się drzewa decyzyjne – Learning Decision Trees, Uczenie maszyn – Machine Learning.

start by stating that Artificial Intelligence is not a homogeneous field of science or technology. My recently published book on this subject entitled "The Archipelago of Artificial Intelligence" (EXIT Publishing, Warsaw, 2021) vividly expresses my belief that the common term "Artificial Intelligence" describes a whole range of separate methods that allow intelligent problem-solving. However, they are based on completely different ideas and are characterized by the fact that there is generally no natural way to transition from one of these methods to another. Having said that, hybrid approaches, combining the methodology and advantages of two or more methods, that are part of this archipelago of artificial intelligence, are also sometimes used. I will allow myself to base the further structure of this lecture on a specific map of Artificial Intelligence shown in Figure 2. In this map, the size of the areas concerning different methods of artificial intelligence represents the frequency of their use, any relationships between methods are indicated by the overlapping of the presented areas, and rounded shapes symbolize methods based on mimicking solutions from real natural intelligence processes (for example, biological ones), while angular shapes symbolize methods based on various formal techniques.

The introductory lecture will indicate – with justification – which of the artificial intelligence methods shown above are used or will be used in medicine in the forthcoming years. We will also consider artificial intelligence methods for improving diagnostic signal acquisition processes (Fig. 3).

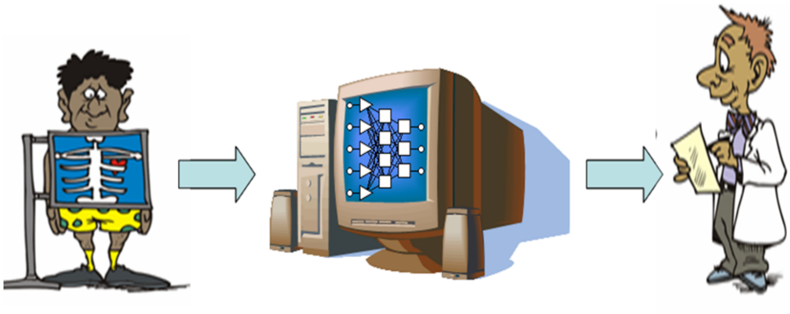


Figure 3: Artificial intelligence as a tool for improving the processes of acquiring diagnostic signals

The acquired symptoms and syndromes coming from the increasingly available diagnostic equipment, developed by advancing biomedical engineering, require thorough analysis and careful understanding by the physician aiming to make a diagnosis. Artificial intelligence can support this analysis and facilitate this understanding. Of course, no technical system can relieve the physician of responsibility for the diagnosis made and the resulting treatment outcomes, so the artificial intelligence systems discussed here cannot be considered diagnostic systems – as they are often mistakenly called. However, they are advisory systems with a very high degree of usefulness – that must be certaintly acknowledged (Fig. 4).

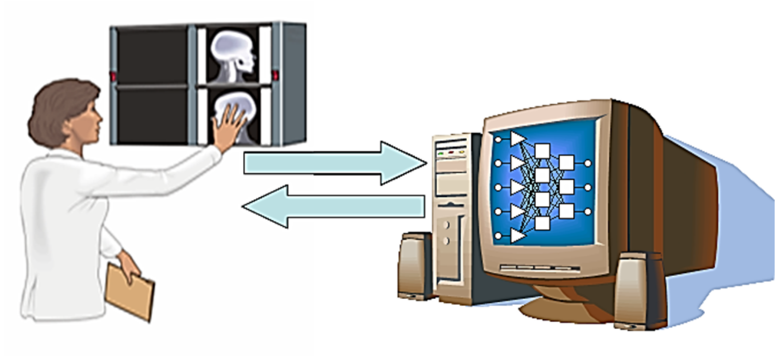


Figure 4: Artificial intelligence as a tool for improving the processes of acquiring diagnostic signals

The next stage of medical practice after making a diagnosis is the implementation and management of the therapeutic process. Therapy may involve taking medications prescribed by the doctor and undergoing recommended therapeutic procedures, and in such cases, artificial intelligence tools can be used to monitor schedules and remind medical staff (nurses) that at a specific time a certain medication needs to be administered or a certain procedure performed. Deep intelligence is not required for this.

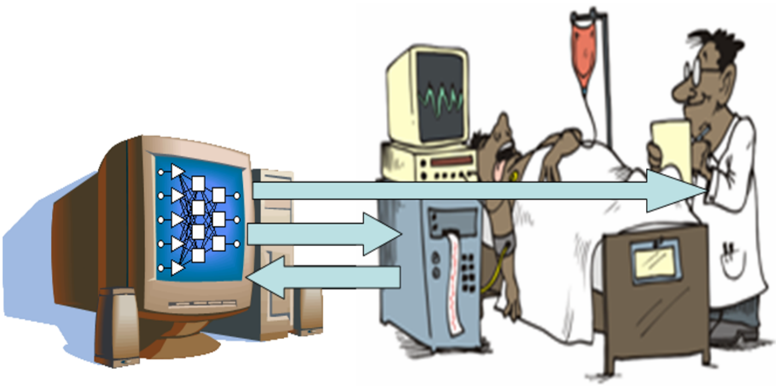


Figure 5: Artificial intelligence as a tool for improving the processes of acquiring diagnostic signals

However, in some, particularly severe cases, artificial intelligence systems might become very useful. This is especially relevant for unconscious patients placed in the ICU (Intensive Care Unit), which is a hospital ward intended for critically ill patients requiring intensive monitoring and support for vital functions such as breathing or circulation. Patients are treated and monitored there with specialized medical equipment, and when operating this advanced equipment, artificial intelligence systems become very useful in controlling the devices and communicating important observations (such as changes in patient condition) to doctors and medical staff (Fig. 5).

The last application of artificial intelligence in medicine discussed in my lecture (although very likely not the last one worth discussing) is the use of this technology to predict treatment outcomes. The predictive capabilities of artificial intelligence have been demonstrated in hundreds of cases in business applications (for example, stock market forecasts), social issues, and the military. It is purposeful and possible to apply them in medicine as well. An artificial intelligence system, by collecting data on doctors' therapeutic efforts and the results of these efforts in the form of changing patient health parameters, can provide a forecast of likely treatment outcomes. It may determine that after a certain period, the patient will be able to leave the hospital to joyfully rush around and gather precursors for another heart attack, or it may detect that despite using the best available treatment methods and measures, the patient will unfortunately die and costly and painful palliative procedures can be discontinued. This way of using artificial intelligence is illustrated in Figure 6.

The short outline of the lecture presented here, opening the proceedings of the "Intelligent Systems and Information Technologies" (ISIT 2025) Conference, may allow conference

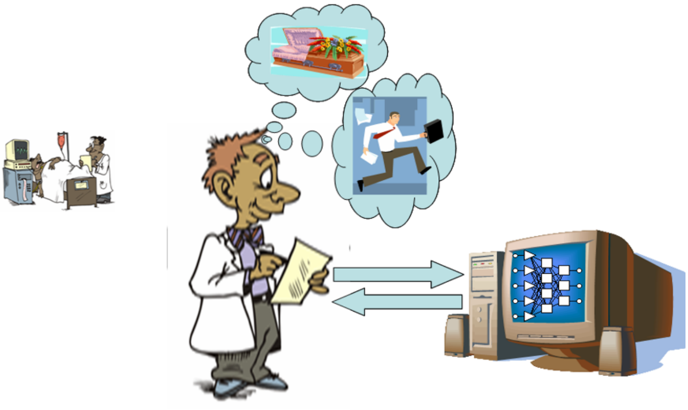


Figure 6: Artificial intelligence as a tool for improving the processes of acquiring diagnostic signals

participants to decide whether it is worth attending this lecture. In any case, I warmly invite you!

BIOGRAPHY

Ryszard Tadeusiewicz, Professor at the Stanisław Staszic AGH University of Science and Technology in Krakow, was elected Rector of this University three times between 1998 and 2005. Doctor Honoris Causa of 15 Universities. Full member of the Polish Academy of Sciences (PAN) and from 2010 to 2015 President of the Krakow Branch of PAN. Since 1998, also a member of the Polish Academy of Arts and Sciences (PAU), founder and first chairman of the PAU Commission of Technical Sciences. Computer scientist, automation engineer, and Biocyberneticist, author of 1,402 scientific publications and 126 books. Editor of 67 scientific monographs. Involved in the training of scientific staff, supervisor of 76 doctoral students, reviewer of 356 doctoral dissertations, 181 habilitations, and 167 professorial applications. Popularizer of science – wrote 1,427 popular science columns. For 10 years, he has been hosting popular science programs twice a week on RMF Classic. A more detailed biography and list of achievements can be found at: www.Tadeusiewicz.pl

Dr Habil. Eng. Marcin Paprzycki
ORCID: 0000-0002-8069-2152

Systems Research Institute
Polish Academy of Sciences
ul. Newelska 6, 01-447 Warsaw, Poland

marcin.paprzycki@ibspan.waw.pl

Extended Green Cloud Simulator – Leveraging Multi-Agent Systems for Green Cloud Simulations

Abstract. While cloud infrastructures offer multiple advantages for the delivery of computing and data resources, the environmental impact of large data centers is a reason for concern. To address this issue, companies started to seek more sustainable alternatives and proposed the concept of carbon-aware clouds, where renewable energy sources are to be utilized. However, since cloud workloads are inherently dynamic and unpredictable, their effective orchestration introduces additional challenges. Consequently, there is a need for solutions that enable the evaluation of orchestration methods before deployment into production. Among them are simulation models, which are the focus of this presentation. In particular, an Extended Green Cloud Simulator (EGCS), an agent-based framework that supports the modeling and simulation of carbon-aware clouds will be presented. Discussion will cover the following topics: (1) how agents were used in EGCS for modeling system components, (2) how their knowledge and decision-making processes are represented, and (3) how rule-based expert systems facilitated the creation of modular and adaptive orchestration strategies. Finally, the presentation will briefly demonstrate the practical use of EGCS in evaluating different resource allocation mechanisms.

Keywords: Software agents, Agent systems, Semantic technologies, High performance computing, Data analytics.

BIOGRAPHY

Associate Professor Marcin Paprzycki is employed at the Systems Research Institute of the Polish Academy of Sciences. He has an MS degree from the Adam Mickiewicz University in Poznań, Poland, a Ph.D. from the Southern Methodist University in Dallas, Texas, USA, and a Doctor of Science degree from the Bulgarian Academy of Sciences, Sofia, Bulgaria. He is a Senior Member of the ACM, a Senior Member of the IEEE, he was a Senior Fulbright Lecturer, and an IEEE CS Distinguished Visitor. His original research interests were in the area of high performance computing / parallel computing / computational mathematics. Over time they shifted towards intelligent systems, software agents and agent systems, and application of semantic technologies, among others. Currently he serves as IEEE Poland Section Conference Coordinator. He has contributed to more than 500 publications, and was invited to the program committees of over 1000 international conferences. He is on the editorial boards of 12 journals.

More information can be found in the presentation available on the ISIT 2025 Conference website: https://isit.uws.edu.pl/2025 and on the homepage: https://www.ibspan.waw.pl/~paprzyck/.

Prof. Dr Habil. Eng. Włodzisław Duch
ORCID: 0000-0001-7882-4729

Department of Informatics
Institute of Engineering and Technology
School of Physics, Astronomy and Informatics
Nicolaus Copernicus University
ul. Grudziadzka 5, 87-100 Toruń, Poland

wduch@umk.pl

Physics-informed artificial intelligence

Abstract. Understanding the neural basis of mental phenomena remains a great challenge. Attractor neural networks are our best models linking mental states with the physical properties of the brain. Information (from senses and memory) is embedded in high-dimensional patterns of neural activity and can be visualized by fMRI scans. Although biological and technological computational substrates are different at the functional level there are many similarities. Generative neural models based on diffusion processes resemble formation of brain patterns, illustrating associative aspects of thinking in large language models. Transitions between attractor states simulate chains of thought. Higher-level processes needed for reasoning strategies are inspired by cognitive science. Large neural models (LLMs) internalize information in rich contexts, learning from all kinds of data. Physics-informed systems can be first trained to respect constraints based on the laws of physics, before successful applications to predict complex real-world phenomena. Attractor network simulations show how temporo-spatial processing disorders can be related to properties of networks and individual neurons, and offer a neural interpretation of psychological phenomena. They may also explain strange observations of emergent personalities reported recently by a number of people.

Keywords: Cognitive science, Computational intelligence, Neural networks, Informatics, Neuroinformatics.

BIOGRAPHY

Professor Włodzisław Duch, in addition to heading the Neurocognitive Laboratory, is also the head of the Neuroinformatics and Artificial Intelligence team at the University Centre of Excellence in Dynamics, Mathematical Analysis, and Artificial Intelligence, and a faculty member of the Department of Applied Informatics at Nicolaus Copernicus University in Toruń. He has also worked at the School of Computer Engineering at Nanyang Technological University in Singapore as a Nanyang Visiting Professor and visiting professor (2003-2012). He served two terms as president of the European Neural Networks Society (2006-2008-2011) and was elected to the College of Fellows of the International Neural Networks Society in 2013. He was Deputy Minister responsible for science in Poland (2014-2015) and Vice-Rector for Research and Informatization at Nicolaus Copernicus University (2012-2014).

More information can be found in the presentation available on the ISIT 2025 Conference website: https://isit.uws.edu.pl/2025 and on the homepage: https://is.umk.pl/~duch/indexpl.html.

Prof. Dr Habil. Zbigniew Józef Michalewicz
ORCID: 0000-0001-5609-5457

Complexica
Level 2, 9 Charles Street
West Lakes, SA 5021
Australia

zbigniew.michalewicz@adelaide.edu.au

AI-based Business Applications for Transforming Data into Decisions

Abstract. The talk would cover a few AI-based business applications that transform data into decisions, based on work conducted for three companies (NuTech Solutions, SolveIT Software, and Complexica) over the last 20 years. A few general concepts (Adaptive Business Intelligence, Travelling Thief Problem, Larry – the Digital Analyst) would be discussed and illustrated by a few examples. The final part of the talk would present Complexica's approach for increasing revenue, margin, and customer engagement through automated analysis.

Keywords: Decision Support Systems, Adaptive Business Intelligence, Optimization & Analytics, Business Automation, AI Business Applications

BIOGRAPHY

Zbigniew Michalewicz received his M.Sc. from the Technical University of Warsaw (1974), Ph.D. from the Polish Academy of Sciences (1981), and D.Sc. in Computer Science (1997). He is Emeritus Professor at the University of Adelaide, Australia, and Chief Scientific Officer at Complexica, a provider of AI-powered solutions for optimizing business decisions.

He is internationally recognized for his work in evolutionary computation, having authored over 300 technical papers and several influential books, including Genetic Algorithms

+ Data Structures = Evolution Programs, Adaptive Business Intelligence, and How to Solve It: Modern Heuristics. His publications have received more than 50,000 citations.

Prof. Michalewicz has served as editor-in-chief of the Handbook of Evolutionary Computation and chaired the first IEEE Conference on Evolutionary Computation. He has received numerous honors, including the title of Professor from the President of Poland (2002), the Order of Polonia Restituta (2013), and the IEEE Evolutionary Computation Pioneer Award (2019). His applied research has led to major collaborations with organizations such as General Motors, Ford, the U.S. Department of Defense, and PKN Orlen.

Part II

Regular papers

Jerzy Tchórzewski[1]
ORCID: 0000-0003-2198-7185

Wojciech Nabiałek[1]
ORCID: 0000-0002-8361-2816

Kamil Kwaśniak[2]
ORCID: 0009-0009-7239-0371

[1] University of Siedlce
Faculty of Exact and Natural Sciences
Institute of Computer Science
ul. 3 Maja 54, 08-110 Siedlce, Poland

[2] Student at University of Siedlce
Faculty of Exact and Natural Sciences
Institute of Computer Science
ul. 3 Maja 54, 08-110 Siedlce, Poland

jerzy.tchorzewski@uws.edu.pl, wojciech.nabialek@uws.edu.pl, kk86192@stud.uws.edu.pl

Simulation Studies of the Quality of the Neural Model of the National Electricity Demand Forecasting System

Abstract. This paper presents research results on the use of neural models of the National Electric Power Demand Forecasting System (KSZM) trained on annual data from 2019, 2020, and 2022, respectively. The annual neural models are MIMO models with 48-hour inputs related to the total generation of the so-called JWCD (uniform system) and the so-called nJWCD (non-uniform system), and 24-hour outputs related to the national electric power demand forecasting system. The KSZM neural model, as well as a simulation model for testing the model's forecasting accuracy, were designed and implemented in MATLAB and Simulink. Many interesting research results were obtained.

Keywords: Artificial Neural Networks, Forecasting Demand for Electrical Power, Matlab, Simulink and Deep Learning Toolbox Environment, National Power System, Simulation Studies.

1 Introduction

The trained ANN model was exported to Simulink, where it was assigned as an m-file to one of the Simulink pages. A simulation model was then built that met the comparative requirements for comparing the outputs generated by the model with those generated by the system. In this simulation model, the inputs are common and concern the total generation of CDGUs and nCDGUs on an hourly basis (48 inputs). Therefore, the ANN simulation model outputs 24 values of forecasted electric power demand (one for each hour of the day). Hourly data on actual electric power demand is simultaneously collected, allowing for comparison of the forecast with observations. In this situation, discrepancies between the corresponding outputs are first sought, and based on these, the absolute error, relative error, mean relative error, and MAPE error are calculated.

2 Artificial neural networks as a model of the system

To create a model of the electric power demand forecasting system, artificial neural networks were used as a modeling method [5, 8, 9, 11-15, 17] and the MATLAB and Simulink environments with a library of m-files called Deep Learning Toolbox [1, 15, 18] were selected as the modeling tool. In this way, for annual and five-year periods, based on data from 2019-2023, neural models of the electric power demand system were obtained, both in the form of mathematical models and an object model with the possibility of using it, e.g., in Simulink [5, 13]. The Artificial Neural Network (ANN) algorithm involved numerous design and programming activities, including: selecting a training pair file for training, testing, and validation, i.e., pairs of input variables and their associated output variables; selecting the ANN architecture and its parameters for research experiments, i.e., determining the number of ANN layers and the number of neurons in each layer; and the type of connections between neurons, including assigning appropriate parameters such as weights and biases, neuron activation functions, etc. [5, 8, 11-15]. The next stage of design was the equally important selection of a training method for the modeling experiment to achieve sufficiently high learning quality, etc. [5, 8, 12, 15]. In this regard, the following were also determined: learning rate, learning accuracy measure, forgetting behavior, learning inertia behavior, etc., whereby the learning process was repeated until the ANN learning quality measure was met, i.e., the expected level of accuracy of the neural model was achieved, specifically the model outputs relative to the expected data from the electrical power demand forecasting system [5, 10]. As a result of training artificial neural networks, for example, using numerical data for the years 2019-2023, results were obtained, including the MSE error curves and the R coefficient of determination, which are presented for the entire period under study in Tables 1 and 2, respectively. The MSE error curves are presented for the entire period under study in Figure 1.

Table 1: Perceptor ANN training results for data from 2019-23 (five-year period). Source: Own study using DLT in MATLAB and Simulink environment based on data regarding the National Electric Power Demand forecasting system in the National Power System [1, 5, 18].

Unit	**Initial Value**	**Stopped Value**	**Target Value**
Epoch	0	38	1000
Elapsed Time	-	00:03:03	-
Performance	0.00344	1.59e-08	0
Gradient	0.00353	9.91e-08	1e-07
Mu	0.001	1e-10	1e+10
Validation Checks	0	0	6

Table 2: Summary of the Preceptor ANN learning model for data from 2019 to 2023 (five-year period). Source: Own study using DLT in MATLAB and Simulink environment based on data from the National Electric Power Demand forecasting system in the National Power System [1, 5, 18].

Description	**Observation**	**MSE**	**R**
Training	1279	1.5916e-08	1.0000
Validation	273	3.0653e-08	0.9999
Testing	273	3.1580e-08	0.9999

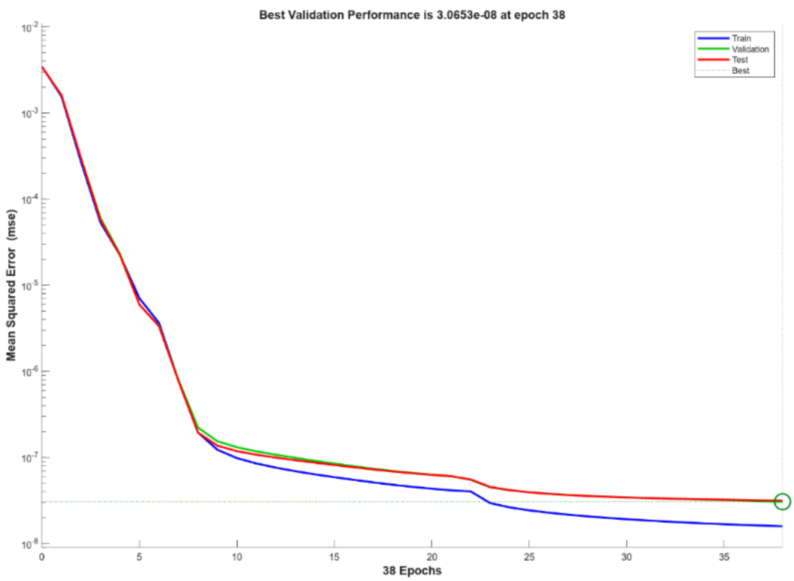


Figure 1: The MSE learning error curve of the Preceptor ANN as a model of the Electric Power Demand Forecasting System for the years 2019-2023 (five-year period). Source: Own study using DLT in MATLAB and Simulink environment [1, 5, 18].

The analysis of the results obtained shows, among other things, that the MSE learning error decreased by as much as five orders of magnitude, that is, from 0.00344 to 1.59 e-8. Meanwhile, the coefficient of determination R for learning, testing, and validation was as follows: learning process - 1.0000, testing process - 0.9999, validation process - 0.9999, that is, in all ANN learning processes at a level close to 1.0.

3 The essence of the simulation model

As part of the presented research, each trained Perceptron ANN model was exported to Simulink as a Simulink m-file, that is, for both the five-year period and individual annual periods. The simulation models were then used to build comparative models to compare the outputs generated by the model with similar outputs generated by the Electric Power Demand Forecasting System. In this simulation model, the inputs are common and refer to the total generation of CDGUs and nCDGUs on an hourly basis (48 inputs). The following are examined: discrepancy, absolute error, relative error, mean relative error, and MAPE error between the model output and the corresponding output from the Electric Power Demand Forecasting System. Therefore, the ANN simulation model produces 24 values of the predicted electric power demand (one for each hour of the day). Hourly data on actual electric power demand are collected in parallel, allowing a comparison of the forecast with observations. In this situation, the first step is to look for discrepancies between the corresponding outputs and, based on them, determine the absolute error, relative error, mean relative error and MAPE error. An illustrative example of the simulation model is shown in Fig. 2, described in detail in [13].

4 Research results

The training data used to train the neural model of the system covered the period of 2019-2025 (five-year period) and the annual periods: 2019, 2020, 2021, 2022. As a result of simulation studies using neural models of the Electric Power demand forecasting system, results were obtained regarding, among others, the MAPE error for the years 2019-2023 (five-year model) and the annual models from the years: 2020, 2021 and 2022, which are presented in Figure 3 for the five-year period and in Table 3 and Figure 4 for the annual periods, respectively. The simulation results confirm that most of the electricity demand forecast errors are within the range of approximately 0.05% to 0.26%, demonstrating the good ability of the artificial neural network to generalize data and accurately estimate demand under typical operating conditions of the National Power System. The forecasting method presented using an artificial neural network is highly effective and can find practical applications in the process of planning the operation of the National Power System.

5 Summary and Conclusions

The simulation results confirm that most of the MAPE errors obtained for forecasting the National Electric Power Demand range from approximately 0.04% to 0.26%, which demonstrates the high ability of the appropriate ANN to generalize learning and accurately estimate

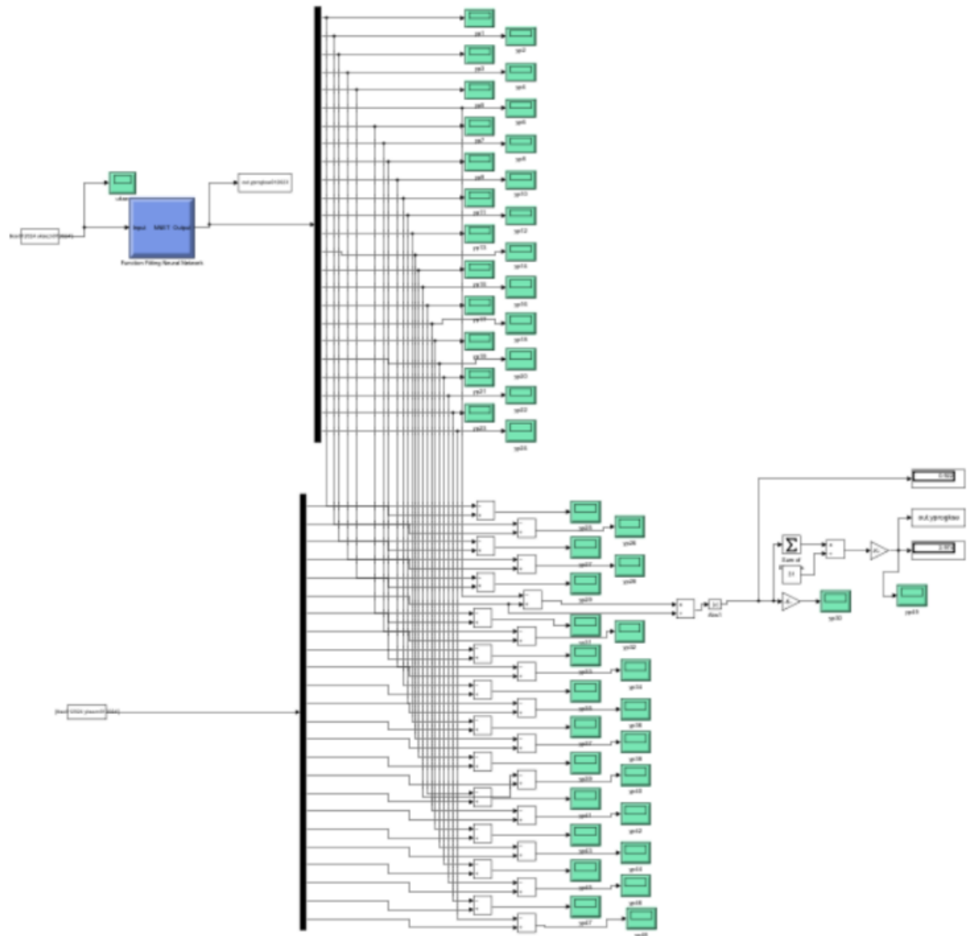


Figure 2: Illustrative simulation model for comparative testing of the Electric Power Demand Forecasting System in the National Power System for an ANN trained on a system model using data from 2023. Notation: block marked in blue – ANN as a neural model, blocks marked in green – corresponding measurement blocks, other notations in the text. Source: Own study using Simulink [1, 13, 18].

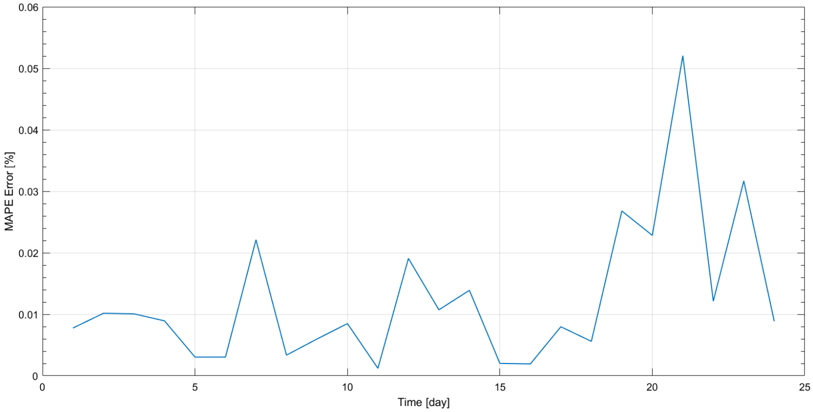


Figure 3: MAPE error results for forecasting electricity demand in 2023 obtained for individual hours of the day between actual data from 2023 and forecast data for 2023 using a neural model trained on data from 2019–23. Source: Own study using MATLAB and Simulink environments [1, 5, 18].

*

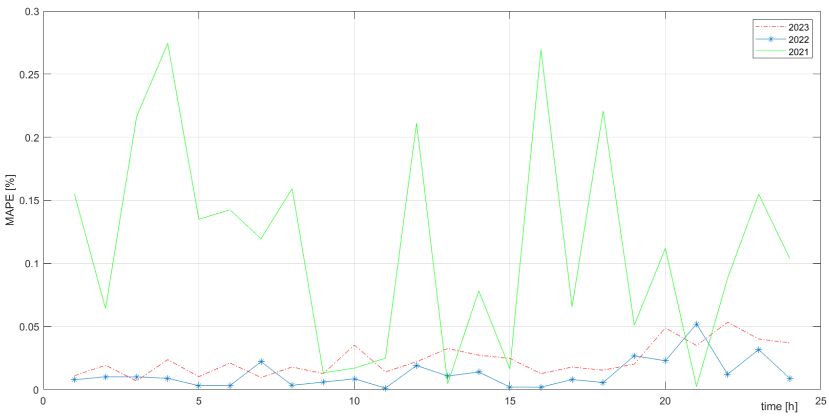


Figure 4: MAPE errors obtained for individual hours of the day for forecasts for 2021, 2022 and 2023 using models trained on annual data from 2020, 2021 and 2022, respectively [1, 5, 18].

Table 3: MAPE errors [%] for annual forecasting model data from 2021, 2022 and 2023. Source: Own study using DLT in MATLAB and Simulink environment [1, 5, 18].

Hours	**MAPE2021 Error [%)]**	**MAPE2022 Error [%]**	**MAPE2023 Error [%]**
1	0.0110445830522703	0.00777468545275929	0.155072134198624
2	0.0192895192632424	0.0101739930307191	0.0642135543006303
3	0.00685130136336092	0.0100732765598863	0.216344584838537
4	0.0237698255100997	0.00893305110746070	0.27438198353779
5	0.0102109077054381	0.00304635476371653	0.13495831473629
6	0.0210498461394865	0.00304451462908511	0.142487135976657
7	0.00955341532354505	0.0221112274867285	0.119533312554318
8	0.0179043562083209	0.00335845147276132	0.159281131919875
9	0.0126317332940480	0.00597855305587992	0.0132469570677441
10	0.0353269493519934	0.00848774959826684	0.0169587010154181
11	0.0140588777431826	0.00121233580375960	0.0248676157606924
12	0.0220800793362918	0.0190834849970303	0.211078201159486
13	0.0325344593545945	0.0107280322387975	0.00457579693869501
14	0.0272791154106499	0.0138984230983028	0.078264688557722
15	0.0247001426219411	0.00201494219930423	0.0163852300124409
16	0.0126435580133183	0.00193163538493869	0.269387590013637
17	0.0178315888276852	0.00797188277584793	0.0656180549761345
18	0.0153680847839492	0.00560090842050023	0.22078553724125
19	0.0202272881048632	0.0267990966140036	0.0509700709866889
20	0.0489071564593587	0.0228323964394754	0.111953394868906
21	0.0349013868554876	0.0520497058534259	0.00232968840997236
22	0.0534179798028544	0.0121359975496944	0.0883129665388355
23	0.0400389527247830	0.0316866850239359	0.154819908329847
24	0.0370105715782120	0.00885700344027261	0.103836771792651

forecasting the National Electric Power Demand using them under the operating conditions of the National Power System. The neural forecasting models presented in this article, i.e., the five-year model and the annual models, are characterized by high model quality and, therefore, can find practical applications in forecasting the electrical power demand of the National Power System.

References

1. M. Beale, M. Hagan, H., Deep Learning Toolbox User's Guide, The MathWorks, Natick 2020-2022.
2. T. Ciechulski, S. Osowski, „Prognozowanie zapotrzebowania mocy w KSE z horyzontem dobowym przy zastosowaniu zespołu sieci neuronowych", Przegląd Elektrotechniczny, R. 94, Nr 9, pp. 108-112, 2018.
3. I. Filipiak, W. Mielczarski, Energetyka w okresie transformacji (in Polish), in English: Power engineering in the period of transformation, WN PWN, pages 295, Warszawa 2023.
4. F. N. Humble, J. K. Gene, Przyspieszenie. Lean i DevOps w rozwoju firm technologicznych (in Polish), Helion, 2019, pages 232.
5. K. Kwaśniak, Model neuronowy i implementacja Krajowego Systemu Zapotrzebowania na Moc elektryczną w środowisku MATLABA i Simulinka, praca magisterska napisana w Instytucie Informatyki na Wydziale Nauk Ścisłych i Przyrodniczych, Uniwersytetu w Siedlcach, Siedlce 2025.
6. R. Marlęga, "Correction of the parametric model of the Day-Ahead Market system using the Artifical Neural Network", Studia Informatica. Systems and Information Technology, Vol. 1(26)2022, pp. 85-105.
7. A. Obuchowicz, Optimization of Neural Network Architectures, Chapter 9, [in:] Intelligent Systems, [eds] Wilamowski B. M., Irvin J. D., The Industrial Electronics Handbook, Second Edition, Taylor and Francis Group, LLC, pp. 9.1-9.24, 2011.
8. S. Osowski, Sieci neuronowe do przetwarzania informacji (in Polish), OW PW, Warszawa, pages 422, 2013.
9. J. Patterson, A. Gibson, Deep Learning. Parctitioner's Aproach, O'Reilly Media, Sebastopol, pages 83, 2017. 10. Polskie Sieci Elektroenergetyczne SA, Wielkości podstawowe raportów dobowych z pracy KSE (in Polish), https://www.pse.pl/dane-systemowe/funkcjonowanie-kse/raporty-dobowe-z-pracy-kse (access: 26.01.2023, 08.05.2022, 26.01.2023, 02.06.2024).
10. D. Ruciński, "Neural modeling of electricity prices quoted on the Day-Ahead Market of TGE S.A. Shaped by environmental and economics factories", Studia Informatica. Systems and Information Technologies, Vol. 1-2, No. 24, pp. 25-35, 2020.
11. R. Tadeusiewicz, Archipelag sztucznej inteligencji (in Polish), OW EXIT, Warszawa, pages 126, 2022.
12. J. Tchórzewski, „Model neuronalny 24-godzinowego jednoczesnego zapotrzebowania na moc z wyprzedzeniem dobowym w KSE", Przegląd Elektrotechniczny, R. 101, Nr 2/2025, pp. 228-232.
13. J. Tchórzewski, S. Rosochacki, "Selection of the programming environment for neural modeling of the power and electricity demand generation systems in terms of unmanned factories", Studia Informatica. Systems and Information Technology, Vol. 2, No. 29, pp.95-106, 2023.
14. J. Tchórzewski, Metody sztucznej inteligencji i informatyki kwantowej w ujęciu teorii sterowania i systemów (in Polish), Wydawnictwo Naukowe UPH, Siedlce, pages 343, 2021.
15. J. Tchórzewski, Rozwój systemu elektroenergetycznego w ujęciu teorii sterowania i systemów (in Polish), Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław, pages 190, 2013.
16. J. Tchórzewski, M. Kłopotek, „A case study in neural network evolution", Prace Naukowe IPI PAN, Tom 943, pp. 1-12, 2001.
17. MathWorks, MATLAB I Simulink, https://www.mathworks.com [dostęp: 1992-2025].

Slavko Dragic[1]
ORCID: 0009-0008-7280-9972

Radosław Marlęga[1]
ORCID: 0000-0003-0174-5074

Jerzy Tchórzewski[2]
ORCID: 0000-0003-2198-7185

[1] Instytut Techniki Górniczej KOMAG
Instytut Badawczy
ul. Pszczyńska 37, 44-101 Gliwice

[2] University of Siedlce
Faculty of Exact and Natural Sciences
Institute of Computer Science
ul. 3 Maja 54, 08-110 Siedlce, Poland

sdragic@komag.eu, rmarlega@komag.eu, jerzy.tchorzewski@uws.edu.pl

Hard Coal Market System Modeling a Case Study Using Regression Machine Learning

Abstract. This article describes a systematic approach to developing the Coal Platform (SPW). After production implementation, the platform will gain artificial intelligence features. Utilizing AI methods, such as regression machine learning, aims to enhance the platform's capabilities through predictive analysis of Poland's hard coal market. The new AI-supported SPW will help adjust annual hard coal production to forecasted demand. This includes a 12-month planning cycle and the expected consumption of hard coal. With rapid economic changes, evolving energy policy, and demand variability, AI becomes crucial for managing resources and avoiding overproduction or shortages. The article also presents quantitative research on modeling an electricity production system using hard coal. This research is still ongoing.

1 Introduction

In these times of unprecedented speed of information flow, and thus economic, administrative, and social communication, there is a growing need for methods that take into account the speed of this flow, primarily in economic processes, including the analysis of the market system for energy resources such as hard coal.

This is followed by a growing demand for both methods for analyzing the raw materials market system and for methods for modeling the market system for energy resources such as hard coal. The presented research results cover the second part of this important issue: methods for modeling the energy resources market system, using the example of the hard coal market system in Poland as an example. This was verified in a case study using the example of hard coal consumption for electricity production in commercial thermal power plants.

Modeling methods, in this case predictive modeling, are currently used in various types of prediction methods. In addition to analytical methods (e.g., polynomial regression) [1-3], attempts are being made to utilize machine learning methods (e.g., using the ARX regression method and the ARIM or AR regression method, the SVM support vector method, etc.) [3-13] and neural learning methods (e.g., Percepton ANNs, Recurrent ANNs) [14-18], as well as other SZI methods, including fuzzy methods [19-20] and methods inspired by quantum computing methods [21-22]. Currently, great hopes are also placed in Deep ANNs, especially generative and convolutional neural networks [23].

Deep learning methods are particularly readily used when researchers have time series data, which is associated with the presence of sequential relationships between the data. However, in the case of Recursive ANNs and Associative ANNs, it seems preferable to rely on Input-Output methods, as they allow for the generation of models of hard coal market systems, rather than direct forecasts of coal or hard coal prices. Obviously, the best practical solution is hybrid solutions, hence the development of various hybrid methods that combine different solutions focused on specific types of methods [15-16, 21, 24].

It is believed that hybrid solutions lead to models that generate forecasts with much greater accuracy than single-source methods, which is one of the requirements for forecasting models used to create hard coal market system models. Furthermore, the expected solution for achieving a hard coal system model is the flexibility of the predictive system planned for development at the Coal Platform, which, by dynamically adjusting production plans to changing demand forecasts, can significantly reduce both overproduction and shortages. This approach not only increases economic efficiency but also contributes to a reduced carbon footprint by limiting unnecessary resource exploitation, reducing energy losses, and optimizing logistics and storage. Not to mention the modeling effectiveness itself, which should be very high.

2 Methodology

Therefore, an important task is to conduct a comparative analysis of, among other things, the effectiveness and efficiency of selected models spanning several years, years, quarters, and even shorter time horizons (e.g., monthly, weekly, daily, hourly, and even 15-minute). This involves obtaining relevant historical data on, for example, the production, consumption, import, and export of hard coal at various stages of its movement from source to end users, as well as macroeconomic and regulatory factors such as the Gross Domestic Project, money supply, state budget revenues and expenditures, inflation, unemployment, etc.

In this modeling approach, it is also necessary to analyze the impact of data quality, seasonality, and energy policy variability on the accuracy of hard coal system models, e.g., using cluster analysis methods [25-26]. Therefore, the main goal of the research is not only to present the hard coal supply chain from its source to the end user, but also to build, following J. Konieczny [27-30], at each stage of the supply chain, the so-called praxeological model of systems linked to the security relationship and the operational relationship, and based on this, the obtained cybernetic models, and consequently, mathematical models of hard coal supply and storage systems.

Since such an approach is not possible without the use of appropriate artificial intelligence methods, the first and fundamental stage of the research, after obtaining numerical data, is the selection of artificial intelligence methods for building the SWK system, and then, after conducting preliminary simulation studies, to check the impact of this type of modeling on decision-making processes in the raw materials sector and on the potential social and economic implications resulting from its implementation, which is a very demanding task both in terms of modeling using effectiveness and efficiency criteria, which were originally called robust modeling (more broadly: robust) [5, 30].

3 Characteristics of the Research Object

Total hard coal consumption in Poland, including in commercial power plants, is declining year by year due to procedures limiting its use in the commercial power sector. A significant decline in hard coal consumption in the Polish economy was recorded in 2024, falling by 14% compared to 2023 and by as much as 33% compared to 2015. This was followed by a decline in the total share of hard coal in electricity production, which was the lowest in Poland's history, reaching 56.2% in 2024. However, in the first half of 2025, domestic thermal hard coal production reached approximately 21.5 million tons. Annual hard coal consumption in Poland, covering the needs of the energy sector, industry, and households, is estimated at 56-58 million tons.

Hard coal imports account for approximately 9.5 million tons, mainly from Colombia, South Africa, Indonesia, and Australia. Nevertheless, Poland remains the leader in European coal mining, accounting for a whopping 97% of hard coal production in the European Union. Hard coal is a key raw material in the Polish energy sector, accounting for approximately 59% of total consumption. Industry and construction use another 23.8%, and households 17.2%. Hard coal, alongside lignite, has a huge impact on the Polish economy, but its negative environmental impacts are equally significant. Due to the increasing use of renewable energy sources and the need to reduce greenhouse gas emissions, the coal sector is facing new challenges, with Poland focusing on the gradual transformation of the coal sector while simultaneously promoting the sustainable development of renewable energy [31-34].

4 A Case Study on Hard Coal System Modeling

The hard coal balance sheet, available in the Statistical Yearbooks of the Republic of Poland, covers annual hard coal revenues and outflows [31-34]. Revenues include revenues from coal extraction from domestic sources (extraction and reduction of stocks) and imported coal. Expenditures include domestic hard coal consumption (processing into other energy carriers, including in power plants, combined heat and power plants, and heating plants, as well as direct consumption), increased stocks, exports, and balance losses and differences. A quantitative presentation of this data for 2023 is shown in the table 1.

Table 1: Hard coal balance in Poland in 2023. Source: [31-34].

SPECIFICATION	**2015**	**2020**	**2022**	**2023**
	in thousand tonnes			
Supply	81112	68102	73297	65508
From domestic sources	72823	55279	53145	48600
production	72898	54714	53145	48600
decrease in stocks	137	565	—	—
Imports	6289	12823	20152	16006
Use	81112	68102	73297	65508
Domestic consumption	72742	62872	64835	55595
transformation into other energy commodities	55723	47742	52737	44533
of which in thermal power plants, combined heat and power plants and district heat plants	41261	35495	39781	32637
direct consumption	17019	16130	12098	11063
Increase in stocks	—	—	3098	5313
Exports	9192	4626	5489	4561
Losses and statistical differences	-622	604	-125	39

To conduct modeling using the arx [4] regression machine learning method, the case study used domestic hard coal consumption for electricity generation in power plants, combined heat and power plants, and heating plants (u1) [thousand tons] as input, and electricity generation in thermal power plants (y1) [GWh] as output. A quantitative presentation of these data for 2023 is shown in Table 2 [31-34]. The numerical example uses data from 1970-2023, i.e., the last 53 years. The curves of the u1 and y1 variables are shown in Figures 1, respectively.

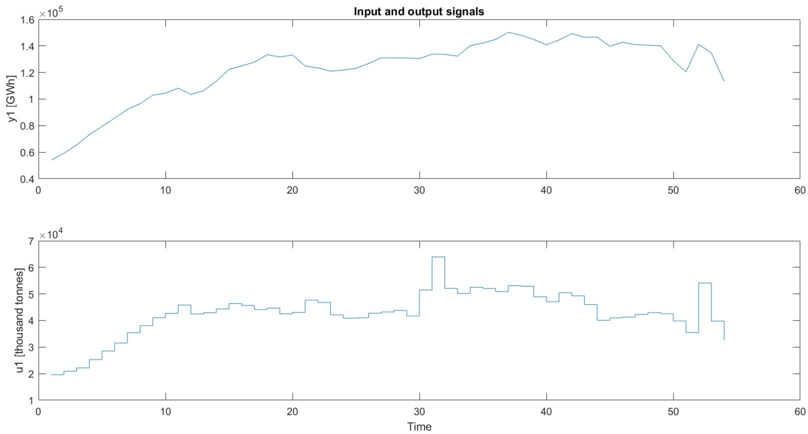


Figure 1: The course of the variable u1(t), i.e. hard coal consumption for electricity production and of the variable y1(t), i.e. electricity produced from hard coal. Source: Own study in MATLAB and Simulink using SIT [4, 35-36].

As a result of regression machine learning, a discrete model of the electricity production system in professional thermal power plants fueled with hard coal was obtained with an accuracy of 82.76% of the form (Fig. 2):

$$\text{arxec702310101} = \text{A(z)}y(t) = \text{B(z)}u(t) + e(t), \tag{1}$$

where:

$$\text{A(z)} = 1 - 0.4269z^{-1} - 0.0394z^{-2} - 0.3694z^{-3} + 0.2757z^{-4} - 0.0559z^{-5} - 0.195z^{-6} + \\ + 0.5229z^{-7} - 0.2519z^{-8} - 0.1914z^{-9} + 0.03991z^{-10}, \tag{2}$$

$$\text{B(z)} = 0.3529z^{-1} - 0.5104z^{-2} + 0.6895z^{-3} + 0.01455z^{-4} + 0.4258z^{-5} + 0.1147z^{-6} + \\ + 0.2267z^{-7} - 0.2066z^{-8} - 0.2899z^{-9} + 0.1047z^{-10}, \tag{3}$$

z – a time shift operator that delays the function argument by the appropriate number of units, e.g., $u(t)z^{-1} = u(t-1)$.

Table 2: Electricity balance in Poland in 2023. Source: [31-34].

SPECIFICATION	**2015**	**2020**	**2022**	**2023**
	in GWh			
Supply	179403	181858	178667	136986
From domestic sources — generation	164944	163990	158043	179748
of which power plants:				
public thermal power plants	142904	129006	120513	134657
hydroelectric power plants[a], wind and solar PV power plants	13350	15764	20695	31108
Import[b]	14459	17868	20024	15238
Use	179403	181858	178667	136986
Domestic consumption	154076	165662	161315	162500
for pumping water in pumped-storage hydroelectric power plants	900	1033	1181	1501
direct consumption	153176	164629	160134	167789
of which supplies to final users from the interests of distribution systems operators	126521	135157	131523	135988
of which to households[c]	28280	29393	30027	28860
Export[b]	14793	7245	7357	16915
Losses and statistical differences	10534	8951	9995	8811

a) Include pumped-storage hydroelectric power plants, run-of-river hydroelectric power plants. b) Including barrier transactions. c) Concerns consumption for community living purposes.

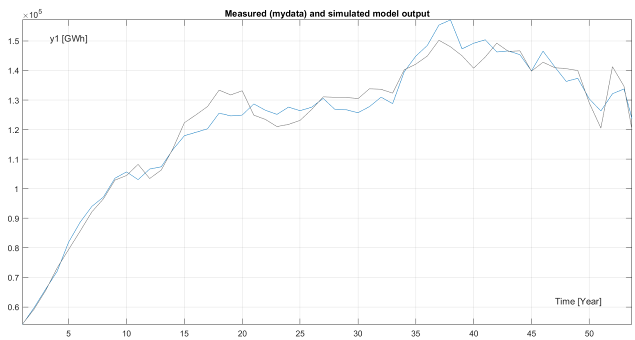


Figure 2: The volume of electricity produced in commercial thermal power plants in 1970–2023. Denomination: y1 - power energy [GWh], Symbols: blue – system data, red – model data. Source: Own study using MATLAB and SIT [4, 35-36].

This model was exported to Simulink as an object for use in further research on the praxeological sequence of using hard coal and other energy sources implemented in the country. The Polish energy sector in the period of systemic transformation requires many coordinated and sustainable actions, as the authors of the publication, among others, emphasize [37].

5 Conclusions and directions for further research

A concept for modeling the praxeological sequence of hard coal use in the technological process of Polish industry was developed. A case study of hard coal consumption for electricity generation demonstrates how to obtain a system model using the ARX regression machine learning method. A relatively high degree of model accuracy (82.76%) was achieved for the system. The model was exported to Simulink as an object, enabling its use in further intended simulation studies. Research is already underway to develop a MIMO model for hard coal use, and will also continue on other subsystems related to hard coal consumption in Poland, as well as on the use of other energy resources such as lignite, natural gas, crude oil, and renewable energy sources.

References

1. Czapaj R., Kamiński J., Benlcazar P., „Prognozowanie 15-minutowego szczytowego dobowego zapotrzebowania na moc w KSE z wykorzystaniem metody najmniejszych kwadratów" [in Polish], Zeszyty Naukowe Wydziału Elektrotechniki i Automatyki PG, Nr 62, XIX KN APE '19 (2019), pp. 45-48.
2. Czapaj R., Benalcazar P., Kamiński J., „Prognozowanie krótkoterminowe zapotrzebowania na energię elektryczną w KSE z wykorzystaniem metody MARSpines" [in Polish], Przegląd Elektrotechniczny (2019), R. 97, No. 7, pp. 133-136.
3. Dudek G., Modele wygładzania wykładniczego do krótkoterminowego prognozowania obciążeń systemów elektroenergetycznych [in Polish], Rynek Energii. Nr 3(106) (2013), pp. 15-19.
4. Zimmer A., Englot A., Identyfikacja obiektów i sygnałów. Teoria i praktyka dla użytkowników MATLABA, PK (2005), Kraków, pages 239.
5. Tchórzewski J., Rozwój systemu elektroenergetycznego w ujęciu teorii sterowania i systemów [in Polish], Oficyna Wydawnicza Politechniki Wrocławskiej (2013), Wrocław, pages 190.
6. Czapaj R., Kamiński J., Sołtysik M., "A Review of Auto-Regressive Methods Applications to Short-Term Demand Forecasting in Power Systems", Energies, 15, 6729 (2022), `https://doi.org/10.3390/en15186729` [access: 22.04.2024], pp. 1-31.
7. Tchórzewski J., Marlęga R., Nabiałek W., "Modeling an Intelligent System Using Regression Machine Learning on the Example of the Electric Power Demand System", (Extended Abstract), [in:] Niewiadomski A., Wawrzyńczak-Szaban A. (ed.): Proceedings of the 2nd Conference on Intelligent Systems and Information Technologies. Logic, Knowledge, and Reasoning in Intelligent Systems. Extended abstracts (2024), Siedlce, UPH w Siedlcach, pp. 106-113.
8. Tchórzewski J., Development models of electric power system in MATLAB environment using to System Identification Toolbox, XIV Międzynarodowa Konferencja Naukowa pt. Aktualne Problemy w Elektroenergetyce (APE). PG, Politechnika Gdańska (2009), Jurata – Gdańsk.
9. Chen G., Hu Q., Wang J., Wang X., Zhu Y., Machine-Learning-Based Electric Power Forecasting, Sustainability (2023), No. 15, 11 279, pp. 1-21.
10. Maurya D., A. K. Tangirala A. K. and Narasimhan S., Identification of Output-Error (OE) Models using Generalized Spectral Decomposition, w 2019 Fifth Indian Control Conference (ICC) (2019), New Delhi, pp. 28-53.

11. Marlęga R., “A methodology of identification and metaidentification research on the example of Day Ahead Market System, Studia Informatica. Systems and Information Technology”, Vol. 2(27)2022, pp. 109-137.
12. Marlęga R., “Hourly identification and simulation of the TGE SA Day-Ahead Market system”, Control and Cybernetics, Vol. 51 (2022) No. 4, pp. 523-555.
13. Tchórzewski J., Nabiałek W., Demiańczuk M., “Regression Machine Learning of an Hourly Model of National Electric Power Demand System”, IEEE Digital Library, PAEE, Kościelisko (2025), pp. 1-5.
14. Tchórzewski J., „Model neuronalny 24-godzinnego jednoczesnego zapotrzebowania na moc z wyprzedzeniem dobowym w KSE” [in Polish], Przegląd Elektrotechniczny, Przegląd Elektrotechniczny (2025), R. 101 NR 2, pp. 228-232.
15. Ciechulski T., Osowski S., „Badanie jakości predykcji obciążeń elektroenergetycznych za pomocą sztucznych sieci neuronowych SVM, RBF i MLP” [in Polish], Przegląd Elektrotechniczny (2014), R. 90, Nr 8, pp. 148-151.
16. Ciechulski T., Osowski S., „Prognozowanie zapotrzebowania mocy w KSE z horyzontem dobowym przy zastosowaniu zespołu sieci neuronowych” [in Polish], Przegląd Elektrotechniczny (2018), R. 94, Nr 9, pp. 108-112.
17. Siwek K., Osowski S,, Szupiluk R., “Esemble neural network approach for accurate load forecasting in a power system”, International Journal of Applied Mathematics and Computer Science (2009), Vol. 19, No. 2, pp. 303-313.
18. Tchórzewski J., “Systemic Neural Modeling of Hourly Power Demand in the National Power System,” 2024 Progress in Applied Electrical Engineering (PAEE), Koscielisko, Poland, IEEE Digital Library (2024), pp. 1-6.
19. Parol M., “Prognozowanie ultrakrótkoterminowe mocy generowanej w odnawialnych źródłach energii z wykorzystaniem logiki rozmytej” [in Polish], Przegląd Elektrotechniczny (2014), nr. 6, R. 90, pp. 265-26.
20. Ciechulski T., Osowski S., „Prognozowanie zapotrzebowania na moc w KSE z zastosowaniem grupowania rozmytego” [in Polish], Przegląd Elektrotechniczny (2009), R. 95, Nr 10, pp. 185-189.
21. Tchórzewski J., Metody sztucznej inteligencji i informatyki kwantowej w ujęciu teorii sterowania i systemów [in Polish], Wydawnictwo Naukowe UPH, Siedlce (2021), pages 343.
22. Tchórzewski J., Ruciński D., „Quantum Inspired Evolutionaty Algorythm based on Day Ahead Market of the Polish Electricity Power Exchange”, Studia Informatica. Systems and Informaion Technology, Vol. 2(31)2024, pp. 15-34.
23. Patterson J., Gipson A., Deep Learning. Praktyczne wprowadzenie [in Polish]. Helion. Gliwice 2018, pages 532.
24. Tadeusiewicz R., Archipelag sztucznej inteligencji [in Polish], OW EXIT, Warszawa (2022), pages 126.
25. Wierzchoń S., Kłopotek M, Algorytmy analizy skupień [in Polish], PWN, Warszawa 2017, pages 308.
26. Tchórzewski J., Kania T., Comparative Study of Cluster Analisis Methods Used to Evaluate Systems in the Power Industry, [Chapter 1 in:)] Theory and Application of Artificial Intellligence Methods, Monograph No. 1 in Series Intelligent Systems and Information Technologies, [ed. Tchórzewski J., Świtalski P.], pp. 11-42.
27. Konieczny J., Inżynieria systemów działania [in Polish], WNT, Warszawa 1983, pages 386.
28. Tchórzewski J., Inżynieria rozwoju systemów [in Polish], Wydawnictwo WSRP w Siedlcach, Siedlce 1990, pages 279.
29. Tchórzewski J., Cybernetyka życia i rozwoju systemów [in Polish], Wydawnictwo WSRP w Siedlcach, Siedlce 1992, pages 408.

30. Marlęga R, Identyfikacja i metaidentyfikacja systemu zarządzania Towarową Giełdą Energii Elektrycznej [in Polish], Rozprawa doktorska, promotor: dr hab. inż. Jerzy Tchórzewski, promotor pomocniczy: dr hab. inż. Arkadiusz Jurczuk, Politechnika Białostocka, Wydział Inżynierii Zarządzania, Białystok 2022, pages 238.
31. Roczniki Statystyczne Rzeczypospolitej Polskiej [in Polish and English], GUS, Warszawa 1970-2025.
32. Rocznik Statystyczny Przemysłu [in Polish and English], GUS, Warszawa, 2000, 2024.
33. Informacja o zasobach węgla kamiennego, stan na dzień 31.12.2024 roku, ARP S.A., Warszawa, czerwiec 2025, pages 20.
34. Polskie Sieci Elektroenergetyczne SA, Wielkości podstawowe raportów dobowych z pracy KSE (in Polish), https://www.pse.pl/dane-systemowe/funkcjonowanie-kse/raporty-dobowe-z-pracy-kse [access: 26.01.2023, 08.05.2022, 26.01.2023, 02.06.2024, 30.05.2025].
35. MathWorks, MATLAB i Simulink, https://www.mathworks.com [access:1992-2025].
36. Ljung L., System Identification Toolbox™ User's Guide (1988–2025), The MathWorks, Inc, Naticks, pages 1794.
37. Filipiak I., Milczarski W., Energetyka w okresie transformacji [in Polish], in English: Power engineering in the period of transformation, WN PWN (2023), Warszawa, pages 295.

Dariusz Ruciński

ORCID: 0000-0001-5458-9170

University of Siedlce
Faculty of Exact and Natural Sciences
Institute of Computer Science
ul. 3 Maja 54, 08-110 Siedlce, Poland

dariusz.rucinski@uws.edu.pl

Data Analysis for Selecting the Structure and Parameters of the Power System Model as an Intelligent System

Abstract. The electricity market and, following it, the National Power System (KSE) face dynamic transformations driven by the growing integration of renewable energy sources, changing consumer attitudes, and increasingly complex market mechanisms on the one hand, and increasingly intelligent equipment of power system devices on the other hand [5]. In this context, artificial intelligence (AI) plays a key role as a tool enabling not only technical optimization of energy systems [2-3], such as smart grids and energy consumption forecasting, but also supporting the development of new market models and real-time decision-making [4-5]. Such activities make AI not only an analytical tool but also a subject of interdisciplinary economic and regulatory research in the energy market [7]. In light of these trends, comprehensive analysis of energy market data requires integration of technical and economic approaches and the use of advanced forecasting models, including technical constraints [4]. This work represents the first conceptual and initial stage in designing and implementing a model for the functioning of the National Power System in selected areas of its operation, based on existing models using neural networks, machine learning regression models, and data illustrating the operation of this system [14, 18].

Keywords: National Power System, power system modeling, power system as an intelligent system, analysis of selected power system variables, data preparation for modeling, artificial intelligence models for modeling purposes.

1 Review of existing power system models

The construction of electricity market models and related power system models is the subject of research by many scientific centers worldwide, including in Poland [2, 6, 7, 11, 14, 16-21].

There are advanced systems using artificial neural networks, regression models, or hybrid models to model the behavior of power systems from various perspectives, including specialized resilience modeling systems:

1) HELICS (Hierarchical Engine for Large-scale Infrastructure Co-Simulation), which is a co-simulation platform from PNNL[1] and NREL[2] that utilizes machine learning elements, among others, to model interdependencies between energy, transportation, and telecommunications systems [3].
2) Pacific Northwest National Laboratory - an American national laboratory managed by Battelle Memorial Institute for the U.S. Department of Energy. [2] National Renewable Energy Laboratory - an American national laboratory focused on renewable energy and energy efficiency.
3) ResilNet, which is a system using recurrent artificial neural networks to predict cascading failures in power systems, taking into account geographical dependencies [7].

Research on the application of AI in electricity market modeling and power system modeling is also conducted in Poland, including:

1) ENERGA Living Lab as a project conducted by the ENERGA group, utilizing artificial intelligence elements to simulate consumer behavior and energy infrastructure on a local scale, with the first laboratory in Poland being an experiment conducted on 300 households in Gdynia, where innovative energy solutions were tested.
2) Numerous system studies focused on metamodels identifying development directions for the power system in Poland [15, 19, 21].

Polish systems are usually more specialized and focus on specific aspects (e.g., energy consumption prediction or modeling specific infrastructure elements rather than comprehensive representation of the power system, but the trend is moving toward more comprehensive solutions.

The described power system models were created for various purposes, from studying technical dependencies through economic relations, security, reliability, etc. The approach proposed in this work focuses on modeling issues, i.e., obtaining power system models and the possible consequences of economic, organizational, and technical decisions that can be achieved through their use in a macrosystem perspective.

[1]Pacific Northwest National Laboratory - an American national laboratory managed by Battelle Memorial Institute for the US Department of Energy.

[2]National Renewable Energy Laboratory - an American national laboratory focused on renewable energy and energy efficiency.

2 Assumptions for research on the power system model

The described power system models were created for various purposes, from studying technical dependencies through economic relations, security, reliability, etc. The approach proposed in this work focuses on modeling issues, i.e., obtaining power system models and the possible consequences of economic, organizational, and technical decisions that can be achieved through their use in a macrosystem perspective

Numerous attempts have been made so far to build a power system model, ranging from modeling attempts based on circuit theories [14], through attempts using control and systems theory [23], to attempts at creating models using artificial intelligence methods [15]. Below is a proposal for describing assumptions for a model of power system functioning. Based on these experiences, the aim of the undertaken research is to propose a model of the National Power System functioning as an intelligent system.

2.1 Purpose of obtaining the model

The power system model is being developed for use in conducting research on complex situations that need to be examined before implementing them in system reality, especially regarding system stability in the face of frequent changes occurring in the electricity market.

The model is ultimately intended to serve for analyzing dependencies between key technical and economic indicators affecting the functioning of the power system, such as: employment, installed capacity, raw material consumption, electricity import and export, and network losses, as well as quantities characterizing electrical infrastructure such as: number of turbine units, length of power lines, etc. The use of artificial neural networks as a method for building a power system model will allow for capturing dependencies and dynamic connections in the system. At the same time, the artificial neural networks themselves will be subjected to separate analysis to assess their usefulness and effectiveness in energy market modeling, which will enable optimization of their architecture and parameters.

This type of approach used for modeling power system development using machine learning regression methods is contained in work [14] and used for modeling the management system of the Polish Power Exchange in work [21]. It is worth noting here that both of the aforementioned works are based on control and systems theory, which enables the introduction of assessment of system functioning and development from the perspective of, among others, their stability, control level, and degree of internal organization [22-23].

2.2 Analysis of available data

For building the power system model, monthly data from 2017-2023 [8-9] were used, it covers both technical data (including installed capacity, number of turbine units, transmission line length) and economic data (including fuel consumption, energy import/export), which are available and regularly published by transmission system operators, ARE, URE, GUS, etc.

Before proceeding with model construction, an analysis of the collected data was conducted, including:

1. **Data normalization** concerning standardization of their values to ensure comparability and stability of the artificial neural network learning process,

2. **Aggregation and completion of missing data** through the application of interpolation methods to minimize the impact of data gaps on model quality.
3. **Comparison of models for all inputs with the PCA[3] model** by using a recurrent network.

Appropriately prepared data were then used for training, testing, and validation of artificial neural networks to obtain neural models of the National Power System.

These models can also be used for forecasting, which involves prior assessment of their quality by obtaining a relative error such as the determination index R^2, as well as measures such as MSE and MAPE, in the context of assessing their accuracy in relation to existing data that were used to build neural models..

Initial analysis the data used in neural modeling conducted within this research consists of 15 variables obtained on a monthly basis in the years 2017-2023 (12 x 7 = 84 observations for each variable). Due to the nature of the National Power System (KSE) functioning, the following were adopted for modeling:

1. **Input variables:** employment [number of persons], installed capacity [MW], number of turbine units [units], overhead line length [km], cable line length [km], hard coal consumption [t], lignite consumption [t], consumption of other raw materials (peat and wood, biogas, waste fuels, etc.) [26], natural gas consumption [million m^3], nitrogen gas consumption [million m^3], electricity import [GWh].
2. **Output variables:** achievable capacity [MW], electricity consumption [GWh], electricity export [GWh], electricity losses [GWh].

3 Data aggregation and interpolation

Due to different time intervals in which data are published (e.g., hourly, monthly, annual), in order to unify them, it was decided to analyze them in monthly time intervals by summing appropriate data published in intervals smaller than a month and by interpolating data published in longer intervals (e.g., annual) such as employment or number of turbine units.

Data aggregation is a simple problem, while data estimation methods are a more complex issue.

There are many methods, from the simplest linear interpolation through polynomial to more complex ones such as Lagrange or Hermite. Due to the nature of the data (relatively small dataset) that captures existing trends, the polynomial method was chosen for interpolation [13].

The polynomial method is implemented in the Matlab environment and is performed by the function *polyfit()*, which fits a polynomial using the least squares method, finding the coefficients of a polynomial of degree *n* that best describes the relationship between independent variables *x* and their corresponding values *y*. For a polynomial of degree *n*, it is represented as follows:

$$P(x) = p_1 x^n + p_2 x^{n-1} + \ldots + p_{n+1} \tag{1}$$

[3]Principal Component Analysis (PCA) is a dimensionality reduction technique that transforms high-dimensional data into a lower-dimensional space while preserving as much of the original variance as possible.

where its essence lies in minimizing the sum of squared deviations:

$$\min \sum_{i=1}^{n} \left[y_i - P\left(x_i\right)\right]^2 \tag{2}$$

However, some data, such as the number of turbines and the length of overhead lines, change "step by step", so it was assumed that the monthly data for a given year would be the values reported for that year.

3.1 Normalization

The process of data normalization involves transforming data to a common scale so that different variables become comparable and no single variable dominates the analysis due to its range of values. Many normalization methods are used. For the purposes of the modeling studies, Z-Score normalization (Standardization) was selected, which calculates data according to the following formula::

$$Z = \frac{X - \mu}{\delta}, \tag{3}$$

where:

X - original value of the variable

μ - arithmetic mean,

σ - standard deviation.

As a result of this normalization, the data are transformed into a distribution with a mean of zero and a standard deviation of 1, while the shape of the original distribution is retained. Positive values indicate points above the mean, and negative values indicate points below the mean. [24].

4 Principal Component Analysis

An attempt was also made to identify the factors that have the greatest impact on the National Power System. In the literature, many models are available for use in this type of research, including Principal Component Analysis (PCA), Independent Component Analysis (ICA), Empirical Mode Decomposition (EMD), Non-negative Matrix Factorization (NMF), Wavelet, etc. In the research conducted, the PCA method [24] was chosen to examine the correlations, as it is a dimensionality reduction technique that involves transforming a set of potentially correlated variables into a smaller set of uncorrelated variables called principal components, quite effectively finding the directions of greatest variance in the data and projecting the data onto these directions. Its mathematical form is as follows:

$$\mathbf{Y} = \mathbf{X}\mathbf{W}, \tag{4}$$

where:

Y - matrix of principal components,

X - matrix of input data after standardization,

W - matrix of eigenvectors (loadings).

The computational process of the method proceeds as follows:

1) data standardization according to the formula:

$$X_{std} = \frac{X - \mu}{\sigma} \tag{5}$$

2) determining the covariance matrix according to the formula:

$$W = \frac{1}{n-1} X_{std}^T X \tag{6}$$

3) decomposition according to the formula:

$$\mathbf{C} = \mathbf{W \Lambda W}^T \tag{7}$$

where:

Λ - matrix of eigenvalues,

W - matrix of eigenvectors.

Due to the fact that the data subjected to analysis are characterized by high multidimensionality, there is a high probability of correlation between corresponding variables. PCA reduces data dimensionality by transforming it into a lower-dimensional space while maintaining as much data variance as possible (MATLAB has a built-in function pca() that returns the variances of the principal components).

PCA - Results for input variable:
PC1: 57.5% variance (cumulative: 57.5%)
PC2: 20.0% variance (cumulative: 77.5%)
PC3: 15.6% variance (cumulative: 93.1%)
PC4: 3.0% variance (cumulative: 96.0%)

Factor analysis reveals which primary variables have the strongest influence on each component. As can be seen in the case of the first component, PC1, which accounts for 57.5% of the variance. Below are the weights showing the share of the individual component factors that make up the PCA directions.

PC1 = 0.415×CableLines + (-0.396)×NumberTurbineUnits + 0.392×InstalledPower + (-0.366)×Employment + 0.342×OtherMaterials + ...

PC2 = 0.653×OverheadLines + 0.514×GasNatural + (-0.393)×GasNitrogenized + 0.277×HardCoal + 0.135×BrownCoal + ...

PC3 = 0.541×HardCoal + 0.420×BrownCoal + (-0.411)×NumberTurbineUnits + (-0.401)×ImportEE + (-0.369)×GasNatural + ...

The above results indicate the possibility of dimensionality reduction. The first 3 principal components explain 93.1% of the total data variance, which means that from 11 original variables, one can effectively reduce to 3 calculated values without significant information loss.

Most variables (employment, capacity, infrastructure) are correlated with each other, suggesting a variance distribution (57.5% →20.0% → 15.6%). Meanwhile, the first 4 components preserve 96.0% of variance, which almost constitutes complete information.

5 Implementation of the KSE system by modeling Recurrent ANN

After performing data analysis to verify their suitability for KSE system modeling, they were subjected to neural modeling using recurrent ANNs (RNNs) for later use of the model to predict key system variables. In design and implementation, 84 data samples were used that characterize various aspects of the functioning of the KSE system, from human resources to variables characterizing technical infrastructure and consumption of raw materials. energy.

Recurrent ANN Architecture and the obtained KSE system model

The following Recurrent ANN parameters were selected for training, testing, and validation: Time delays – 2, one hidden layer with 32 neurons, learning function - Bayesian Regularization, data split: 70% training, 15% validation, 15% testing. The ANN learning process was characterized by high efficiency, achieving convergence after 26 epochs in 30.21 seconds, with the best validation result at $1.61 \cdot 10^{-3}$ after the 16th epoch for 11 original input values and training time 71.10 seconds, number of epochs 84, best validation result $2.895659e^{-3}$ for PC inputs.

General prediction quality indicators

The trained KSE neural system model using Recurrent ANN showed satisfactory prediction quality, both for the entire dataset and data limited to PC1 and PC2.

Table 1: Table 2. Summary of RNN learning results for all inputs and for PC1 and PC2 inputs. Source: own elaboration in MATLAB and Simulink environment [1, 12].

Metric	All 11 inputs	PC -2 inputs	11 inputs vs 2 inputs
MSE	68544.79	494705.738604	-426 160.95
RMSE	261.81	703.353210	-441.54
MAE	149.4	316.088062	-166.62
MAPE	5.89%	7.02%	-1.13%
R^2	0.9988	0.9543	-0.0445

As can be seen, the network training results for the 11 original inputs and the two PCA inputs are satisfactory. The 11-input model still performs slightly better than the model trained with PC1 and PC2 data, including R^2. Prediction errors are slightly lower, and R^2 is closer to 1 (0.9998 vs. 0.99543). Nevertheless, both models explain the energy market system relatively well. As a result of the conducted model behavior research, the following were obtained in the case

Table 2: Table 3. Summary of RNN learning results for all inputs and for PC1 and PC2 inputs. Source: own elaboration in MATLAB and Simulink environment [1, 12].

Output features	**All 11 inputs**		**PC - 2 inputs**	
	R^2	**RMSE:**	**R^2**	**RMSE:**
Available Power	0.9955	477.10	0.9937	1381.75
Consumption	0.9970	14.92	0.9981	2.97

Output features	All 11 inputs R^2	RMSE:	PC - 2 inputs R^2	RMSE:
Electricity Export	0.6642	215.20	0.6777	263.80
Losses	0.9948	3.79	0.9967	1.16

For individual results, training errors are also similar; the difference is relatively small and depends on individual network training attempts. This suggests that for large, multidimensional datasets that are difficult to model, PCA can be used for dimensionality reduction.

6 Conclusions and directions for further research

The research objective has been achieved, as a preliminary data analysis was conducted to select the architecture and parameters of artificial neural networks (ANN) for training a neural model of the National Power System (KSE) as an intelligent system. The model's strengths include:

- high overall accuracy, as the neural model of the (KSE) demonstrates exceptionally high capability to reproduce complex dependencies in the power system, confirmed by an R^2 value of 0.9998,
- very good computational efficiency characterized by a very short training time of 30 seconds, which, combined with achieving high prediction quality, makes the model a practical tool for system operators in the prediction process,
- convenient accuracy differentiation, as the model best predicts parameters directly related to technical infrastructure (power, consumption, losses), which is particularly important in operational planning when using the model for prediction.

However, there are certain limitations and areas for further work and analysis, particularly regarding the output variable **energy export**, which is characterized by a low coefficient of determination R^2 value of 0.6642, indicating the need to include additional external variables, such as prices on international markets or energy policies of neighboring countries.

Directions for further research include:

1. the possibility of expanding the dataset, including the incorporation of additional external variables, particularly to improve energy export prediction,
2. conducting temporal analysis on models with a greater number of time delay constants to better capture system dynamics,
3. implementing operational deployment to demonstrate that the obtained neural model of the (KSE) can be effectively used for short-term operational planning, particularly in power management and electricity consumption forecasting.

References

1. Beale M. H., [et al.] , Neural Networsk Toolbox™ User's Guide, by The MathWorks, Inc (1992-2025).
2. Diao []et al.] , Autonomous Voltage Control for Grid Operation Using Deep Reinforcement Learning.(2019) https://doi.org/10.48550/arXiv.1904.10597, 5 p. [accessed: June 2025].
3. Hardy T. D., Palmintier B., Top P, Krishnamurthy D., and Fuller J. C. HELICS: A Co-Simulation Framework for Scalable Multi-Domain Modeling and Analysis. IEEE Access, No 12, (2024), pp. 24325 – 24347,doi:10.1109/ACCESS.2024.3363615
4. https://www.e3sconferences.org/articles/e3sconf/pdf/2024/121/ [accessed: May 2025].
5. https://energyaction.com.au/strategic-insights-energy-market-analysis/ [accessed: April 2025].
6. https://arena.gov.au/assets/2017/02/CSIRO-Electricity-market-analysis-for-IGEG.pdf [accesed: April 2025].
7. .https://www.neuron-eranet.eu/projects/ResilNet [accessed: April 2025].
8. https://www.pse.pl/dane-systemowe/funkcjonowanie-kse/raporty-dobowe-z-pracy-kse [accessed: April 2025].
9. https://www.are.waw.pl/badania-statystyczne [accessed: April 2025].
10. https://www.rynekelektryczny.pl/moc-zainstalowana-oze-w-polsce [accessed: April 2025].
11. Li, S., [et al.] Cyber-physical cascading failure and resilience of power grid: A comprehensive review. Frontiers in Energy Research Frontiers. (2023). doi: 10.3389/fenrg.2023. 1095303, 17 p.
12. Matlab. (2023). Help for MATLAB and Simulink. Retrieved from [accessed: April 2025].
13. Tan, P-N., Steinbach, M., Kumar, V. (Introduction to Data Mining. Pearson International, 2019), 719 p.
14. Tchórzewski J. Rozwój systemu elektroenergetycznego w ujęciu teorii sterowania i systemów, engl:(Development of the power system in terms of control and systems theory). Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2013, 190 p.
15. Tchórzewski J., Metody sztucznej inteligencji i informatyki kwantowej w ujęciu teorii sterowania i systemów, egl:(Artificial Intelligence and Quantum Computing Methods from a Control and Systems Theory Perspective), Wydawnictwo Naukowe UPH, Siedlce 2021, 343 p.
16. Tchórzewski J., Kania T., Comparative Study of Claster Analysis Methods Used to Evaluate Systems in the Power Industry, [Chapter 1 in:] Theory and Application of Artificial Intelligence Methods, Monograph No. 1 in Series: Intelligent Systems and Information Technology, [ed. J. Tchórzewski, P. Świtalski], Wydawnictwo Naukowe UPH, Siedlce 2022, pp. 11-42.
17. Ruciński D., Forecasting Accuracy of LSTM and Perceptron Models in Day-Ahead Market Prediction on the Polish Power Exchange S.A., [Chapter 2 in:] Design and Implementation of Artificial Intelligence Systems, Monograph No. 3 in Series: Intelligent Systems and Information Technology, [ed. D. Mikułowski, A. Niewiadomski], Wydawnictwo Naukowe UwS, Siedlce 2025, pp. 33-56.
18. Ruciński D., Modeling of the Day-Ahead Market on the Poish Power Exchange on the Example of Selected Neural Networks, [Chapter 1 in:] Theory and Application of Artificial Intelligence Methods, Monograph No. 1 in Series: Intelligent Systems and Information Technology, [ed. J. Tchórzewski, P. Świtalski], Wydawnictwo Naukowe UPH, Siedlce 2022, pp. 85-118.
19. Ruciński D., Selected Aspects of Neural-Evolutionary Modeling of Proces on the Day Ahead Market of TGE S.A., Studia Informatica. Systems and Information Technologies, Vol. 2(31)2024, pp.69-86.
20. Ruciński D., Neural modeling of prices on the Day-Ahead Market at the Polish Power Exchange supported by an Evolutiary Algorithm and inspired by quantum computing, Control and Cybernetics, Vol. 4/2022, pp. 557-584.

21. Marlęga R., Comparative study of selected methods of analytical , identification and neural modeling on example of the Day-Ahead Systems, ., [Chapter 6 in:] Modeling and Analysis of Intelligent Information Systems, Monograph No. 2 in Series: Intelligent Systems and Information Technology, [ed. J. Tchórzewski, P. Świtalski], Wydawnictwo Naukowe UwS, Siedlce 2023, pp. 145-198.
22. Klamka J., Stan badań w zakresie teorii sterowania, engl:(State of the art of control theory research), PAK vol. 57, nr 11/2011, pp. 1436-1440.
23. Kaczorek T., Dzieliński A [et al], Podstawy teorii sterowania, engl:(Basics of control theory), WNT, Warszawa 2005, 497 s.
24. Ralston, A., Rabinowitz, P., Wstęp do analizy numerycznej, engl:(Introduction to numerical analysis). Wydawnictwo Naukowe PWN 1987. 306 p.

Anna Wawrzyńczak[1,2]
ORCID: 0000-0001-8292-6875

Dominik Filipczuk[1]
ORCID: 0009-0006-5327-4277

[1] University of Siedlce
Faculty of Exact and Natural Sciences
Institute of Computer Science
ul. 3 Maja 54, 08-110 Siedlce, Poland

[2] National Centre for Nuclear Research
ul. Andrzeja Sołtana 7, 05-400 Otwock-Świerk, Poland

anna.wawrzynczak-szaban@uws.edu.pl, dominik.filipczuk@uws.edu.pl

Improving Airborne Contamination Forecasting with Tree-Structured Parzen Estimator-Based Neural Network Tuning

Abstract. The rapid development of artificial intelligence has facilitated the use of neural networks in complex, time-sensitive areas, including environmental hazard prediction. This paper introduces a method for optimising the hyperparameters of a neural network model to predict the spatial dispersion of atmospheric contaminants in urban environments. The optimisation is conducted using the Tree-Structured Parzen Estimator (TPE), a Bayesian optimisation algorithm that is implemented through the Optuna framework.
The study evaluates the performance of models trained with manually selected hyperparameters compared to those optimised automatically. Results indicate that TPE significantly enhances the model's predictive performance, achieving higher accuracy while requiring fewer computational resources by effectively pruning less promising configurations. The optimised model demonstrates its potential for real-time environmental monitoring and emergency response systems, providing a faster and more accurate alternative to traditional dispersion models.

Keywords: deep neural network, Bayesian optimisation, emergency preparedness, contaminant atmospheric transport

1 Introduction

The study in this paper aims to develop an emergency-response system capable of real-time localisation of airborne toxic release sources in urban areas, using sparse concentration data from a sensor network. Quick identification of the source location is crucial for effective crisis response and mitigation. In the literature, this task is referred to as source term estimation (STE) [1], and it involves reconstructing unknown source parameters by comparing model predictions with sensor observations, typically through optimisation or Bayesian inference [2].

Urban areas introduce substantial complexity to this problem due to turbulent airflow and interactions with dense building structures, significantly hindering real-time localisation. Traditional approaches often require repeatedly running computational fluid dynamics (CFD)-based dispersion models, which is too time-consuming for operational use.

To address this, previous studies [3, 4, 5] proposed using artificial neural networks (ANNs) as fast surrogates for dispersion models. These ANNs were trained to predict contaminant concentration fields based on input parameters such as source coordinates, emission time, wind speed, and direction. However, the hyperparameters of the networks—such as architecture, learning rate, and regularisation settings—were selected manually, which may limit the model's generalizability and performance.

This paper extends the earlier work by introducing a systematic approach to ANN hyperparameter tuning using the Tree-Structured Parzen Estimator (TPE) algorithm, a Bayesian optimisation method designed for efficient exploration of complex hyperparameter spaces [6, 7]. The goal is to improve model accuracy and robustness while preserving its suitability for real-time STE applications in urban environments.

2 Neural network architecture optimisation by TPE algorithm

The artificial neural network (ANN) was employed in this study to predict the concentration of a contaminant released in an urban environment. The goal of the network is to estimate the concentration at a given location and time for a specific release scenario. The input layer of the ANN consists of neurons encoding the parameters of the release and the observation point:

$$\text{Input}_i \equiv \{X_s, Y_s, Q, d, x, y, v, t\},$$

where (X_s, Y_s) denote the coordinates of the contamination source (in meters within the domain), Q is the release rate, d is the duration of the release (in seconds), v is the wind direction, (x, y) are the coordinates of a sensor, and t is the time (in seconds) since the beginning of the release. Given this input, the network outputs:

$$\text{Output}_i \equiv C_i^{S_j(x,y)}(t),$$

representing the contaminant concentration C at sensor S_j located at (x, y), at time t after the onset of the release.

To ensure high predictive performance, the architecture of the ANN must be carefully tuned. This includes selecting an appropriate number of hidden layers, the number of neurons per layer, activation functions, learning rate, and other hyperparameters. These choices are not learned during training and can significantly affect the model's accuracy and generalisation. Manual tuning is time-consuming and suboptimal, especially in high-dimensional and non-convex search spaces. Therefore, automated hyperparameter optimisation (HPO) techniques are increasingly used.

In this work, we adopted the TPE algorithm [7] to optimise the ANN architecture. TPE is a Bayesian optimisation (BO) method that is particularly effective for complex search spaces involving continuous and categorical variables. Unlike traditional BO, which models the posterior distribution $p(y|x)$ of the objective value y given parameters x, TPE reverses this and models $p(x|y)$. It splits the set of evaluated configurations into two groups based on a performance threshold y_γ: the "good" group D_l with $y \leq y_\gamma$ and the "bad" group D_g with $y > y_\gamma$. For each group, TPE fits separate kernel density estimators (KDEs), denoted $p(x|D_l)$ and $p(x|D_g)$.

To propose new hyperparameter configurations, TPE maximises the following ratio:

$$r(x|D) = \frac{p(x|D_l)}{p(x|D_g)},$$

which indicates how likely a candidate x belongs to the good-performing group relative to the bad-performing one. This acquisition function allows the algorithm to efficiently balance search space exploration with the exploitation of high-performing regions.

TPE supports various kernel functions and weighting strategies. Gaussian kernels are commonly used for continuous parameters, while Aitchison–Aitken kernels are applied for categorical variables. In our case, the TPE algorithm was used to tune the number of hidden layers and neurons in the network to maximise the coefficient of determination R^2, which served as the objective function during optimisation.

The dataset used to train the network was generated based on a simulated scenario of neutral gas dispersion over the central part of London. Details regarding the simulation setup, computational domain, and dataset generation process are provided in [4]. This realistic, urban-based scenario served as a representative test case for evaluating the predictive capabilities of the ANN under complex atmospheric dispersion conditions.

3 Results and Analysis

A total of 12 optimisation trials were conducted, each corresponding to a unique configuration of the neural network model. Hyperparameters were dynamically selected using the TPE algorithm to maximise the coefficient of determination (R^2) on the validation set. Each model was trained and validated on the same dataset to ensure fair comparison. The optimised hyperparameters included the number of hidden layers (ranging from 6 to 12), the number of neurons per hidden layer (between 250 and 600), and the learning rate (from 10^{-5} to 10^{-1}). All models used the Adam optimiser and were trained using supervised learning with MSE loss and R^2 as evaluation metrics. The history of individual tests is visible in Figure 2 presenting the objective function value throughout the learning.

The best result was achieved in Trial 10, where the network reached an $R^2 = 0.905 \pm ??$ after 50 epochs. This configuration featured six hidden layers with 513, 250, 465, 426, 352,

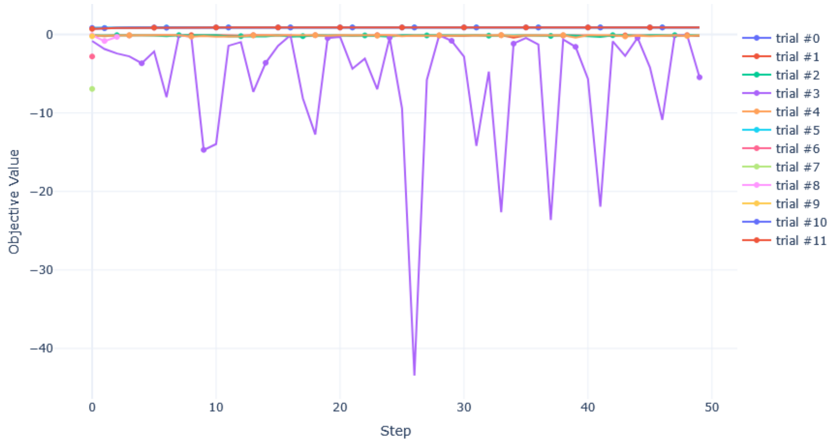


Figure 1: The course of the objective function training of individual models

and 350 neurons, respectively, and a learning rate of 1.27×10^{-5}. The model was trained for an additional 50 epochs to enhance further performance, which improved the MSE from 0.021 to 0.019 and increased the validation R^2 to 0.9165. Final evaluation on the test set confirmed the model's strong generalisation, as shown in Figure 2.

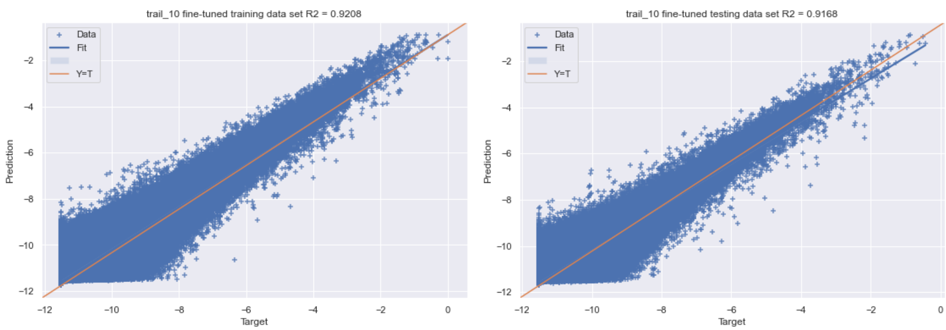


Figure 2: The scatter plots for the selected best model (Trial 10) for the training and testing datasets, and corresponding determination coefficient R^2. The dashed line represents the ideal fit.

All experiments were implemented in Python using the open-source Optuna framework, which provides a flexible interface for automated hyperparameter optimisation based on the TPE algorithm [8].

4 Conclusions

The conducted experiments demonstrated that using the Tree-Structured Parzen Estimator (TPE) algorithm for neural network hyperparameter optimisation leads to a significant improvement in predictive model performance compared to traditional manual tuning approaches presented in [4] and [5]. Higr determination coefficient ans smaller MSE values have been achieved, leading to better prediction of the contaminant concentration level. The application of the Optuna framework provided not only better model accuracy but also improved the efficiency and organisation of the experimental process. A key advantage of the TPE algorithm lies in its ability to estimate the potential quality of new configurations without requiring complete evaluation. By modelling the distribution $p(x|y)$ and prioritising configurations with a high likelihood of belonging to the top-performing group, TPE efficiently narrows the search space, reduces the number of ineffective training runs, and optimises the use of available computational resources.

Acknowledgements The computations presented in this study were carried out on the DELL PowerEdge T640 computing server owned by the Institute of Computer Science at the University of Siedlce.

References

1. Hutchinson, M., Oh, H., Chen, W.: A review of source term estimation methods for atmospheric dispersion events using static or mobile sensors. Inf. Fusion **36**, 130—148 (2017)
2. Kopka, P., Wawrzynczak, A.: Framework for stochastic identification of atmospheric contamination source in an urban area. Atmospheric Environment **195**(1), 63–77 (2018) https://doi.org/10.1016/j.atmosenv.2018.09.035
3. Wawrzynczak, A., Berendt-Marchel, M.: Can the Artificial Neural Network Be Applied to Estimate the Atmospheric Contaminant Transport? In: Dimov, I., Fidanova, S. (eds) Advances in High Performance Computing. HPC 2019. Studies in Computational Intelligence vol 902, pp. 132–142. Springer, Cham. (2021)
4. Wawrzynczak, A., Berendt-Marchel, M.: Feedforward neural networks in forecasting the spatial distribution of the time-dependent multidimensional functions. 2022 International Joint Conference on Neural Networks, IJCNN, pp. 1-8. Padua, Italy (2022). https://doi.org/10.1109/IJCNN55064.2022.9892001
5. Wawrzynczak, A., Berendt-Marchel, M. (2023). How to Select Superior Neural Network Simulating Inner City Contaminant Transport? Verification and Validation Techniques. In: Mikyška, J., de Mulatier, C., Paszynski, M., Krzhizhanovskaya, V.V., Dongarra, J.J., Sloot, P.M. (eds) Computational Science – ICCS 2023, Lecture Notes in Computer Science, vol 14074. Springer, Cham. (2023)
6. Bergstra, J., Bardenet, R., Bengio, Y., & Kégl, B. (2011). Algorithms for hyper-parameter optimization. *Advances in Neural Information Processing Systems*, **24**, 2546–2554. Retrieved from https://proceedings.neurips.cc/paper_files/paper/2011/file/86e8f7ab32cfd12577bc2619bc635690-Paper.pdf
7. Watanabe, S. (2023). Tree-structured parzen estimator: Understanding its algorithm components and their roles for better empirical performance. arXiv preprint arXiv:2304.11127.
8. Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna: A next-generation hyperparameter optimization framework. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 2623–2631). Association for Computing Machinery.https://doi.org/10.1145/3292500.3330701

Agnieszka Siłuszyk[1]
ORCID: 0000-0002-5723-9923

Agnieszka Gil[1,2]
ORCID: 0000-0002-1291-5908

Renata Modzelewska[1]
ORCID: 0000-0002-9669-7716

Marek Siłuszyk[3]
ORCID: 0000-0001-6384-2086

Anna Wawrzaszek[2]
ORCID: 0000-0001-9946-3547

Anna Wawrzyńczak[4,5]
ORCID: 0000-0001-8292-6875

[1] University of Siedlce, Faculty of Exact and Natural Sciences, Institute of Mathematics
ul. 3 Maja 54, 08-110 Siedlce, Poland

[2] Space Research Centre of Polish Academy of Sciences
ul. Bartycka 18A, 00-716 Warszawa, Poland

[3] Polish Air Force University
ul. Dywizjonu 303 35, 08-521 Deblin, Poland

[4] University of Siedlce, Faculty of Exact and Natural Sciences, Institute of Computer Science
ul. 3 Maja 54, 08-110 Siedlce, Poland

[5] National Centre for Nuclear Research
ul. Soltana 7, 05-400 Otwock, Poland

agnieszka.siluszyk@uws.edu.pl, agnieszka.gil-swiderska@uws.edu.pl,
renata.modzelewska@uws.edu.pl, marek.siluszyk@uws.edu.pl,
sanna@cbk.waw.pl, anna.wawrzynczak-szaban@uws.edu.pl

Machine Learning for Geomagnetic Storm Analysis: Uncovering Patterns in Solar Cycle 24 Variables

Abstract. We present a machine learning–driven framework for analyzing geomagnetic storms during Solar Cycle 24 (2010–2021). Using Principal Component Analysis (PCA) on a 13-dimensional feature set—including heliospheric magnetic field components, solar wind metrics, and geomagnetic indices—we extract latent structures and correlations in storm dynamics. To enhance interpretability, we apply perception-based visual analytics with Chernoff Faces, mapping key parameters (Kp, B_z, Dst) to visual features for intuitive clustering. The results reveal that the strongest storms align with low solar wind temperatures, suggesting links to magnetic cloud events. This work illustrates how dimensionality reduction, pattern recognition, and human-centered visualization can accelerate discovery in space weather research, bridging computational intelligence and physical insight.

Keywords: Machine Learning, Space Weather, Principal Components Analysis, Visual Analytics, Geomagnetic Storms, Pattern Recognition

1 Data and Methodology

A time of data that we can live in, observe, collect information from, and process for our needs is an extraordinary and outstanding world. Nowadays, alongside developments in physics, mathematics, chemistry, biology, and industrial sectors, we witness numerous studies focused on human life, health, nature, and behaviors using intelligent systems and information technology. It is only natural to seek to understand the mechanisms and technologies that can address pressing issues, ultimately enabling people to live as comfortably as possible. One critical area of interest relates to space weather phenomena. Two primary domains under closer examination are solar activity levels and their geoeffectiveness during the period from 2010 to 2021, specifically within Solar Cycle 24.

We investigate the connections between solar, heliospheric, and geomagnetic parameters during intense geomagnetic storms. For this purpose, we consider data with one-hour resolution, particularly examining the heliospheric magnetic field (HMF) magnitude, B, and HMF components B_y, and B_z; solar wind parameters such as density (SWd), temperature (SWT), dynamic pressure (SWp), and speed (SWs); as well as geomagnetic indices including Kp, ap, Dst, and AE sourced from the OmniWeb Data Explorer (`https://omniweb.gsfc.nasa.gov/`). Additionally, we utilize the computed geoelectric field, E_c [5], derived from the one-minute geomagnetic data measured at the Belsk observatory, which is part of the international INTERMAGNET network [3].

Our objective is to capture the similarities and characterize geomagnetic storms using the simplest possible features. More specifically, we studied 38 geomagnetic storms from 2010 to 2021, addressing only a few aspects of the vast space weather theory and its associated challenges. The selection and examination of parameters such as Kp, B_z, and Dst provide a solid foundation for discussing and adapting the general theory of geomagnetic storms.

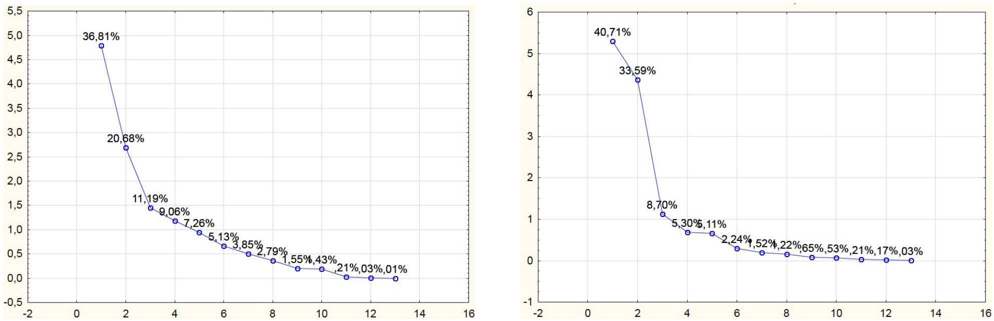


Figure 1: Scree tests of the number of principal components for data parameters of geomagnetic storm 12 Semptember 2014, 3 days before $Kp \geq 5$ (left), 3 days during the storm (right)

Notably, throughout the 2010-2021 timeframe, we identified particular events characterized by $B_z < -10nT$ and $Dst < -100nT$.

In line with our research methodology, we categorized all parameters into two groups, demarcating the periods before and after geomagnetic storms: before the time when $Kp \geq 5$ and afterward. This approach enabled us to study the dynamic structure of the vectors of 13 parameters across two separate datasets. We employed Discriminant Analysis, an effective method for classifying data, in accordance with [1]. Additionally, we explored Pattern Recognition and Neural Networks, which have gained popularity in contemporary research. Cluster Analysis was employed to identify relationships and similarities among the variable sets, producing graphical results that present both interesting and challenging insights.

Despite the myriad of known data processing methods, there continues to be a quest for techniques that can effectively group, classify, and elucidate large data sets. In this study, we focus on the behaviors of geomagnetic and heliospheric parameters that describe geomagnetic storms, applying one machine learning technique, Principal Component Analysis (PCA) [2], [6]. This method utilizes a covariance or correlation matrix based on the data values of the parameters. When the data values exhibit a similar order of magnitude, we employ the covariance matrix; otherwise, we utilize the correlation matrix. In general, we can write

$$\begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} \tag{1}$$

The notation of eq.(1) means that a set of variables $\{X_1, X_2, \dots, X_n\}$ is converted to a vector having new orthogonal variables $(Z_1, Z_2, \dots, Z_n)$, whereby $\sum_{j=1}^{n} a_{nj}^2 = 1$, and $\sum_{j=1}^{n} a_{mj}a_{nj} = 0$, $m \neq n$. In what follows, $X_j = (x_{ij})^T$, for $i = 1, \dots, N$, where x_{ij} are the primary data of X_j, thus $S = (s_{jk})$, $1 \leq j, k \leq n$ expresses the covariance matrix of X_j

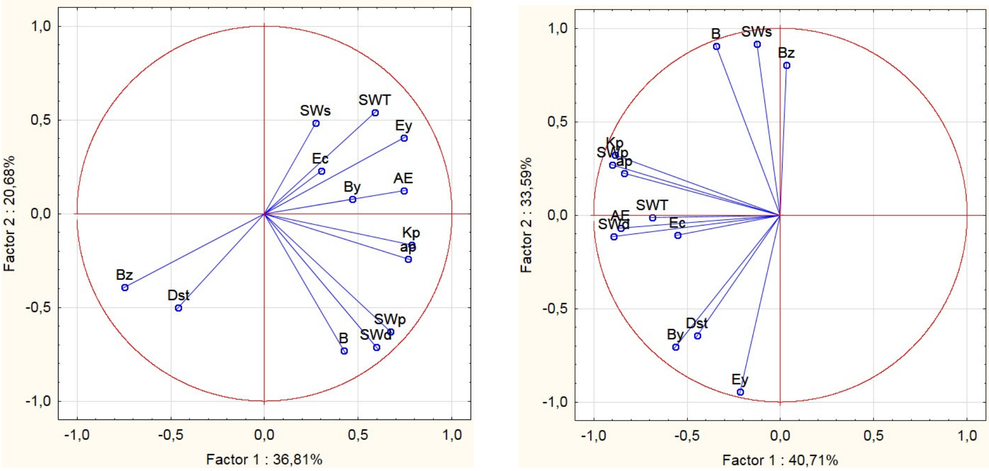


Figure 2: Projection in the plane OZ_1Z_2 and arrangement of parameters for geomagnetic storm 12 Semptember 2014, 3 days before $Kp \geq 5$ (left), 3 days during the storm (right)

for

$$s_{jk} = \frac{1}{N-1} \sum_{i=1}^{N} \left(x_{ij} - \bar{X}_j\right) \left(x_{ik} - \bar{X}_k\right) \tag{2}$$

$$\bar{X}_j = \frac{1}{N} \sum_{i=1}^{N} x_{ij}. \tag{3}$$

According to formula (1) we have a vector $\alpha_j = (a_{j1}, a_{j2}, \ldots, a_{jn})$, $1 \leq j \leq n$ which is the solution of the characteristic equation

$$|(S - \lambda I)\, \alpha_j| = 0. \tag{4}$$

Here, I means identity matrix, λ-eigenvalue. In the case where some eigenvalues λ_j are equal, the choice of the corresponding eigenvectors is not unique. Thus, we compute

$$\sum_{j=1}^{n} Var\left(X_j\right) = tr\left(S\right) = \sum_{j=1}^{n} \lambda_j = \sum_{j=1}^{n} Var\left(Z_j\right). \tag{5}$$

Under the above equation, we obtain

$$\frac{Var\left(Z_j\right)}{\sum_{j=1}^{n} Var\left(X_j\right)} = \frac{\lambda_j}{\lambda_1 + \lambda_1 + \ldots + \lambda_n}. \tag{6}$$

In PCA, as mentioned above, we can study a correlation matrix instead of a covariance one without loss of generality by using the standardisation process. For the general case, the correlation matrix will have the form $\begin{pmatrix} 1 & r\left(X_k, X_j\right) \\ r\left(X_j, X_k\right) & 1 \end{pmatrix}$ where $r\left(X_k, X_j\right) =$

$r(X_j, X_k)$ is the correlation coefficient between X_j and X_k.
Continuation and complement to the PCA method is Factor Analysis, which is a Multivariate Exploratory Technique used to detect the structure and general regularities in the relationships between variables, reducing the number of variables, verifying the detected relationships, describing and classifying the studied variables in new spaces defined by new factors. In this work, factor analysis is not the purpose of our study.

2 Visualisations

An essential aspect of this study is the clear and effective presentation of data. Well-designed visualizations—such as informative charts, diagrams, and images—are crucial for conveying key insights and enabling meaningful interpretation. Data should be presented in both tabular and graphical forms: while tables provide precise numerical values, graphical representations allow readers to quickly perceive patterns, identify anomalies, and form an intuitive understanding of the problem. Such visual insight often facilitates the construction of accurate models.

Selecting an appropriate visualization method requires careful consideration of its suitability and limitations. Principal Component Analysis (PCA) is particularly valuable in this context, as it offers both analytical rigor and a strong graphical component, making it well-suited for exploring complex, multidimensional datasets. As shown in Eq. (1), we begin with n variables X_j and transform them into n new variables Z_j through computation. While these transformed variables may be more complex, the aim is to reduce their number while preserving most of the original variability. From Eq. (6) it follows that each successive principal component explains a smaller proportion of the variance. For further theoretical details, see [2].

Three commonly used criteria for determining the appropriate number of principal components are:

- Variance criterion: retain components explaining at least 5% of total variance.
- Kaiser's criterion: retain components with $\lambda_j \geq 1$, $j = 1, \ldots, n$.
- Cattell's criterion, also known as the *scree test*.

Figures 1 and 2 present the PCA results for the geomagnetic storm of 12 September 2014. Figure 1 displays scree plots for the data parameters, comparing three days before the event with three days during the storm ($K_p \geq 5$). The plots indicate a clear "elbow" at the third component, suggesting that three principal components are sufficient to capture the majority of the data's variance. Figure 2 displays the biplots—projections in the OZ_1Z_2 plane—illustrating the arrangement of parameters over the same periods. Figures 1 and 2 together provide the PCA results for the geomagnetic storm of 12 September 2014, highlighting both variance distribution and the spatial arrangement of parameters in the principal component space. Fig. 2 shows that the relationships between individual parameters before and during the storm are subject to reordering. During the storm, the parameters are clearly organized into three groups, while before they are somewhat more dispersed.

To enhance this analysis, we utilize an Icon Plot technique known as Chernoff Faces [4], which encodes multivariate data into facial features. In this approach, each observation is

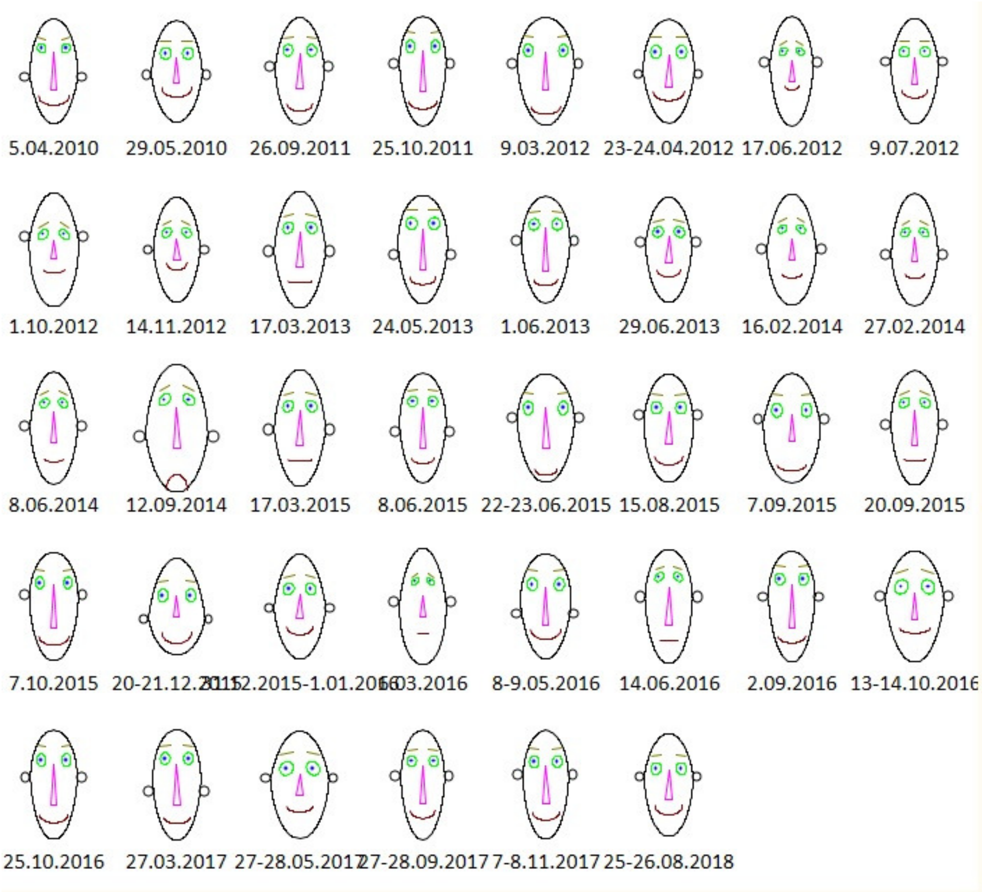


Figure 3: Chernoff Faces showing means of parameters describing 38 geomagnetic storms, in the time when $Kp \geq 5$, in period 2010-2021, in Solar Cycle 24. Meaning: face/width=B, ear/level=B_y, halfface/height=B_z, upface/eccentricity=SWT, loface/eccentricity=SWd, nose/length=SWs, mouth/centering=SWp, mouth/curve=E_y, mouth/length=Kp, eyes/height=Dst, eyes/separation=ap, eyes/slant=AE, eyes/eccentricity=E_c

represented by a unique "face," where the shape and size of features such as nose length, eyebrow angle, and face width correspond to the relative values of selected variables [7]. This visualization enables patterns and groupings to emerge naturally through human perception. Figure 3 displays the aggregated characteristics of 13 parameters during geomagnetic storms, calculated as averages over 72 data points.

The Kp, B_z, and Dst indices are fundamental to characterising geomagnetic storms. $\boldsymbol{Kp}$ measures global activity, $\boldsymbol{B_z}$ (the north-south component of the interplanetary magnetic field) is a key storm driver when negative (southward), and $\boldsymbol{Dst}$ quantifies the strength of the ring current, with more negative values indicating stronger storms. Using these three core parameters, we re-evaluate the storm groups. In our Chernoff faces visualisation, **mouth length** represents Kp, **halfface height** represents B_z, and **eye height** represents Dst.

- **Group 1: The Strongest Storms (High $\boldsymbol{Kp}$, Negative $\boldsymbol{B_z}$, Very Low $\boldsymbol{Dst}$)**
  This group is characterised by faces with a long mouth, a short “half-face” height, and low-set eyes. This combination visually represents an intense geomagnetic storm driven by a very high Kp index, a prolonged and significant southward B_z component, and a deep minimum in the Dst index. This aligns perfectly with the accepted physical conditions for major geomagnetic storms. The storms on September 12, 2014, and September 7, 2015, are key examples. These two faces are visually the most distinct and similar to each other, clearly demonstrating their extreme nature.
- **Group 2: Strong to Moderate Storms with Variable $\boldsymbol{B_z}$ and $\boldsymbol{Dst}$**
  This is a large and diverse group. The faces are unified by a significant mouth length, indicating Kp values at or above the storm threshold of 5, but they show a wide variety of halfface heights (B_z) and eye heights (Dst). Some faces may have an upward-curving mouth (positive E_y), while others have a straight mouth ($E_y \approx 0$), further differentiating them. This group contains storms of varying intensity, where other parameters like solar wind speed and pressure may have played a more significant role. The storms in this group are those not classified in Group 1 or Group 3.
- **Group 3: Weaker Storms ($\boldsymbol{Kp}$ at Threshold, Positive $\boldsymbol{B_z}$)**
  Faces in this group typically have a relatively short mouth, a high “halfface” height, and eyes that are not particularly low-set. This combination suggests a storm at the lower end of the intensity scale for this dataset. This group represents the “milder” storms, where a significant southward B_z was not the dominant factor. The storms from September 26, 2011, through June 8, 2014, are potential members of this category. While some of these faces have a straight mouth, a deeper analysis would be needed to precisely confirm their B_z and Dst values from the visual representation.

Adding Solar Wind Temperature to our classification provides an additional layer of information that helps us physically interpret the storms. We found that the most intense storms, which are represented by faces with long mouths and a "frowning" expression, strongly correlate with low solar wind temperature. This suggests that magnetic clouds primarily caused these extreme events.

In contrast, the group of storms with a positive Ey parameter, or "smiling" faces, is much more diverse, containing both high and low-temperature cases. This indicates that their origins and driving mechanisms were more varied. The Chernoff faces method visually reveals these correlations, allowing for the quick identification of key relationships between different parameters without the need for detailed numerical analysis.

3 Summary

A machine learning–based framework was developed to investigate geomagnetic storms during Solar Cycle 24 (2010–2021). PCA was applied to a 13-dimensional dataset, uncovering latent structures and correlations in storm dynamics. The results show that three principal components are sufficient to capture the majority of the data's variance.

To improve interpretability, we employed Chernoff Faces—a perception-driven visual analytics method—mapping key parameters (Kp index, B_z component, and Dst index) to distinct facial features, enabling intuitive clustering of storms by their characteristics.

The analysis revealed a strong association between the most intense storms and low solar wind temperatures, pointing to a likely link with magnetic cloud events. These findings demonstrate how the integration of dimensionality reduction, pattern recognition, and human-centered visualization can accelerate knowledge discovery in space weather research, effectively bridging computational intelligence with physical interpretation.

Acknowledgements The authors acknowledge NASA and the OMNIWeb team (`omniweb.gsfc.nasa.gov`) for providing solar, heliospheric, and geomagnetic data. Geoelectric field was computed based on the geomagnetic field data from Central Geophysical Observatory in Belsk.

References

1. Krzyśko, M.: Analiza dyskryminacyjna. WNT, Warszawa (1990)
2. Jolliffe, I.T.: Principal component analysis. 2nd edn., New York (2002)
3. Jankowski, J. and Sucksdorff, C. and International Association of Geomagnetism and Aeronomy: IAGA Guide for Magnetic Measurements and Observatory Practice, International Association of Geomagnetism and Aeronomy (1996)
4. `https://www.statsoft.pl/rozwiazania/nauka/internetowy-podrecznik-statystyki/`
5. Wawrzaszek, A., Gil, A., Modzelewska, R., Tsurutani, B.T., Wawrzaszek, R.: Analysis of Large Geomagnetically Induced Currents During the 7–8 September 2017 Storm: Geoelectric Field Mapping 21, 3, e2022SW003383 (2023)
6. Morozova A., and Rebbah R., e: Principal component analysis as a tool to extract Sq variation from the geomagnetic field observations: conditions of applicability, (2022)
7. Nagi, F.M., Mohamed Al-Fardi M., Visualizing Language:A Chernoff Face Approach for Interpreting Embeddings From Online Mental Health Posts, (2025)

Mieczysław A. Kłopotek
ORCID: 0000-0003-4685-7045

Sławomir T. Wierzchoń
ORCID: 0000-0001-8860-392X

Bartłomiej Starosta
ORCID: 0000-0002-5554-4596

Piotr Borkowski
ORCID: 0000-0001-9188-5147

Dariusz Czerski
ORCID: 0000-0002-3013-3483

Institute of Computer Science of Polish Academy of Sciences
ul. Jana Kazimierza 5, 01-248 Warszawa, Poland

{klopotek, stw, barstar, dcz, p.borkowski}@ipipan.waw.pl

On Applicability of Explanation Evaluation Methods for Clustering Results in Textual Domain

Abstract. A new method of evaluating the quality of explanations for the clustering process based on Graph Spectral Clustering methodology is proposed due to shortcomings of current approaches to evaluation of explanations.

Keywords: Explainable AI, Graph Spectral Clustering, Evaluation of Explanations.

1 Introduction

Artificial Intelligence methods gained importance over the last decades because they are able to provide more and more accurate answers. This is accompanied by the new developments in the domain of Large Language Models (LLM). However, the improved accuracy is achieved increasingly at the expense of model complexity which become non-comprehensible, turning in fact to black-box models. As vital decisions are going to be based on AI results, like those in business, military or health care, there exists an urgent need to find ways to convince the end users that AI suggestions are rational.

This problem led to the development of so-called Explainable AI (XAI) that would present reasons for the decision to the end user. This solution gives rise to the next issue: how to convince that the explanations are valid. Therefore, there is a surge in efforts to create methods of Quality of Evaluation of ML Explanation (QEMLE).

There exist multiple works ([1, 2, 3, 12, 11]) and surveys on assessing the quality of explanations ([7, 18, 14, 5]). Most quality evaluation methods require either direct interaction with human user or some predefined sets with "ground truth" explanations.

We shall argue in this paper that in case of document clustering methods a different pathway is required: an analytical justification why the explanations are trustworthy.

But let us first expose the way how the issues have been approached so far.

As summarized by [18], the goal of QEMLE is either to assess either interpretability or fidelity of the explanation method or both. Hereby these terms are to understood as follows [8]:

– interpretability - clarity, parsimony, and broadness
– fidelity - completeness and soundness

According to [9], the QEMLE can be divided into three categories:

1. Application-grounded evaluation
2. Human-grounded evaluation
3. Functionality-grounded evaluation

Application-grounded evaluation requires experiments with end-users. For predefined complex tasks, reflecting the real-world application, one seeks to answer the question how well explanations support domain experts trying to complete these tasks.

Human-grounded evaluation requires experiments with lay humans. Simpler human–subject experiments are conducted that reflect the essence of the target application, but are not that complex. Engagement of lay humans lowers the costs and allows for experiments with a broader set of subjects. This evaluation approach is claimed to depend only on the quality of the explanation, regardless of the types of explanations and the accuracy of the associated prediction.

Functionality-grounded evaluation does not require human experiments. Instead, a comparison is made with some approximation of the formal definition of interpretability, for example the depth of the decision tree being outputted as explanation.

There exist both qualitative and quantitative metrics to evaluate explanations for human experiments. Qualitative metrics include asking about the usefulness of, satisfaction with, confidence in, and trust in provided explanations by means of interviews or questionnaires [15, 4, 19, 17]. Quantitative metrics include measuring human-machine task performance in terms of accuracy, response time needed, likelihood to deviate, or ability to detect errors [15, 10] and physiological responses from humans during experimental tasks [16].

In the last category, functionality-grounded evaluation, the evaluations of explanations depend on the types of explanation which themselves can be divided into: model-based explanations (e.g. decision trees), attribution based explanations (e.g. the ranked lists of attributes contributing to the result) and example-based explanations (e.g. selecting instances that are well predicted or not well predicted by the model as explanations). Respective evalution metrics are summarized in the paper [18].

Issues in Evaluating Explanations for Clustering of Textual Documents

The evaluation methodologies mentioned in the literature are not well suited for evaluating explanations for Graph Spectral Clustering methods [6] applied to textual documents. It has to be stated that these clustering methods until recently were considered black-box methods. Only recently an explanation method was elaborated for these methods [13]. Hence evaluation of such explanations remains a hot topic. Neither Human-grounded evaluation nor Application-grounded evaluation are feasible as the explanation methodology still waits for being incorporated in widely used (clustering) systems. Hence, no human subject experiments are feasible.

Methods of Functionality-grounded evaluation are not applicable. Investigations of explanation complexity are pointless because the explanations are simplistic in nature: a simple list of ranked top words, characterizing a cluster. Tests of leaving out some words from the documents are pointless due to the nature of natural language here the same concept can be expressed by different words. Stability tests are also not very helpful because of the nature of the underlying classical clustering algorithm (k-means). k-means seeks in a stochastic way the local minimum of the loss function and therefore clustering results may (slightly) differ from run to run. Also the GSC methods are in fact approximations of graph cut methods constituting another source of instability between runs. Other methods from this area are not applicable either.

2 A Proposal of Evaluation Method for Explanations for Clustering of Textual Documents

The mentioned problems urged us to develop a new approach to evaluation of explainability, to be presented in the paper, which may be called assumptions-based. The validity of explanations relies on a chain of nearly equivalence relations (see [13]). So, the quality of explanations relies in fact on the degree of fitness to the various assumptions (e.g. the cluster balance, impacting both the explanation process and the clustering process). The new methodology means then defining deviation degrees of individual steps and their combination.

Hence, we can distinguish here the following types of quality measures:

- evaluation of the final result of explanation process
- evaluation of fitness of the assumptions of individual explanation steps to the actual data
- evaluation of the final result of explanation process versus the "ideal" process
- evaluation of sensitivity of the explanation to the leave-one-out testing.

2.1 Evaluation of the final result of explanation process

After the explanation process, each cluster C_i is assigned an ordered list of words L_i where the ordering depends on the weight of a given word for the cluster center. The simplest statistics is the presence of explanation words in the cluster documents which can be summarized in different ways, like

- the positions of the words from the documents from cluster C_i on the explanation list L_i of a given cluster; in particular one seeks the word from a document d_j that has the highest rank $r_i(d_j)$ on L_i; one then computes either the average of $r_i(d_j)$ over all documents $d_j \in C_i$, or computes the histogram, or some other statistics;
- the positions of the words from the documents on the explanation list of a given cluster $r_i(d_j)$ versus presence in the other clusters $r_{i'}(d_j)$, $i' \neq i$; one would expect that $r_i(d_j)$ is numerically lower than $r_{i'}(d_j)$; one would then compute averages, histograms etc. of the difference $r_{i'}(d_j) - r_i(d_j)$ or of $sgn(r_{i'}(d_j) - r_i(d_j))$

2.2 Evaluation of the final result of explanation process versus the "ideal" process

We take two explanations, one based on the result of the clustering and another based on intrinsic clustering and computes measures mentioned above for each of them. Then one compares how much do these measures disagree between the two clusterings.

2.3 Evaluation of fitness of the assumptions of individual explanation steps to the actual data

This approach relies on details of the clustering process and include (in case of the clustering methods described in [13]:

- comparison of the actual k-means fitness measure against the ideal one (best result over a long series of clusterings)
- verification of the assumption of how much the cluster sizes differ from uniform clusters
- histograms of similarities to the one's own cluster center versus other (closest) cluster centers

2.4 Evaluation of sensitivity of the explanation to the leave-one-out testing

Generally, leave-one-out testing is deemed to run comparisons when dismissing single documents from the data. Our approach is, however, to leave entire cluster out and perform a clustering into a lower number of clusters. The expectation is that ranks of words shall not differ too much between the original clustering and the clustering in reduced document set.

2.5 Experimental setup for testing the evaluation methods

The team has collected over years tweets which will be the basis for the verification process. We define an intrinsic cluster as the set of documents that has one single hashtag assigned to them. Documents with two or more hashtags are dismissed from consideration.

Sets of 5,10,15 and 20 intrinsic clusters from cluster pool will be randomly selected. Then the clustering and explanation process will be performed. Subsequently, the evaluations, described in sections 2.1-2.4, will be performed.

References

1. Balagopalan, A., Zhang, H., Hamidieh, K., Hartvigsen, T., Rudzicz, F., Ghassemi, M.: The road to explainability is paved with bias: Measuring the fairness of explanations. In: 2022 ACM Conference on Fairness Accountability and Transparency. p. 1194–1206. FAccT '22, ACM (Jun 2022). https://doi.org/10.1145/3531146.3533179, http://dx.doi.org/10.1145/3531146.3533179
2. Hoffman, R.R., Mueller, S.T., cand Jordan Litman, G.K.: Measures for explainable ai: Explanation goodness, user satisfaction, mental models, curiosity, trust, and human-ai performance. Front. Comput. Sci., Sec. Theoretical Computer Science, 06 February 2023 **5** (2023). https://doi.org/https://doi.org/10.3389/fcomp.2023.1096257, https://www.frontiersin.org/journals/computer-science/articles/10.3389/fcomp.2023.1096257/full
3. Holzinger, A., Carrington, A.M., Müller, H.: Measuring the quality of explanations: The system causability scale (SCS). comparing human and machine explanations. CoRR **abs/1912.09024** (2019), http://arxiv.org/abs/1912.09024
4. Huysmans, J., Dejaeger, K., Mues, C., Vanthienen, J., Baesens, B.: An empirical evaluation of the comprehensibility of decision table, tree and rule based predictive models. Decis. Support Syst. **51**(1), 141–154 (Apr 2011). https://doi.org/10.1016/j.dss.2010.12.003, https://doi.org/10.1016/j.dss.2010.12.003
5. Linardatos, P., Papastefanopoulos, V., Kotsiantis, S.: Explainable ai: A review of machine learning interpretability methods. Entropy **23**(1) (2021). https://doi.org/10.3390/e23010018, https://www.mdpi.com/1099-4300/23/1/18
6. von Luxburg, U.: A tutorial on spectral clustering. Statistics and Computing **17**(4), 395–416 (2007)
7. Löfström, H., Hammar, K., Johansson, U.: A Meta Survey of Quality Evaluation Criteria in Explanation Methods, p. 55–63. Springer International Publishing (2022). https://doi.org/10.1007/978-3-031-07481-3_7, http://dx.doi.org/10.1007/978-3-031-07481-3_7
8. Markus, A.F., Kors, J.A., Rijnbeek, P.R.: The role of explainability in creating trustworthy artificial intelligence for health care: a comprehensive survey of the terminology, design choices, and evaluation strategies. CoRR **abs/2007.15911** (2020), https://arxiv.org/abs/2007.15911
9. Michel, A.H.: The black box, unlocked: Predictability and understand-ability in military ai; (2020), https://unidir.org/files/2020-09/BlackBoxUnlocked.pdf
10. Poursabzi-Sangdeh, F., Goldstein, D.G., Hofman, J.M., Vaughan, J.W., Wallach, H.M.: Manipulating and measuring model interpretability. CoRR **abs/1802.07810** (2018), http://arxiv.org/abs/1802.07810
11. Rudin, C.: Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead (2019), https://arxiv.org/abs/1811.10154
12. Sovrano, F., Vitali, F.: An objective metric for explainable AI: how and why to estimate the degree of explainability. CoRR **abs/2109.05327** (2021), https://arxiv.org/abs/2109.05327
13. Starosta, B., Kłopotek, M.A., Wierzchoń, S.T., Czerski, D., Sydow, M., Borkowski, P.: Explainable graph spectral clustering of text documents. PLoS One **20(2):e0313238** (Feb 2025). https://doi.org/10.1371/journal.pone.0313238
14. Tamajka, M.: How to measure the quality of explanations of ai predictions (2022), https://kinit.sk/how-to-measure-the-quality-of-explanations-of-ai-predictions/
15. Zhou, J., Chen, F.E.: Human and Machine Learning: Visible, Explainable, Trustworthy and Transparent. ; Human–Computer Interaction, Springer: Berlin/Heidelberg, Germany (2018)
16. Zhou, J., Sun, J., Chen, F., Wang, Y., Taib, R., Khawaji, A., Li, Z.: Measurable decision making with gsr and pupillary analysis for intelligent user interface. ACM Trans. Comput.-Hum. Interact. **21** (2015)
17. Zhou, J., Arshad, S.Z., Yu, K., Chen, F.: Correlation for user confidence in predictive decision making. In: Proceedings of the 28th Australian Conference on Computer-Human Interaction. p.

252–256. OzCHI ’16, Association for Computing Machinery, New York, NY, USA (2016). https://doi.org/10.1145/3010915.3011004, https://doi.org/10.1145/3010915.3011004
18. Zhou, J., Gandomi, A.H., Chen, F., Holzinger, A.: Evaluating the quality of machine learning explanations: A survey on methods and metrics. Electronics **10**(5) (2021). https://doi.org/10.3390/electronics10050593, https://www.mdpi.com/2079-9292/10/5/593
19. Zhou, J., Li, Z., Hu, H., Yu, K., Chen, F., Li, Z., Wang, Y.: Effects of influence on user trust in predictive decision making. In: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems. p. 1–6. CHI EA ’19, Association for Computing Machinery, New York, NY, USA (2019). https://doi.org/10.1145/3290607.3312962, https://doi.org/10.1145/3290607.3312962

Stanisław Ambroszkiewicz
ORCID: 0000-0002-8478-6703

University of Siedlce
Faculty of Exact and Natural Sciences
Institute of Computer Science
ul. 3 Maja 54, 08-110 Siedlce, Poland

stanislaw.ambroszkiewicz@uws.edu.pl

Functionals in the Clouds: a Cloud-Native Function-Level Programming

Abstract. Cloud Native Application CNApp (as a distributed system) is a collection of independent components (micro-services) interacting via communication protocols. This gives rise to present an abstract architecture of CNApp as dynamically reconfigurable acyclic directed multi graph where vertices are microservices, and edges are the protocols. Generic mechanisms for such reconfigurations evidently correspond to higher-level functions (functionals). This implies also internal abstract architecture of microservice as a collection of event-triggered serverless functions (including functions implementing the protocols) that are composed into event-dependent data-flow graphs, and dynamically reconfigured at the runtime. Again, generic mechanisms for such compositions and reconfigurations correspond to functionals and higher order type theory like Coq [1] Our contribution is strictly theoretical and relies on the abstract architecture of CNApp that is closely related to the calculus of functionals and relations.

Keywords: Cloud-Native Applications, serverless FaaS and BaaS, abstract architecture, higher order type theory - calculus of functionals and relations

1 Introduction

Contemporary *mainframes* are huge data centers comprising tens or hundreds of thousands of interconnected physical servers and disk arrays. Usually, they are built according to the cloud architecture where almost everything is virtualized starting with IaaS, PaaS and SaaS. The next virtualizations are FaaS (Function as a Service) and BaaS (Backend as a Service) that constitute a new computing paradigm called *Serverless*. It is based on an execution model in which an application consists of interrelated individual functions and backend services that the cloud provider can manage, scale and execute through automatic and dynamic allocation of computing resources. Developers (of the applications) are charged for—only the compute resources and storage needed to execute a particular piece of code.

Using FaaS and BaaS, developers can effectively implement micro-services that comprise the applications. Cloud providers are responsible for deploying the code, and for application runtime environment.

There are many serverless platforms provided by main players like Amazon Web Services, Microsoft Azure, Google Cloud Platform, IBM Cloud, Alibaba, as well as many others. They are trying to persuade developers do create applications mainly out of already existing and proprietary (more or less universal) functions and backend services by composing and orchestrating them into workflows. This gives rise to a new high-level programming paradigm, e.g. Schleier-Smith et al. 2021 [13]:

" This next phase of cloud computing will change the way programmers work as dramatically as the first phase changed how operators work. [...] new general-purpose serverless abstractions will emerge ..."

The physical servers are based on the von Neumann architecture. The programming languages (for developing functions and backend services) are still of von Neumann style. However, the overall high-level architecture of cloud systems, and the emerging (serverless) programming style (see for example Castro et al. 2019 [7]) are not von-Neumann. This resembles the idea of function-level programming postulated by John Backus 1978 [4].

The classic functional programming languages (based on lambda calculus, lazy evaluation and term rewriting) like Clojure, Erlang, Scala, Haskell and many others, are of limited use here. The essence of serverless computing is dynamic orchestration of functions and backend services into workflows where the functions are *black boxes* with clearly defined input and output interfaces and functionalities.

We are going to show that abstract calculus of functionals (higher-order functions) and relations (proposed in [3], and then extended in [2]) may help to understand and formalize the emerging serverless programming paradigm. Functionals and relations correspond to generic mechanisms to compose and reconfigure serverless functions into complex and sophisticated workflows. They are abstractions that allow to cope with such complexity.

Our contribution is strictly theoretical and relies on the abstract architecture of CNApp that is closely related to the calculus of functionals and relations. From this point of view, the serverless computing paradigm may help to realize the idea of function-level programming postulated by John Backus 1978 [4].

The von Neumann computer architecture is still the basic abstract model of computation in computers.

Most of the current programming languages are high-level abstract isomorphic copies of von Neumann computer architectures. There is a vicious cycle caused by this isomorphism. Non-von Neumann computer architectures cannot be developed because of the lack of widely available and effective non-von Neumann languages. New languages cannot be created because of lack of conceptual foundations for non-von Neumann architectures.

Even though Backus stated this almost 50 years ago, it is still a fundamental truth in IT.

Let us stress again, that the overall high-level architecture of cloud systems, and the emerging (serverless) programming style are not von-Neumann.

There are also interesting speculations about the near future of serverless programming like The Serverless SuperComputer proposed by AWS Lambda creator Tim Wagner in 2019, and The Serverless End Game by Garcia et al. 2021 [8] predicting virtualization and transparency of computing resources eventually enabling unlimited flexible scaling.

Some related work on workflows and FaaS are as follows.

HyperFlow, Balis 2016 [5], provides an interesting (although explicitly not based on FaaS) model of computation for workflows. The abstraction there is still at the software level. Hyperflow by Malawski et al. 2020 [11] was applied to serverless execution of scientific workflows.

Triggerflow, see Lopez et al. 2020 [10], is an interesting framework for composing event-based services. Although a formal model of workfow is presented as a Finite State Machine, its architecture and realization are done on the software level.

Bocci et al. 2021 [6] presented an overview of the existing literature on FaaS orchestrations, and their executions environments.

Ristov at al. 2021 [12] proposed an abstract function choreography language (AFCL) for serverless workflow specification. The authors claim that AFCL is at a high-level of abstraction. Its basic constructs correspond to the classic programming control-flow and data-flow expressions, and are: if-then-else, switch, sequence, for, while, parallel, and parallelFor. They are not enough for all possible dynamic (at the runtime) reconfigurations of data-flows, and to cope with complexity if the number of functions is large.

All Cloud providers are now offering cloud orchestration and function composition services, for example, IBM Composer, Amazon Step Functions, Amazon Express Workfows, Azure Durable Functions, and Google Cloud Composer. All of them are essentially at software level, and do not provide higher-order abstractions to cope with complexity and dynamic reconfigurations.

The necessity of such abstraction is the main idea of Open Application Model (`https://oam.dev/`) for defining cloud native apps. It is focused on application rather than container (Docker) or orchestrator (K8s), and brings modular, extensible, and portable design for modeling application deployment with higher level consistent API.

To summarize the Introduction, let us cite Kounev et al. [9] 2023 (Communications of the ACM): *Even though serverless computing has gained significant attention in industry and academia over the past five years, there is still no consensus about its unique distinguishing characteristics and precise understanding of how these characteristics differ from classical cloud computing.*

Hence, the fundamental studies (like the one presented in this paper) are still necessary.

References

1. The Coq Proof Assistant (2022), `https://coq.inria.fr/about-coq`
2. Ambroszkiewicz, S.: Functionals and hardware (2015), `https://arxiv.org/abs/1501.03043`
3. Ambroszkiewicz, S.: Primitive recursion on higher types. In: 2015 Computer Science and Information Technologies (CSIT). pp. 23–32. Proceedings of IEEE (2015). `https://doi.org/10.1109/CSITechnol.2015.7358244`, `https://doi.org/10.1109/csitechnol.2015.7358244`
4. Backus, J.: Can programming be liberated from the Von Neumann Style?: A functional style and its algebra of programs. Commun. ACM **21**(8), 613–641 (Aug 1978)

5. Balis, B.: Hyperflow: A model of computation, programming approach and enactment engine for complex distributed workflows. Future Generation Computer Systems **55**, 147–162 (2016)
6. Bocci, A., Forti, S., Ferrari, G.L., Brogi, A.: Secure FaaS orchestration in the fog: how far are we? Computing pp. 1–32 (2021)
7. Castro, P., Ishakian, V., Muthusamy, V., Slominski, A.: The rise of serverless computing. Communications of the ACM **62**(12), 44–54 (2019)
8. Garcia Lopez, P., Slominski, A., Behrendt, M., Metzler, B.: Serverless predictions: 2021-2030. arXiv e-prints pp. arXiv–2104 (2021)
9. Kounev, S., Herbst, N., Abad, C.L., Iosup, A., Foster, I., Shenoy, P., Rana, O., Chien, A.A.: Serverless Computing What It Is, and What It Is Not? Communications of the ACM **66**(9), 80–92 (September 2023)
10. López, P.G., Arjona, A., Sampé, J., Slominski, A., Villard, L.: Triggerflow: trigger-based orchestration of serverless workflows. In: Proceedings of the 14th ACM International Conference on Distributed and Event-based Systems. pp. 3–14 (2020)
11. Malawski, M., Gajek, A., Zima, A., Balis, B., Figiela, K.: Serverless execution of scientific workflows: Experiments with hyperflow, aws lambda and google cloud functions. Future Generation Computer Systems **110**, 502–514 (2020). `https://doi.org/https://doi.org/10.1016/j.future.2017.10.029`, `https://www.sciencedirect.com/science/article/pii/S0167739X1730047X`
12. Ristov, S., Pedratscher, S., Fahringer, T.: AFCL: An abstract function choreography language for serverless workflow specification. Future Generation Computer Systems **114**, 368–382 (2021)
13. Schleier-Smith, J., Sreekanti, V., Khandelwal, A., Carreira, J., Yadwadkar, N.J., Popa, R.A., Gonzalez, J.E., Stoica, I., Patterson, D.A.: What serverless computing is and should become: The next phase of cloud computing. Communications of the ACM **64**(5), 76–84 (2021)

Cezary Borkowski

Mirosław Szaban
ORCID: 0000-0002-6918-4675

Franciszek Seredyński
ORCID: 0000-0003-0206-8183

University of Siedlce
Faculty of Exact and Natural Sciences
Institute of Computer Science
ul. 3 Maja 54, 08-110 Siedlce, Poland

cb85253@stud.uws.edu.pl, miroslaw.szaban@uws.edu.pl,
franciszek.seredynski@uws.edu.pl

Solving the Problem of Pattern Formation Using Genetic Algorithm

Abstract. This article presents a method for solving the pattern formation problem using a Genetic Algorithm. The approach focuses on pattern formation in a two-dimensional space, where the evaluation function measures the number of correctly positioned pattern elements within the system. The Genetic Algorithm employs tournament selection, multi-point crossover, and a narrowing mutation strategy. The method demonstrates high effectiveness, as confirmed by the presented experimental results.

Keywords: Collective behavior, Cellular Automata, Formation of a Pattern Distribution, Genetic Algorithm.

1 Introduction

Pattern formation is a phenomenon commonly observed in nature, evident in examples ranging from snowflake structures and animal coat patterns to the spatial organization of cells in biological tissues. Understanding and simulating such processes has become a central focus in fields such as artificial life, complex systems, and distributed computing. From a computational standpoint, pattern formation is often interpreted as a form of self-organization, wherein local interactions among simple components give rise to coherent global structures.

Recent approaches to simulating self-organizing behavior frequently utilize models such as Cellular Automata (CA). Several CA-based methods addressing the pattern formation problem have been presented in [1, 7]. Specifically, the studies in [2, 3, 4, 5] examine the arrangement of dominoes within a cellular grid as a specialized instance of pattern formation.

Our research builds on a recently introduced methodology [6], which incorporates three fundamental components: (a) a multi-agent interpretation of the pattern formation problem; (b) the use of a variant of the Spatial Prisoner's Dilemma (SPD) game to model agent interactions; and (c) the implementation of evolutionary, competing CA-based agents functioning as basic reinforcement learning units.

In our recent work [8, 9], we propose a novel game-theoretic multi-agent system for generating desired 2D patterns. The pattern formation task is formulated as a variant of the iterated Spatial Prisoner's Dilemma, where the agents—modeled as CA-based entities—compete evolutionarily while operating as reinforcement learning mechanisms.

In this paper, we propose a evolutionary approach based on a Genetic Algorithm to address the pattern formation problem in a discrete 2D space. Our GA evolves entire candidate solutions through biologically inspired mechanisms—selection, crossover, and mutation—over successive generations. The fitness is calculated by counting the number of matching elements between the candidate and the reference structure, thereby guiding evolution toward the desired global configuration. Our experimental results demonstrate that the proposed GA-based method can consistently reproduce high-quality target patterns, indicating its potential as a promising direction for further exploration in the domain of such kind of systems.

2 Pattern Formation Problem

This study addresses the problem of pattern formation within a two-dimensional discrete space. The space is defined on an $m \times n$ grid, where each cell can be in one of two binary states: 0 or 1. A desired pattern is represented by a binary matrix, and a valid configuration is characterized by the precise placement of state-1 elements. Specifically, a "correctly placed" 1 is defined as a cell in state 1 whose eight neighboring cells (in the Moore neighborhood) are all in state 0. This local condition provides the foundation for a global evaluation function, which quantifies the quality of a pattern by counting the number of such correctly embedded 1s throughout the grid.

Let X denote a binary matrix representing a candidate solution. The optimization goal is to maximize the following objective function:

$$f(X) = \sum_{i=2}^{m-1} \sum_{j=2}^{n-1} \mathbb{I}\left(X_{i,j} = 1 \land \forall_{(p,q) \in \mathcal{N}_{i,j}} X_{p,q} = 0\right), \tag{1}$$

where $\mathcal{N}_{i,j}$ denotes the Moore neighborhood of cell (i, j), and $\mathbb{I}()$ is the indicator function that returns 1, if the condition is true and 0 otherwise.

This optimization problem is inherently NP-hard due to the exponential growth of possible binary configurations. Traditional approaches often rely on deterministic or rule-based heuristics to explore the solution space. However, these methods frequently face limitations in terms of scalability and convergence, especially with larger problem instances.

Consequently, this work addresses the central research question: Can a Genetic Algorithm (GA) efficiently generate globally coherent patterns by maximizing a globally defined evaluation function, starting from a random initial state?

3 Genetic Algorithm for Pattern Formation Problem

To solve the pattern formation problem defined in Section 2, we use a Genetic Algorithm (GA) specifically designed to evolve binary matrices that represent potential solutions.

Each individual in our population is a binary matrix of size $m \times n$, with each cell holding a 0 or a 1. Unlike methods that convert the matrix into a single, one-dimensional string, our approach keeps the two-dimensional structure throughout the entire process. This allows genetic operations like crossover and mutation to be applied directly in the 2D space. By doing this, we can make changes that respect the local spatial relationships, which aligns better with the nature of our pattern formation task.

A solution's fitness is calculated as the ratio of correctly placed 1s to the total number of 1s in the reference matrix. Incorrectly placed 1s, in a position where the reference matrix has a 0, is not adding to the score but are still included in the total count in the denominator. This method ensures that only accurate matches contribute positively to the final fitness score.

Let X be the candidate matrix and R the reference matrix of size $m \times n$. The fitness function is defined as:

$$f(X, R) = \frac{\text{Number of positions where } X_{i,j} = 1 \text{ and } R_{i,j} = 1}{\text{Number of positions where } R_{i,j} = 1 \text{ or } X_{i,j} = 1}. \quad (2)$$

This score is a value between 0 and 1. A value of 1 means that all 1s in the candidate match the reference exactly, with no extra or missing 1s. Lower scores indicate incorrect placements or missing pattern elements.

The algorithm utilizes the following genetic operators:

- Tournament Selection: Individuals are selected for reproduction using binary tournament selection, where two random candidates compete and the one with higher fitness is chosen.
- Multi-point Crossover: N-point crossover is applied to selected parent pairs. N random crossover points are chosen along the chromosome, and the corresponding segment are exchanged between parents to produce offspring.
- Narrowing Mutation: To maintain genetic diversity and avoid premature convergence, a narrowing mutation strategy is used. Initially, the mutation probability is relatively high, allowing broad exploration. Over time, both the mutation probability and the range of mutation gradually decrease, promoting fine-tuning in later generations.

The overall flow of the GA is described as shown in Algorithm 1:

Algorithm 1 Genetic Algorithm for Pattern Formation

1: Initialize population with random binary matrices.
2: Evaluate fitness of all individuals.
3: Repeat for a fixed number of generations is reached or convergence criterion is met:
- Select parents using tournament selection.
- Apply multi-point crossover to produce offspring.
- Apply narrowing mutation to offspring.
- Evaluate new individuals and form the next generation.

4 Experimental Results

To evaluate the performance of our proposed GA for pattern formation, we conducted a series of experiments. Each experiment began with a random binary configuration and ran for a predetermined number of generations, while varying the GA's core parameter settings.

We first performed parameter tuning to identify the most effective values for population size, number of generations, crossover rate, and initial mutation rate.

Using the optimal parameter configuration from the tuning phase, we then executed multiple runs of the algorithm. These repeated experiments allowed us to assess the algorithm's stability, robustness, and overall performance. The results were collected and subjected to statistical analysis.

4.1 Tuning the parameters of GA

To perform the experiments with the genetic algorithm, we defined four key control parameters: N (population size), P_k (crossover probability), P_m (mutation probability), and T (number of generations). We varied each parameter within a specific range using a fixed step, as follows: $N \in [500; 1000]$ with step $= 100$, $P_k \in [0.7; 0.8]$ with step $= 0.05$, $P_m \in [0.004; 0.006]$ with step $= 0.001$ and $T \in [100; 500]$ with step $= 100$.

From that conditions, the most effective configuration for GA we observe for selected parameters values: $N = 1000$, $P_k = 0.70$, $P_m = 0.004$, $T = 200$, where corrected rate of patterns' distribution was: **f_avg_C_corr** $= 0.999$, also $f\_min\_C\_corr = 0.990$ and $f\_max\_C\_corr = 1.0$.

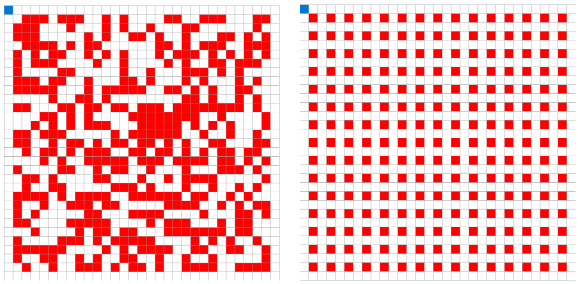
4.2 Analysis of Results

The experiments were conducted using the following configuration: a CA size of $m \times n = 31 \times 31$; a population size of 1000; 200 generations; a crossover rate of 0.7; 8 crossover points; an initial mutation rate of 0.004; and a tournament size of 5, as mentioned earlier.

Based on the parameter tuning results, we selected the best-performing configuration (see, Table 1) to conduct a series of ten independent tests. These tests were designed to evaluate the genetic algorithm's consistency and robustness. As the results show, the GA found the target solution in 9 out of 10 cases. The highest fitness value, which represents the number of correct pattern assignments, was 255 it corresponds to rate equal 1 (see, last column of Table 1).

Table 1: Summary of 10 tests with selected parameters. (Source: own research.)

Test No.	Iteration	f_min_C_corr_N	f_avg_C_corr_N	f_max_C_corr_N
1	Last (Best)	0.965 (0.961)	0.993 (0.988)	0.996 (0.996)
2	Last (Best)	0.965 (0.961)	0.993 (0.988)	0.996 (0.996)
3	Last (Best)	0.969 (0.957)	0.997 (0.989)	1.000 (1.000)
4	Last (Best)	0.969 (0.965)	0.997 (0.993)	1.000 (1.000)
5	Last (Best)	0.969 (0.965)	0.997 (0.993)	1.000 (1.000)
6	Last (Best)	0.965 (0.961)	0.993 (0.988)	0.996 (0.996)
7	Last (Best)	0.957 (0.952)	0.984 (0.979)	0.987 (0.987)
8	Last (Best)	0.969 (0.965)	0.997 (0.993)	1.000 (1.000)
9	Last (Best)	0.965 (0.961)	0.993 (0.988)	0.996 (0.996)
10	Last (Best)	0.969 (0.965)	0.997 (0.993)	1.000 (1.000)



(a) Initial pattern (generation 1) **(b)** Best individual (generation 200)

Figure 1: Example of pattern evolution: from the initial random configuration to the best evolved individual after 200 generations. (Source: own research.)

Figure 1 illustrates the pattern evolution process. The Figure 1 (a) shows the initial, random state of an individual at generation 1. The Figure 1 (b) displays the best individual (and our goal) found after 200 generations, which is the result of the GA's optimization.

Figure 2 shows the progression of the average fitness value across generations, confirming a steady improvement and convergence toward a high-quality solution.

Each test case run on a standard personal computer equipped with: Processor: Intel(R) Core(TM) i7-11800H @ 2.30GHz, RAM: 32 GB. For the reference configuration ($N = 1000$, $T = 200$, $m = n = 31$), was completed in under 8 (avg = 7.39) seconds of wall-clock time.

5 Summary

The presented research confirms that a supervised Genetic Algorithm is a successful tool for solving the pattern formation problem in a discrete 2D space.

A key conclusion from our experiments is that evolutionary mechanisms, when properly configured, can lead to the stable and repeatable emergence of structured patterns. The narrowing mutation strategy was particularly effective, helping to preserve diversity in early generations and ensuring convergence to optimal structures in later stages. This adaptive

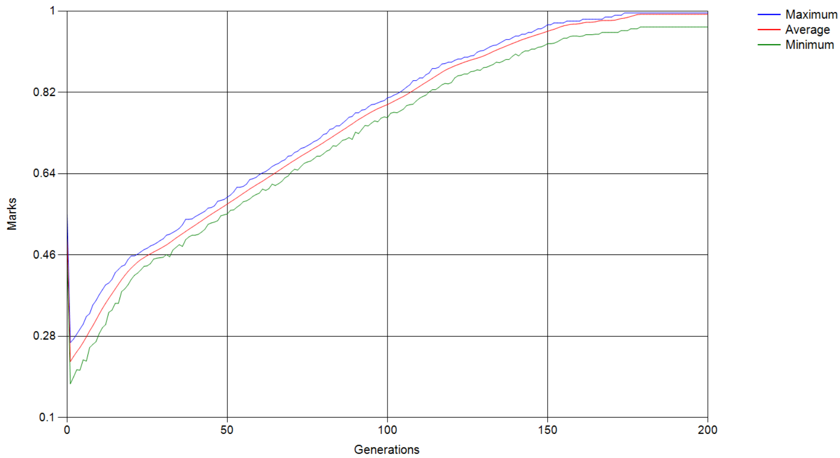


Figure 2: Fitness function values for a GA sample run. (Source: own research.)

balance between exploration and exploitation is crucial for maintaining the algorithm's performance across a wide range of scenarios.

The results demonstrate both high accuracy and robustness. The algorithm repeatedly achieved near-optimal solutions with minimal variance across 10 independent runs, showing strong repeatability and resistance to random fluctuations.

Future directions include hybridizing this approach with agent-based models, generalizing it to multi-class patterns, and exploring its applications in fields such as programmable materials, robotics, and morphogenetic systems.

References

1. Nagpal, R., Programmable pattern-formation and scale-independence, Unifying Themes in Complex Sytems IV, Minai, A.A, Bar-Yam, Y., eds., (2008) 275–282
2. Hoffmann, R., How agents can form a specific pattern, Cellular Automata, Wąs, J., Sirakoulis, G., Bandini, S. eds., LNCS 8751 (2014) 660–669
3. Hoffmann, R., Cellular automata agents form path patterns effectively, *Acta Phys. Pol. B Proc. Suppl.* 9(1) (2016) 63–75
4. Hoffmann, R., Désérable, D., Line patterns formed by cellular automata agents, Cellular Automata, El Yacoubi S., Wąs, J., Bandini, S. eds., LNCS 9863 (2016) 424–434
5. Hoffmann, R., Désérable, D., Seredyński, F., A Cellular Automata Rule Placing a Maximal Number of Dominoes in the Square and Diamond, *The Journal of Supercomputing* 77, 2021, pp. 9069-9087
6. Seredyński, F., Kulpa, T., and Hoffmann, R., Evolutionary Self-optimization of Large CA-based Multi-Agent Systems, *J. of Computational Science* 68 (2023) 101994
7. Yamins, D., Nagpal, R., Automated Global-to-Local programming in 1-D spatial multi-agent systems, Proc. 7th Int. Conf. AAMAS (2008) 615–622
8. Szaban, M., Seredyński, M., Hoffmann, R., Désérable, D., Seredyński, F.: Pattern Formation by Collective Behavior of Competing Cellular Automata-based Agents. LNCS 14978 (2024), 34–44
9. Szaban, M., Seredyński, F.: Arrangement method by the collective behavior of competing cellular automata-based agents. In: 2nd Conference IS&IT, (2024)

Dawid Dziedzic

Mirosław Szaban
ORCID: 0000-0002-6918-4675

University of Siedlce
Faculty of Exact and Natural Sciences
Institute of Computer Science
ul. 3 Maja 54, 08-110 Siedlce, Poland

dawid.dziedzic2359@gmail.com, miroslaw.szaban@uws.edu.pl

Is it Possible to Construct an Artificial Neural Network Using Cellular Automata?

Abstract. This paper presents a concept for employing one-dimensional cellular automata as a tool for constructing artificial neural networks. It introduces the design of an artificial neuron implemented using a converging (narrowing) cellular automaton. Subsequently, a network architecture is proposed by interconnecting multiple instances of the aforementioned neuron model. A training method for this type of network is proposed, based on the appropriate selection of transition rules, and the results of both training and testing are presented. It has been demonstrated that such a construction is feasible and yields promising results; however, further development and refinement of the concept are necessary, along with efforts to optimize the computational efficiency of this type of network.

Keywords: Artificial neural networks, One-dimensional cellular automata, Machine learning, Evolutionary algorithms, Feedforward neural networks.

1 Introduction

The quest for alternative computational paradigms to Artificial Neural Network (ANN) has long been a topic of interest in artificial intelligence and complex systems (lastly, its use in various applications, see [1, 2, 3, 4]. Cellular automata (CA), as discrete dynamical systems characterized by local interaction rules, have been extensively studied for their emergent behaviors, pattern formation capabilities, and potential for parallel computation [5, 6]. Their intrinsic simplicity, combined with the ability to generate complex global phenomena from local interactions, makes CA a promising candidate for neural-like architectures.

The idea of employing CA for neural computation dates back to the early 1980s, notably with Stephen Wolfram's work on computational universality in CA [5]. While Wolfram's 1983 statistical mechanics paper laid foundational groundwork, his 1984 study explicitly linked CA to universal computation, demonstrating how simple rules could emulate complex systems akin to neural processing. Later, Ilachinski [6] reviewed CA models inspired by neural systems, emphasizing their potential for distributed processing and pattern recognition. These approaches often involved designing CA rules to emulate synaptic interactions, drawing inspiration from Hebbian learning principles [7].

Recent research has focused on CA-based neural networks with dynamic training capabilities. Sipper [8] demonstrated how genetic algorithms can evolve CA rules to perform neural-like learning, achieving pattern recognition through emergent attractor states. Similarly, Chua's cellular nonlinear networks (CNNs) [9] formalized a hybrid paradigm combining CA locality with continuous-state neural dynamics. Earlier, Boccara [10] showed that partitioned CA rules could classify patterns by evolving toward category-specific attractors.

Despite these advances, implementing traditional neural architectures in CA remains challenging due to the discrete, local nature of CA rules, which inherently lack explicit weighted connections. Recent approaches bridge this gap by combining CA with evolutionary algorithms [11] or by constructing multi-layer CA systems [12] to realize neural-like processing.

An intriguing concept is presented in Gilpin' work [13], which demonstrates the inverse of the case discussed in this article. He illustrates how a CA can be constructed via a convolutional neural network (CNN).

Our work proposes a CA-based framework for artificial neural networks (ANNs), where CA rules act as neurons within a multilayer architecture. Training involves co-evolving these rules via genetic algorithms (inspired by Sipper [8]) and Hebbian adaptation [7]. This approach leverages CA's decentralized structure to offer inherent parallelism and fault tolerance, facilitates self-organization through rule evolution, and enables neural-like computation without relying on explicit weight matrices, as seen in Chua's CNNs [9].

2 Cellular Automata-based Artificial Neural Network (CA-ANN)

2.1 CA-ANN Architecture

The proposed architecture implements artificial neurons using one-dimensional cellular automata (CA) with rule-based state transitions. Each neuron is modeled as a CA that processes any odd number of bits (neuron input) through iterative rule applications, ultimately converging to a single output bit (see, Figure 1). The behavior of each neuron is governed by an 8-bit rule vector (ranging from 0 to 255), where each specific rule dictates how a cell's next state is determined by its current state and the states of its immediate left and right neighbors.

Figure 1 illustrates the schematic operation of a neuron based on a CA with rule 90. In this model, the neuron (CA) processes binary input sequence by applying CA rule to compute its subsequent states, at last one bit (neuron output).

A CA-ANN layer is a linear structure composed of n neurons, all of which receive the same input sequence. Each neuron processes the data according to its own rule (CA rule) and

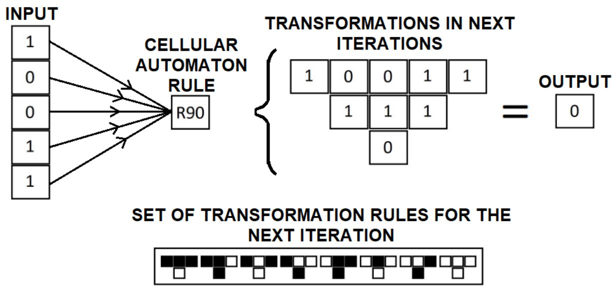


Figure 1: Schema of the neuron operation using cellular automata with rule 90. (Source: own work)

returns a single bit. The resulting sequence of these n bits forms the output sequence. In this way, we can construct deep CA-ANN (see, Figure 2).

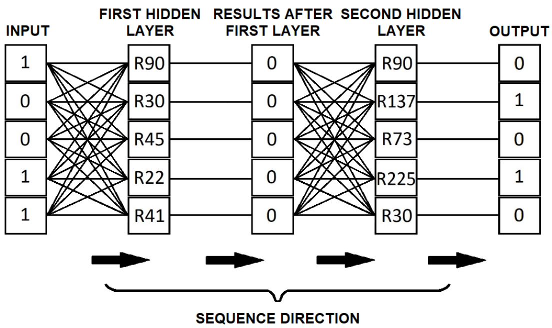


Figure 2: Schema of operation for the cellular automata-based neural network with two hidden layers. (Source: own work)

Within a CA-based CNN, complex transformations are achieved through simple local interactions (see, Figure 2). Each neuron progressively reduces the input sequence length until a single output bit remains, thus effectively implementing a binary decision function.

2.2 CA-ANN Learning Methodology

The training methodology for the cellular automata-based neural network is a two-step process. The first stage is a performance evaluation (or scoring) that analyzes a rule's ability to drive the cellular automaton (CA) to the desired binary state.

Performance Evaluation: The effectiveness of individual rules is assessed based on their capacity to transform input sequences into the desired output patterns accurately. This evaluation involves assigning a performance score to each rule, where successful transformations—those matching the expected output—are awarded positive points, while incorrect transformations incur penalties. Specifically, rules receive +1 point for correct transformations and -1 point for incorrect ones. These scores are accumulated over multiple iterations, enabling continuous monitoring of each rule's contribution to the network's overall performance. The scoring process is conducted during the simulation runs, where the output of the automaton is compared to the target pattern using the Hamming distance as a metric of discrepancy.

In the second stage, a genetic algorithm-inspired optimization framework is used to iteratively refine the transition rules that govern each neuron.

Genetic Algorithm-inspired optimization framework: Rules exhibiting negative performance scores are deemed ineffective and are subject to replacement. Additionally, rules with positive scores may undergo probabilistic modification to promote exploration of the rule space. This replacement and modification process is governed by a strategy that balances exploiting well-performing rules and exploring new configurations. Specifically, rules with negative scores are systematically replaced to prevent stagnation, while positively scoring rules (set at 20% as default) are candidates for replacement by crossover or mutation process. After selection in the replacement process 70% of selected rules are crossovered (classical one-piont crossover) to produce new rule configurations. This operation creates from the top-performing rules to preserve effective patterns and maintain diversity within the rule set. For 30% other rules we apply mutation based on Gaussian distribution with mean value equal to 0 and standard deviation 10.

Termination Conditions: The training process concludes when either of the following criteria is met: (1) the network achieves perfect transformation accuracy, indicated by a Hamming distance of zero between the network output and the expected pattern; or (2) a maximum number of iterations—typically set to 100,000—is reached.

3 Experimental Results

3.1 Experimental Parameters

The evaluation examined a cellular automata-based neural network (CA-ANN) that was realized with the following configuration values: Initial states: 000100010, 001100000, 010001101; Expected outputs: 000010000. Input sequence length: 9 bits; Hidden layers: 5; Rule initialization & Random: (0-255).

3.2 Training Process and Results

The algorithm successfully transform all inputs to specified outputs through evolutionary optimization. In this process was verified the network's ability to learn with additional input patterns using the evolved rule set. Figure 3 demonstrates the scoring mechanism for individual rules, with the top-performing rules (upper part of Figure 3) significantly outperforming ineffective ones (down part of Figure 3). Training shows also percentage chart of best sets of rules transforming initial states to desired expected outputs (Figure 4).

3.3 Testing Performance

The performance of the CA-ANN model was evaluated through a comprehensive testing procedure involving 1,000 test cases. Each of the expected output patterns from the training phase while applying random input sequences, testing the model's ability to produce correct

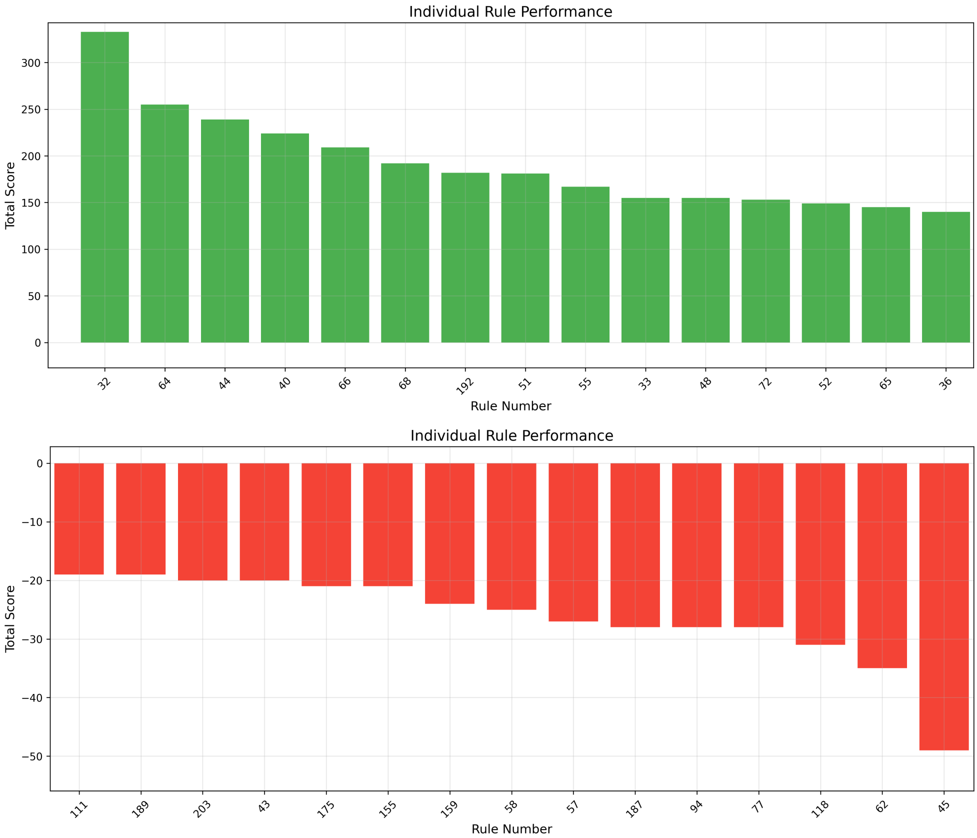


Figure 3: Performance comparison between best (upper part: top 15 performing rules) and worst performing rules (down part: bottom 15 performing rules) (Source: own research)

outputs from novel inputs. The test architecture used a per-file batch processing approach, where each file was iteratively processed to verify that the selected rule set successfully transforms initial states to expected outputs without the need for further training.

The performance of the trained CA-ANN across three different models is summarized in Table 1. The test results were relatively stable and showed convergent values. For the first model, the average performance was over 65%, with a peak of 70%. The subsequent models achieved average performances of nearly 26% and 45%, with their best results reaching approximately 28.7% and 49.6%, respectively. Notably, the minimum performance was not significantly lower than the average in any of the models (see, Table 1).

This pattern-dependent behavior confirms the model's sensitivity to input-output alignment and its capacity for specialized pattern transformation when properly configured.

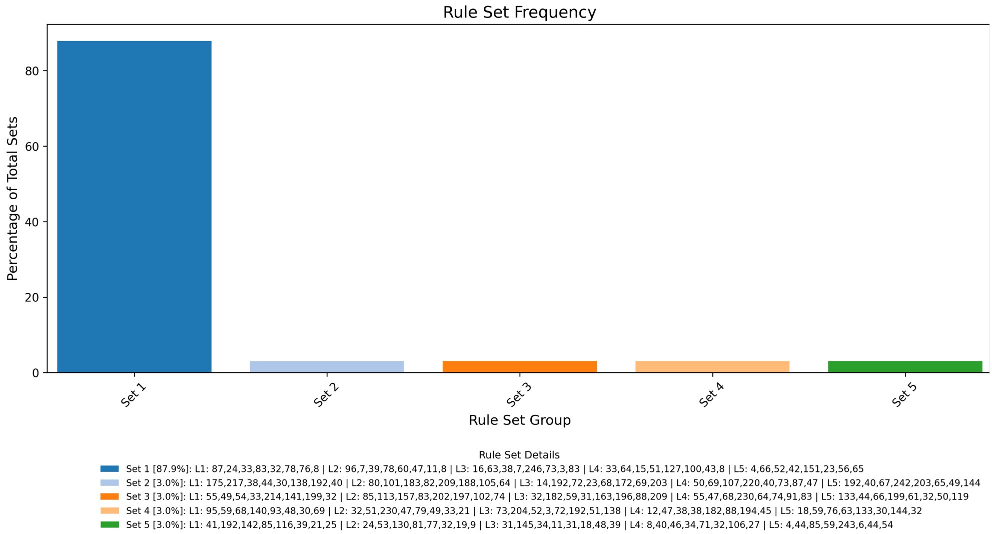


Figure 4: Percentage distribution of the best training sets. (Source: own research)

Table 1: Performance of the trained CA-ANN for three different models, 100 tests per model. (Source: own research)

Performance of CA-ANN	Model 1	Model 2	Model 3
Minimal	61.1%	23.5%	40.7%
Average	65.424%	25.982%	44.77%
Maximal	70%	28.7%	49.6%

4 Conclusions

This paper introduces a new approach to constructing artificial neural networks based on one-dimensional cellular automata (CA-ANN). The core of our network is a neuron created from a tapering one-dimensional cellular automaton. The CA-ANN is structured similarly to a classical ANN, but the traditional neuron is replaced by our proposed CA-based neuron.

Our preliminary experimental results show that CAs can, in fact, form functional artificial neural networks. The results of our tests demonstrated that the proposed CA-ANN can learn and then solve problems represented in binary form.

While this approach shows promise, it requires development in several key areas to obtain more precise scores and allow for its application to a wider range of problems.

References

1. Ruciński, D., Artificial Neural Network based on mathematical models used in quantum computing, Studia Informatica. Systems and Information Technology, 27/2, (2022) 27-48

2. Wawrzyńczak, A., Wysocki, Ł., Kopka,P., Modeling of Continental-Scale Atmospheric Contaminant Dispersion Using Artificial Neural Networks, Proceedings of 2nd Conference (ISIT), (2024)
3. Przychodzki, M., Niewiadomski, A., Application of artificial neural networks and NEAT algorithm to control a quadcopter, Studia Informatica. Systems and Information Technology, 29/2, (2023) 41-56
4. Tchórzewski, J., Rosochacki, S., Selection of the programming environment for neural modelling of the power and electricity demand generation systems in terms of unmanned factories, Studia Informatica. Systems and Information Technology, 29/2, (2023) 95-106
5. Wolfram, S., Universality and Complexity in Cellular Automata, *Physica D*, 10, (1984) 1–35
6. Ilachinski, A., Cellular Automata: A Discrete Universe, World Scientific, (2001)
7. Hebb, D.O., The Organization of Behavior, Wiley, (1949)
8. Sipper, M., Evolution of Parallel Cellular Machines: The Cellular Programming Approach, Springer, Heidelberg, (1997)
9. Chua, L.O., CNN: A Paradigm for Complexity. *Int. J. Bifurcation Chaos*, 31, World Scientific, (1998)
10. Boccara, N., et al., Partitioning CA Rules for Classification Tasks. *Physica D*, (1994)
11. Seredyński, F., Kulpa, T., Hoffmann, R., Evolutionary Self-optimization of CA-based Multi-Agent Systems. *Journal of Computational Science*, 68, (2023)
12. Hoffmann, R., Désérable, D., Path Patterns Formation in Cellular Automata Agents. *Acta Physica Polonica B*, 9(1), (2017) 63–75
13. Gilpin, w., Cellular automata as convolutional neural networks, *Phys. Rev. E*, 100, (2019)

Mathieu Francois[1,2]
ORCID: 0009-0009-7808-771X

Iván Rivero[2]
ORCID: 0009-0006-0466-3065

Cristina Tirnauca[2]
ORCID: 0000-0002-7129-2237

Rafael Duque[2]
ORCID: 0000-0001-8636-3213

[1] ENSEIRB-MATMECA
Bordeaux INP, Avenue du Dr Albert Schweitzer, Talence, France

[2] Departmento de Matemáticas, Estadística y Computación
Universidad de Cantabria, Avda. de los Castros s/n, 39005 Santander, Spain

mathieu.francois@bordeaux-inp.fr, ivan.rivero@unican.es,
cristina.tirnauca@unican.es, rafael.duque@unican.es

An Interactive Platform for Medical Machine Learning Using AutoML and Large Language Models

Abstract. This work introduces an interactive platform that combines Automated Machine Learning and Large Language Models to simplify the application of machine learning in medical research. Designed for healthcare professionals without programming expertise, the system enables users to build, evaluate and interpret predictive models on structured biomedical datasets through a natural language interface. The backend employs AutoGluon for automated model selection and optimization, while SHAP visualizations provide interpretable insights into feature importance. A conversational assistant, powered by Meta's LLaMA-4 model and served via the Groq's application programming interface, guides users throughout the workflow – explaining preprocessing steps, model results and suggesting improvements. The platform supports end-to-end tasks including data upload, training configuration, model deployment and prediction generation, all accessible via a user-friendly web interface. Experimental validation on public biomedical datasets demonstrates the system's ability to produce accurate and interpretable models with minimal user intervention. By lowering technical barriers, this tool empowers domain experts to independently explore machine learning applications, fostering innovation in clinical research and accelerating the integration of artificial intelligence into medical decision-making.

Keywords: AutoML, Large Language Models, Medical Data Analysis.

1 Introduction

Machine Learning (ML) has emerged as a transformative tool in medical research, enabling the extraction of complex patterns from large and heterogeneous datasets. Its applications span from disease diagnosis and prognosis to treatment optimization and healthcare resource management [1]. As biomedical data continues to grow in volume and complexity, ML offers the potential to uncover clinically relevant insights that would be difficult or impossible to detect using traditional statistical approaches.

Despite its promise, one of the main challenges in applying ML in the medical domain lies in the technical barrier faced by healthcare researchers. Most medical professionals and clinical researchers lack formal training in machine learning, programming or data science. As a result, they often rely on collaborations with data scientists or software engineers, which can create bottlenecks, miscommunication and delays in research workflows. This dependency also limits the ability of researchers to iteratively explore hypotheses or adapt models to specific clinical needs.

To address this gap, recent developments in Large Language Models (LLMs) and Automated Machine Learning (AutoML) offer a promising solution. LLMs are advanced deep learning models trained on massive amounts of textual data to understand and generate natural language [2]. Models based on transformer architectures, such as those used in this work, can understand and generate human-like text, enabling intuitive interfaces that significantly lower the barrier to interacting with complex technical systems. AutoML, on the other hand, refers to a suite of techniques that automate the end-to-end process of applying machine learning to real-world problems [3], including data preprocessing, feature selection, model selection and hyperparameter tuning. By reducing the need for manual intervention in the ML workflow, AutoML enables non-experts to build robust and optimized models with minimal technical effort. Together, these technologies can empower domain experts to build and evaluate ML models without extensive programming knowledge.

This work presents a novel system that integrates LLMs and AutoML into a unified web-based platform designed to be accessible to researchers in the medical field. The system allows users to interact with the machine learning pipeline through natural language queries, while AutoML handles the backend model optimization. To demonstrate the feasibility and effectiveness of the proposed approach, we conduct a series of experiments using publicly available biomedical datasets.

2 System Architecture and Implementation

The system enables the construction and interpretation of predictive models from structured biomedical datasets by users without prior training in programming or machine learning. Its design follows four guiding principles: accessibility, automation, explainability and interactivity. Accessibility is provided through a web interface that abstracts implementation details. Automation is handled by an AutoML engine responsible for model selection, training and tuning. Explainability is addressed via visualizations and language-based descriptions. Interactivity is supported through the integration of an LLM that provides contextual assistance during the modeling process.

2.1 Overview of Technologies

The backend of the platform is built using **AutoGluon**, an open-source AutoML framework developed by Amazon that enables users to automatically train and optimize machine learning models with minimal code. It supports tabular data and provides model ensembling, hyperparameter tuning and interpretability features out-of-the-box. The web server is implemented using **Flask**, a lightweight Python framework that facilitates RESTful API development and integration with ML workflows.

To support natural language interaction, the system integrates an LLM –**meta-llama/llama-4-scout-17b-16e-instruct**– served via the **Groq** inference engine. This LLM provides contextual assistance, explanations and guidance throughout the modeling process. While this is the default model, users may also choose other LLMs provided through Groq. For model interpretability, the platform employs **SHAP** (SHapley Additive exPlanations) to generate visualizations that highlight the contribution of each feature to the model's predictions [4].

2.2 System Components

The application follows a **frontend-backend** architecture:

– **Frontend**: Built using HTML, CSS and JavaScript the frontend provides a user-friendly interface for uploading datasets, configuring training parameters, launching model training and visualizing results.
– **Backend**: Implemented in Python using the Flask framework, the backend handles file uploads, model training with AutoGluon, SHAP-based visualizations and communication with the LLM via the Groq API. It is organized into modular components:
  - **app.py**: manages server initialization, API endpoints, session control and orchestration of training and prediction workflows.
  - **model_utils.py**: implements AutoML training logic, performance evaluation and visualization generation.
  - **chatbot_utils.py**: handles LLM queries and formats responses for user interaction.
  - **prompts.py**: stores reusable prompt templates for instructing the LLM to explain datasets, preprocessing steps and model outputs.

Temporary directories (**uploads/**, **models/**, **decompression/**) are used to manage file storage during the lifecycle of user sessions.

2.3 User Workflow

The platform supports two main workflows: training and prediction. In the training phase, users upload a dataset (CSV, Excel or ARFF), select a target variable, configure preprocessing options (e.g., missing value handling) and define a training time limit. AutoGluon then automatically builds and selects the best-performing model. The system displays evaluation metrics, model leaderboards and visualizations such as confusion matrices, ROC curves and SHAP plots (see Figure 2). The trained model can be downloaded for future use.

In the prediction phase, users upload a trained model and a new dataset. The system generates predictions and presents them in a tabular format, with the option to download results as a CSV file.

Throughout both workflows, users can interact with the integrated LLM assistant, which provides real-time explanations and guidance tailored to non-technical users (see Figure 2). A final feature allows users to generate a downloadable PDF report summarizing the entire modeling process, including LLM-generated explanations.

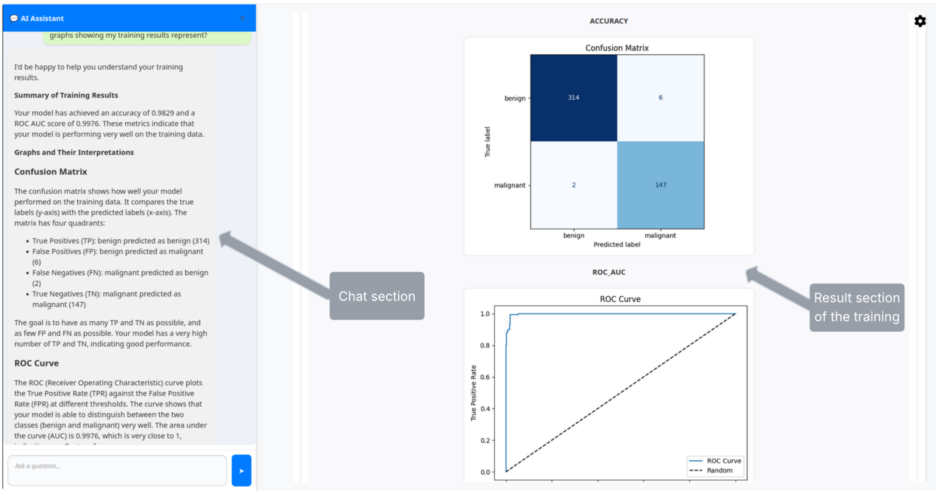


Figure 1: Screenshot of the interface displaying the chat section with a virtual assistant based on LLM for result explainability.

3 Experimental Results

To evaluate the effectiveness of the proposed platform, we conducted experiments on eight publicly available medical datasets from Kaggle and OpenML: four for classification and four for regression. These datasets were selected because they include a predefined train-test split, allowing for consistent and fair evaluation. For each case, the only inputs provided to the system were the training set (along with an indication of the target column) and the test set for generating predictions. A time budget of five minutes was specified for the AutoML search process. The system was run using its default configuration, with minimal user intervention, to simulate realistic usage by non-experts.

For binary classification tasks, performance was evaluated using the *ROC AUC*; for multiclass classification, we used *accuracy*; and for regression tasks, the *root mean squared error (RMSE)*. The results obtained by our system were compared to the best scores available in the public rankings of the respective datasets. Table 1 summarizes the performance of our approach relative to these community benchmarks.

Table 1: Performance comparison on eight medical datasets

Dataset	**Task**	**Metric**	**Our result**	**Best reported**
Breast Cancer (Wisconsin)	Binary classification	ROC AUC	0.980	0.991
Pima Indians Diabetes	Binary classification	ROC AUC	0.785	0.841
Contraceptive M. C.	Multiclass classification	Accuracy	0.565	0.574
Hypothyroid	Multiclass classification	Accuracy	0.997	0.998
Liver Disorders	Regression	RMSE	3.35	3.38
BrisT1D Blood Glucose	Regression	RMSE	2.56	2.36
COVID-19 Death Prediction	Regression	RMSE	299	226
COVID-19 Cases Prediction	Regression	RMSE	1.036	0.869

As shown in Table 1, the proposed AutoML platform achieved competitive results across all datasets, often approaching or matching the best community-reported scores. For several tasks, such as Liver Disorders and Hypothyroid, the system's performance was nearly identical to the top benchmark. In other cases, particularly in more complex regression tasks like COVID-19 predictions, there was a noticeable performance gap which can be explained by the limited training time and by leaderboard effects such as 'hill climbing' in public Kaggle datasets. Nevertheless, the results demonstrate the platform's ability to deliver strong baseline models with minimal effort, highlighting its practicality for rapid, low-expertise deployment in medical prediction tasks.

4 Concluding Remarks and Future Work

This work introduces a practical tool that bridges the gap between medical expertise and ML technology. By combining AutoML with LLM-based assistance, the platform empowers healthcare professionals to independently explore and apply AI to their research. The system's intuitive design, explainable outputs and robust performance make it a valuable asset for interdisciplinary collaboration in medical data science. Future work will explore support for multimodal data, enhanced personalization and integration with electronic health record systems to further expand its impact in clinical settings. We also plan to conduct user evaluation studies to assess the quality of LLM explanations.

Acknowledgements This work has been partially supported by grants PID2022-139237NB-I00 and PID2023-146243OBI00 funded by MICIU/AEI/10.13039/501100011033 and by "ERDF/EU".

References

1. Sharma, A., Lysenko, A., Jia, S., Boroevich, K.A., Tsunoda, T.: Advances in AI and machine learning for predictive medicine. J. Hum. Genet. 69(10), 487–497 (2024). https://doi.org/10.1038/s10038-024-01234-5
2. Shi, H., Xu, Z., Wang, H., Qin, W., Wang, W., Wang, Y., Wang, H.: Continual learning of large language models: A comprehensive survey. ACM Comput. Surv. (2024). https://doi.org/10.1145/3657572
3. Barbudo, R., Ventura, S., Romero, J.R.: Eight years of AutoML: categorisation, review and trends. Knowl. Inf. Syst. 65(12), 5097–5149 (2023). https://doi.org/10.1007/s10115-023-01894-2
4. Scott M. Lundberg and Su-In Lee. A Unified Approach to Interpreting Model Predictions. In: Proceedings of the 31st International Conference on Neural Information Processing Systems (NeurIPS), pp. 4765–4774 (2017)

Khaoula Ballouk
ORCID: 0009-0002-3834-7133

Ismail Lagrat
ORCID: 0009-0008-0652-8742

Laboratory of Advanced Systems Engineering
Ibn Tofail University
Kenitra, Morocco

khaoula.ballouk@uit.ac.ma, ismail.lagrat@uit.ac.ma

Advancing Emergency Department Operations through AI-Based Triage Systems

Abstract. Medical emergency services often struggle to meet patient needs promptly amid overcrowding and limited resources, with traditional triage methods relying on human judgment that may falter under pressure. Artificial intelligence (AI) presents a promising alternative by automating case prioritization through real-time clinical data analysis. This study adopts a hybrid approach, combining systematic and narrative review methodologies. Drawing on a selection of peer-reviewed articles from 2019 to 2024 found via targeted searches on Scopus, this review highlights the strengths of AI-based triage systems—including improved patient sorting, reduced wait times, and more efficient resource use. To ensure responsible adoption, ongoing efforts must focus on refining AI models, integrating wearable technologies and training healthcare professionals. As healthcare demands grow, AI-driven triage stands out as a vital innovation for enhancing emergency department operations, improving patient outcomes, and supporting clinicians in high-pressure settings.

Keywords: Emergency care unit, Healthcare transformation, Machine learning.

1 Introduction

Emergency departments (EDs) are a critical pillar of modern healthcare, providing urgent care for a wide range of conditions—from minor injuries to life-threatening emergencies [1, 2]. As a primary entry point for many patients lacking immediate alternatives, EDs must respond swiftly and effectively. However, rising demand combined with limited resources has led to chronic overcrowding, compromising care delivery, delaying treatment, and increasing pressure on medical staff [3]. This environment contributes to clinician fatigue and diminished job satisfaction, further straining ED operations.

Traditional triage systems, such as the Manchester Triage System (MTS), are designed to prioritize patients based on symptom severity [2, 4]. While standardized, these models depend on clinical judgment and can introduce variability, particularly during peak hours or large-scale incidents. Inconsistencies, cognitive bias, and the time required for manual evaluation often hinder timely care.

To address these shortcomings, AI- and ML-based triage solutions have emerged as promising alternatives [5]. These technologies can rapidly process large volumes of patient data—vital signs, symptoms, history, demographics—to support more objective and consistent triage. Unlike conventional methods, AI systems apply algorithmic models that reduce subjectivity and improve decision accuracy.

Research shows that AI can evaluate multiple clinical variables in seconds, outperforming traditional protocols in identifying high-risk patients and expediting interventions [6]. For example, patients with pre-existing conditions may be more accurately triaged using AI models that integrate historical and current data. Additionally, dynamic adjustment of patient priority based on real-time ED capacity enables better use of staff and resources, improving patient outcomes and reducing workflow strain [7, 8].

This hybrid review article aims to investigate recent progress in AI-powered triage solutions, highlighting their practical effects on emergency department operations and their capacity to enhance patient care outcomes. Through an analysis of contemporary studies and AI applications in the triage process, it seeks to offer a thorough insight into how these innovations may boost departmental efficiency, overcome the drawbacks of conventional triage approaches, and fundamentally transform emergency medical services. Furthermore, the review addresses the implementation hurdles associated with integrating AI into healthcare, emphasizing that despite its promising potential, successful adoption depends on meticulous planning, comprehensive validation, and a steadfast dedication to maintaining high standards of care amid technological evolution.

2 Methodology and analysis

A structured hybrid systematic and narrative literature review was performed to investigate the use of AI-powered triage systems within emergency departments. From an initial pool of 601 records, 19 studies were selected based on their relevance to the research scope, while non-pertinent publications were excluded. The review process was conducted using a systematic and structured approach, encompassing several key stages. It began with the formulation of research questions, aimed at defining the core issues explored in the study, followed by the development of a search strategy detailing the academic database consulted and the specific keywords used to identify relevant publications. Finally, a set of selection criteria was applied to determine which studies would be included or excluded based on their relevance and methodological quality. To establish a solid foundation for the review, an initial exploration of the field was carried out, focusing on how AI technologies are integrated into triage processes in emergency care. This helped shape a set of guiding research questions aimed at identifying current practices, the functionalities of AI systems, their demonstrated advantages, existing barriers to adoption, and recommendations for future development and deployment, and these guiding questions formed the basis for the subsequent analysis.

What are the possibilities and open questions regarding the integration of AI technologies into triage processes in emergency care ?

- **Q1** : What are the main functionalities offered by AI-powered triage systems to support clinical decision-making in emergency departments ?
- **Q2** : What demonstrated benefits have AI-powered triage systems provided in terms of patient prioritization, care efficiency, and resource optimization ?
- **Q3** : What are the key barriers hindering the widespread adoption of AI-powered triage systems in emergency care settings, and how might these be overcome ?

To address the research questions, a structured and progressive data collection strategy was employed. The process began with a broad investigation aimed at establishing a solid understanding of the general domain, which was then refined to target more specific issues. An inclusive and well-formulated search query was developed to capture a wide range of studies relevant to our topic :

("artificial intelligence" OR "machine learning") AND ("emergency department" OR "emergency services" OR "emergency room") AND (triage OR diagnosis OR prediction OR "patient flow") AND (systematic OR review OR application OR implementation)

All referenced data were sourced from digital scientific repositories. The primary database utilized was Scopus, selected for its extensive coverage of peer-reviewed publications related to our research question: How can AI-powered triage systems be effectively integrated into emergency departments to enhance patient prioritization and operational efficiency, while addressing barriers to adoption? A detailed overview of the data collection process is presented in **Figure 2**.

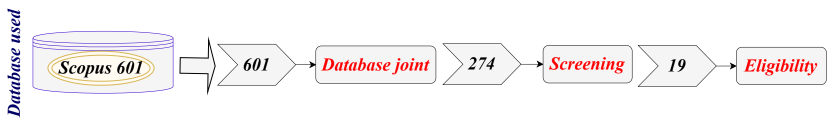


Figure 1: Filter process of selected articles.

To ensure methodological rigor, the PRISMA was followed, providing a structured approach for conducting and reporting systematic reviews. This methodology encompasses key stages such as study identification, selection criteria, quality assessment, and outcome reporting. The application of the PRISMA framework in this review followed these main phases :

- **Identification** : Relevant studies were gathered from multiple sources, including electronic literature databases, citation tracking, theses, and research reports, with the goal of capturing all potentially relevant work within the scope of the research topic.
- **Screening** : Retrieved studies were preliminarily evaluated based on their titles, abstracts, and keywords to determine their alignment with predefined inclusion criteria. Studies that did not meet the basic relevance threshold were excluded from further review.
- **Eligibility** : Full-text assessments were conducted on the remaining articles to evaluate methodological quality and compliance with inclusion and exclusion criteria. Studies that lacked sufficient detail or failed to meet the standards were removed from consideration, while those that fulfilled the criteria were included for in-depth analysis.

In line with systematic review protocols, non-scientific sources were excluded, and duplicate records across databases were eliminated to ensure the integrity and relevance of the final dataset.

3 Essential components of AI-powered triage systems

AI-based triage systems integrate advanced technologies aimed at improving patient assessment speed and accuracy in emergency departments [5]. Core components include structured data processing, predictive algorithms and real-time analytics [9]. Studies show these systems outperform traditional methods in identifying critical cases, such as sepsis, by accelerating intervention and improving detection precision [10]. In practice, hospitals report notable improvements, including up to 30% reductions in wait times and more efficient resource management through adaptive triage based on patient vitals and ED capacity [9]. These integrated technologies (see figure 2) form the backbone of AI-driven triage, offering promising solutions to reduce overcrowding, optimize patient flow, and enhance care delivery [1, 11].

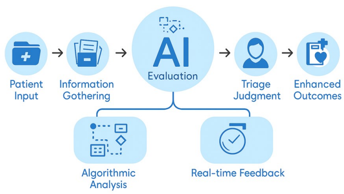


Figure 2: Filter process of selected articles.

3.1 Data acquisition and processing

A core element of AI-powered triage is the aggregation and processing of diverse patient data, which serves as the basis for clinical risk prediction [12]. These systems typically incorporate inputs such as vital signs, symptoms, medical history, and demographics to build detailed patient profiles. Wearable devices, for instance, provide real-time metrics like heart rate, respiratory rate, and oxygen saturation—key indicators for identifying early signs of clinical deterioration [13]. The combination of real-time monitoring and electronic health record (EHR) data enhances model adaptability across varied patient scenarios. Integrating past medical histories with current observations has also been shown to improve prediction accuracy, particularly for patients with chronic conditions presenting mild symptoms. By synthesizing these data, AI systems support more precise and individualized triage decisions.

3.2 Computational models

Machine learning algorithms form the backbone of AI-powered triage systems, enabling the classification of patient risk levels through complex data pattern recognition. Models such as decision trees, regression techniques, and neural networks are commonly employed, each trained on extensive historical clinical data to detect indicators of serious conditions. Among these, neural networks—especially deep learning architectures—have shown strong capabilities in identifying subtle, non-linear trends in patient profiles, thus improving triage precision [14]. In high-pressure settings, decision tree algorithms are also valued for their transparency, offering clinicians clear reasoning behind prioritization outcomes.

3.3 Real-time data analytics and feedback

AI-enabled triage improves emergency department operations by delivering real-time insights that support rapid clinical decision-making, particularly during overcrowded periods or large-scale emergencies [15]. These systems can detect acute changes in patient vitals—such as early signs of sepsis—and trigger immediate alerts for intervention. Additionally, real-time analytics aid in dynamic resource allocation, adjusting patient priority based on current load, bed occupancy, and specialist availability. This responsiveness reduces delays and enhances patient flow. By automating status monitoring, AI also eases clinicians' cognitive burden, allowing them to concentrate on high-impact decisions [15].

4 Enhancing Emergency Care: Clinical and Operational Benefits of AI-Driven Triage

A key advantage of AI-based triage is its capacity to rapidly identify and prioritize critical patients, significantly reducing wait times for those in urgent need [16]. By processing vital signs, symptoms, and medical history in real time, these systems can detect high-risk cases—such as strokes or cardiac events—within seconds. This speed is crucial, as early intervention greatly improves outcomes in acute scenarios. Evidence shows that AI triage can cut treatment delays by up to 20%, leading to faster care and improved prognoses [5, 9]. Ultimately, such efficiency not only enhances clinical quality but also boosts patient satisfaction. In addition to improving clinical responsiveness, AI-driven triage systems play a vital role in optimizing resource management within emergency departments, where demand fluctuates and capacity is often constrained. These systems support operational planning by predicting patient flow and clinical outcomes, enabling better allocation of staff, beds, and diagnostic equipment [17]. Drawing on both historical trends and real-time data, AI can anticipate surges—such as those during seasonal outbreaks—and inform proactive adjustments in staffing or treatment space. This predictive approach reduces overcrowding, improves patient throughput, and contributes to a more efficient and adaptive emergency care environment [11].

5 Prospects and recommendations

The adoption of AI triage depends largely on clinician trust and understanding. Effective training programs should explain system functionality, data inputs, decision factors, and showcase real-world benefits. Equally vital is educating on AI limitations and biases, enabling clinicians to critically evaluate and, if needed, override AI outputs. Research shows that transparency and evidence-based training increase clinician confidence and promote human-AI collaboration in emergency settings [18, 19].

6 Study limitations

A key limitation of this hybrid review is the article's length constraint (6 pages), which did not allow me to present additional benefits and detailed recommendations. Furthermore, it lacks the methodological rigour of systematic reviews, is restricted to articles published in English between 2014 and 2024 (thus excluding grey literature and non-English publications), and may be affected by publication bias favouring positive results.

7 Conclusion

Triage systems powered by artificial intelligence provide a promising approach to addressing persistent issues in emergency departments, including excessive patient loads and difficulties in prioritising care. Leveraging sophisticated algorithms alongside extensive data, these tools facilitate quicker and more precise evaluations, which enhances both resource utilisation and patient outcomes. Nevertheless, to guarantee their reliability and fairness, it remains crucial to tackle concerns such as data quality, potential biases within algorithms, clinician acceptance, and ethical implications. As these technologies continue to evolve, they hold the potential to significantly reshape emergency care by streamlining resource distribution and improving the delivery of medical services.

References

1. Samadbeik, M., Staib, A., Boyle, J., et al.: Patient flow in emergency departments: a comprehensive umbrella review of solutions and challenges across the health system. BMC Health Services Research **24**, 274 (2024). https://doi.org/10.1186/s12913-024-10725-6
2. Zaboli, A.: Establishing a common ground: the future of triage systems. BMC Emergency Medicine **24**, 148 (2024). https://doi.org/10.1186/s12873-024-01070-2
3. Yadgarov, M.Y., Landoni, G., Berikashvili, L.B., et al.: Early detection of sepsis using machine learning algorithms: a systematic review and network meta-analysis. Frontiers in Medicine **11**, 1491358 (2024). https://doi.org/10.3389/fmed.2024.1491358
4. Zachariasse, J.M., van der Hagen, V., Seiger, N., et al.: Performance of triage systems in emergency care: a systematic review and meta-analysis. BMJ Open **9**, e026471 (2019). https://doi.org/10.1136/bmjopen-2018-
5. Tahernejad, A., Sahebi, A., Abadi, A.S.S., et al.: Application of artificial intelligence in triage in emergencies and disasters: a systematic review. BMC Public Health **24**, 3203 (2024). https://doi.org/10.1186/s12889-024-20447-3

6. Friedman, A.B., Delgado, M.K., Weissman, G.E.: Artificial intelligence for emergency care triage—much promise, but still much to learn. JAMA Network Open **7**(5), e248857 (2024). https://doi.org/10.1001/jamanetworkopen.2024.8857
7. Singla, B., et al.: Revolutionizing the healthcare sector with AI. IGI Global (2024), 26 July.
8. Khang, A.: AI-driven innovations in digital healthcare: emerging trends, challenges, and applications. IGI Global (2024), 9 February.
9. Porto, B.M.: Improving triage performance in emergency departments using machine learning and natural language processing: a systematic review. BMC Emergency Medicine **24**, 219 (2024). https://doi.org/10.1186/s12873-024-01135-2
10. Marsilio, M., Roldan, E.T., Salmasi, L., et al.: Operations management solutions to improve ED patient flows: evidence from the Italian NHS. BMC Health Services Research **22**, 974 (2022). https://doi.org/10.1186/s12913-022-08339-x
11. Lukey, J.: Collection of assessment data. Medicine LibreTexts (2023). Retrieved from https://med.libretexts.org
Quizlet: Chapter 4 Review Flashcards. Retrieved from https://quizlet.com
12. Monnink, S.H., van Vliet, M., Kuiper, M.J., et al.: Clinical evaluation of a smart wristband for monitoring oxygen saturation, pulse rate, and respiratory rate. Journal of Clinical Monitoring and Computing (2024). https://doi.org/10.1007/s10877-024-01229-z
13. Pasini, A., Amendola, S.: A neural modelling tool for non-linear influence analyses and perspectives of applications in medical research. Applied Sciences **14**(5), 2148 (2024). https://doi.org/10.3390/app14052148
14. Porter, B., Grippa, F.: A platform for AI-enabled real-time feedback to promote digital collaboration. Sustainability **12**(24), 10243 (2020). https://doi.org/10.3390/su122410243
15. Yan, M.Y., Gustad, L.T., Nytrø, Ø.: Sepsis prediction, early detection, and identification using clinical text for machine learning: a systematic review. Journal of the American Medical Informatics Association **29**(3), 559–575 (2022). https://d
16. Ouanes, K., Farhah, N.: Effectiveness of artificial intelligence (AI) in clinical decision support systems and care delivery. Journal of Medical Systems **48**, 74 (2024). https://doi.org/10.1007/s10916-024-02098-4
17. Sutskever, I.: Ilya Sutskever predicts the end of pre-training as AI hits 'peak data'. OpenTools AI (2024).
18. Rosenbacke, R., Melhus, Å., McKee, M., Stuckler, D.: How explainable artificial intelligence can increase or decrease clinicians' trust in AI applications in health care: systematic review. JMIR AI (2024) 3.
19. World Health Organization (WHO): WHO outlines considerations for regulation of artificial intelligence for health (2023).

Marcin Stępniak[1]
ORCID: 0000-0002-5113-272X

Wojciech Nabiałek[1]
ORCID: 0000-0002-8361-2816

Artur Niewiadomski[1]
ORCID: 0000-0002-9652-5092

Mateusz Przychodzki[1]
ORCID: 0009-0000-2668-4875

Andrzej Sędek[2]
ORCID: 0000-0003-1486-7417

[1] University of Siedlce
Faculty of Exact and Natural Sciences
Institute of Computer Science
ul. 3 Maja 54, 08-110 Siedlce, Poland

[2] University of Siedlce
Faculty of Social Sciences
Institute of Pedagogy
ul. Żytnia 39, 08-110 Siedlce, Poland

{marcin.stepniak, wojciech.nabialek, artur.niewiadomski, mateusz.przychodzki, andrzej.sedek}@uws.edu.pl

Physiological Monitoring Devices for Virtual Reality Therapy: Technologies, Applications, and Future Directions

Abstract. Virtual reality (VR) therapy has emerged as a promising approach for treating psychological conditions, particularly anxiety and stress-related disorders. The integration of physiological monitoring hardware with VR environments enables real-time biofeedback and personalized therapeutic interventions. This article reviews physiological monitoring hardware technologies used in VR therapy, including cardiovascular sensors, electrodermal activity sensors, respiratory monitoring devices, and emerging metrics such as voice analysis and eye tracking. It examines hardware requirements for stress measurement applications, highlighting the advantages and limitations of various monitoring systems and their integration with VR platforms. Within the development of the VR TherapyS platform, there has been an evolution of hardware from basic ESP microcontroller-based sensor configurations in early stages to advanced devices like the Neurobit Optima+ for multi-modal physiological monitoring, with ongoing efforts toward an integrated wristband device combining multiple sensors in a user-friendly wearable format. Future developments focus on advanced sensor integration, miniaturization, and improved wireless connectivity to enhance accessibility, accuracy, and integration of physiological monitoring for VR-based therapeutic interventions.

Keywords: Physiological Monitoring, VR TherapyS, Virtual Reality Therapy, Social anxiety.

1 Introduction

Virtual reality (VR) therapy has transformed psychological treatment by creating immersive, controlled environments for exposure-based interventions, particularly for stress and anxiety disorders. The incorporation of physiological monitoring hardware enhances this approach by providing objective data on stress responses, allowing therapists to tailor sessions in real time. This review explores a broad range of hardware technologies for measuring physiological parameters in VR contexts, emphasizing stress assessment and therapeutic applications.

Key parameters include heart rate (HR), heart rate variability (HRV), galvanic skin response (GSR), respiratory patterns, lower oxygen saturation, or subtle muscle tremors. They are an excellent starting point for building devices that will be able to detect and classify the level of emotional tension in real time. Evolution of the hardware used for VR Therapys platform has progressed from wired, laboratory-based systems to wireless, wearable devices, improving accessibility and reducing intrusiveness. Initial phases used basic sensors connected to ESP microcontrollers for GSR and HR monitoring, evolving to the Neurobit Optima+ for multi-modal data collection, with current efforts focused on a wristband prototype integrating multiple sensors [1, 2].

In practice, multiparameter measurements not only provide higher reliability but also allow the analysis to be adapted to the individual reactivity of the body.

The new device (VR-feedback project) will be developed with users in mind who need an affordable, real tool for assessing stress levels, whether in therapeutic applications, scientific research, or training in virtual environments.

This article draws on scientific literature to review hardware technologies, their applications in VR therapy, advantages, limitations, and future directions.

2 Review of Physiological Data Collection Devices

Stress responses in VR therapy manifest as elevated HR, decreased HRV, and changes in electrodermal or respiratory signals, detectable by specialized hardware [18, 5].

2.1 Wrist-Worn Devices

Modern wearables collect HR/HRV (ECG/PPG), GSR/EDA, temperature, and motion. For example, Empatica wristbands (E4, EmbracePlus) record PPG, EDA, and movement, ensuring high-quality data in stress research [9]. Fitbit and Apple Watch also capture HR and HRV; however, most models restrict real-time data export, which limits use in biofeedback scenarios [11].

2.2 VR-Integrated Sensors

Headset or controller-embedded sensors (e.g., emteqGO, SensCon) provide seamless biosignal monitoring in VR, measuring EDA, PPG, and movement [10, 6]. Add-ons like HTC Vive with Tobii eye-tracking further support objective assessment in VR therapies [16].

2.3 Medical and Specialized Devices

High-precision devices such as the Neurobit Optima+, Bittium Faros 360, or BIOPAC BioNomadix enable advanced HRV, GSR, and respiration data collection, mainly for research [4]. Wireless respiratory belts and contactless radar-based sensors address usability challenges in VR [15].

2.4 Prototypes and Emerging Developments

New prototypes focus on miniaturization and VR adaptation. Projects like SensCon integrate PPG/EDA in controllers [6]; nano-wearables with biochemical sensors for cortisol enable deeper stress monitoring [8].

3 VR-feedback — Concept of a Device for Measuring Physiological Stress Parameters

The VR-feedback device will be based on a modern ESP-WROOM-32E microcontroller, which thanks to its dual-core architecture, Wi-Fi/Bluetooth connectivity, and TinyML support, will introduce functionalities to biomedical signal processing that until now were reserved mainly for large stationary systems.

3.1 Components Used

ESP-WROOM-32E is the future "brain" of the system. It will provide enough computing power to process signals from several sensors simultaneously. It will allow running an embedded TinyML model to classify stress levels in real time. Wi-Fi connectivity and REST API will enable the user to control the prototype remotely, without the distraction of manual operation.

Skin conductance measurement is one of the oldest but still effective tools for monitoring stress reactions. The use of a simple and inexpensive **LM324 operational amplifier** will provide sufficient signal amplification and filtering while keeping the prototype cost low. In commercial solutions, ready-made GSR sensors often cost much more, while our circuit, despite being budget-friendly, will provide comparable sensitivity.

MAX30102 — heart rate and oxygen saturation sensor — Reading pulse and blood oxygenation will make it possible to detect typical stress reactions such as increased heart rate. Compared to large medical pulse oximeters, this sensor stands out for its compact size and low power consumption, which is crucial in a mobile project.

The **AHT10, temperature and humidity sensor**, will provide not only high measurement accuracy but, above all, high measurement resolution (up to 0.01°C and 0.024% RH), which is crucial for tracking microfluctuations in skin temperature and humidity that may indicate sweating as a stress reaction. In commercial solutions, skin humidity is often ignored, although it carries valuable diagnostic information.

The **6DOF IMU module (LSM6DSOTR**) will detect subtle movements, muscle tensions, or micro-movements of the hand that are difficult to notice but often indicate internal tension.

3.2 Construction and Power Supply

The VR-feedback prototype will be designed as a compact PCB measuring approximately 6×4 cm, worn on the wrist like a sports band. The GSR sensors will be connected via wires with electrodes that the user will wear on two fingers or the wrist. Power will be supplied via a USB-C port, eliminating the need for charging from a 240 V mains supply and simplifying compliance with directives such as RoHS and CE safety. The lack of a battery and mains charger will also reduce the device's weight, increasing comfort during longer sessions.

3.3 Embedded Software

Sensor data will be cyclically sampled and processed by an embedded TinyML model that enables real-time baseline learning, stress detection, and continuous monitoring. Processed results will be served via REST API, ensuring data privacy on device; integration with external AI models allows for advanced trend analysis and calibration.

4 Applications, Challenges, and Future Directions

Physiological hardware enables objective stress measurement in VR, facilitating personalized exposure therapy. In TSST-VR adaptations, sensors monitor responses to virtual social stressors, showing comparable cortisol and HRV changes to real-world tests [13]. The articles describe social anxiety therapy in VR, emphasizing stress measurement (HR, HRV, GSR, respiration) in scenarios like TSST-VR. These devices can be used for real-time stress monitoring, and integrate with VR TherapyS, enabling biofeedback and scenario personalization (e.g., public speaking simulations). For social anxiety, hardware like Empatica tracks GSR during virtual interactions, allowing therapists to adjust scenarios [7]. VR TherapyS uses this for adolescent therapy, with pilot studies confirming effective stress induction [1]. Data collection in everyday settings is supported by devices like Fitbit and Apple Watch. Hardware requirements include low latency (<50ms) for biofeedback and compatibility with HMDs like

Oculus Quest. These applications align with VR integration, where devices like emteqGO and SensCon eliminate cables, improving immersion in therapies.

Advantages of current hardware include portability, real-time data, and non-invasiveness, enhancing VR immersion [12]. Limitations encompass motion artifacts, battery life, and cost, potentially hindering widespread adoption [3]. In VR TherapyS, the shift to wearables addresses intrusiveness, though signal quality in dynamic VR remains a challenge.

Technical challenges include ensuring signal integrity in VR motion and data privacy in wireless systems [17]. Cost barriers limit access, though microcontrollers like ESP offer affordable entry points.

In the future, VR-feedback could be improved with additional EKG sensors for more precise detection of micro-changes in the heart signal, breath analysis (for example, via an additional chest pressure sensor), as well as automatic generation of reports and visualizations for clinical or psychotherapeutic purposes.

Future hardware will incorporate AI-driven sensor fusion and nanomaterials for ultra-miniature devices. Enhanced connectivity via 5G could enable cloud-based processing [14]. VR TherapyS's wristband prototype exemplifies this, promising integrated voice and eye-tracking sensors for advanced stress detection [2]. In the last 10 years, there has been rapid development of wearable devices, from simple trackers to advanced biosensors integrating with VR. Devices like Empatica E4, Neurobit Optima+, and wristband prototypes fit ideally into stress therapy applications described in the articles, enabling continuous monitoring and personalization. Future directions include AI and nanomaterials for even higher precision.

5 Conclusion

Physiological monitoring hardware is pivotal for advancing VR therapy, providing tools for precise stress assessment and personalized interventions. From basic ESP systems to advanced devices like Neurobit Optima+, the evolution enhances therapeutic efficacy. As exemplified by VR TherapyS, ongoing developments toward integrated wristbands will further democratize access, supported by rigorous scientific validation.

The VR-feedback project will prove that it is possible to achieve results comparable to commercial systems costing several thousand euros, even in the budget device sector. Thanks to the simplicity of its construction, flexible software, and open architecture, the device could become a real tool for popularizing biofeedback and stress research — both in laboratories and in everyday practice.

References

1. Bartyna, W., Stepniak, M.: The vr therapy platform for social anxiety treatment. Studia Informatica. System and information technology **29**(2), 27–39 (2023). `https://doi.org/10.34739/si.2023.29.02`
2. Bartyna, W., Stepniak, M.: Data collection and analysis in social anxiety vr therapies. Studia Informatica **31**(2), 62–68 (2024). `https://doi.org/10.34739/si.2024.31.05`
3. Betti, S., Lova, R.M., Rovini, E., Acerbi, G., Santarelli, L., Cabiati, M., Ry, S.D., Cavallo, F.: Evaluation of an integrated system of wearable physiological sensors for stress monitoring in

working environments by using biological markers. IEEE Transactions on Biomedical Engineering **65**(8), 1748–1758 (2018). https://doi.org/10.1109/TBME.2017.2764507
4. BIOPAC: Bionomadix wireless wearable physiology. https://www.biopac.com/product-category/research/bionomadix-wireless-physiology/, accessed: 2025-07-15
5. Boucsein, W.: Electrodermal Activity. Springer Science & Business Media (2012)
6. Chiossi, F., Kosch, T., Menghini, L., Villa, S., Mayer, S.: Senscon: Embedding physiological sensing into virtual reality controllers. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies **7**(MHCI) (Sep 2023). https://doi.org/10.1145/3604270, https://doi.org/10.1145/3604270
7. Emmelkamp, P., Meyerbröker, K., Morina, N.: Virtual reality therapy in social anxiety disorder. Current Psychiatry Reports **22** (05 2020). https://doi.org/10.1007/s11920-020-01156-1
8. Feng, X., Huda, A.: Energy efficient nano wearable devices optimised with integration of gaining sharing knowledge algorithm for realtime health monitoring. Journal of Nano Molecular Intelligence and Virtual Health Systems (01 2025). https://doi.org/10.70023/nmivhs.251104
9. Garbarino, M., Lai, M., Tognetti, S., Picard, R., Bender, D.: Empatica e3 - a wearable wireless multi-sensor device for real-time computerized biofeedback and data acquisition. pp. 39–42. IEEE (12 2014). https://doi.org/10.4108/icst.mobihealth.2014.257418
10. Gnacek, M., Seiss, E., Kostoulas, T., Balaguer-Ballester, E., Nduka, C.: Remote collection of physiological data in a virtual reality study collection of physiological data in a fully remote setting via bio signal enabled virtual reality headset (05 2021)
11. Google: How do I get started with fitbit charge 5? - fitbit help center. https://support.google.com/fitbit/answer/14237922?hl=en, accessed: 2025-07-15
12. Han, H.J., Labbaf, S., Borelli, J.L., Dutt, N., Rahmani, A.M.: Objective stress monitoring based on wearable sensors in everyday settings. J Med Eng Technol **44**(4), 177–189 (jun 2020)
13. Kothgassner, O.D., Felnhofer, A., Hlavacs, H., Beutl, L., Palme, R., Kryspin-Exner, I., Glenk, L.M.: Salivary cortisol and cardiovascular reactivity to a public speaking task in a virtual and real-life environment. Computers in Human Behavior **62**, 124–135 (2016). https://doi.org/https://doi.org/10.1016/j.chb.2016.03.081, https://www.sciencedirect.com/science/article/pii/S0747563216302400
14. Kriklenko, E., Kovaleva, A., Klimenko, A., Dukuev, U., Pertsov, S.: Multimodal assessment of changes in physiological indicators when presenting a video fragment on screen (2d) versus a vr (3d) environment. Behavioural Neurology **2022**(1) (2022). https://doi.org/https://doi.org/10.1155/2022/5346128, https://onlinelibrary.wiley.com/doi/abs/10.1155/2022/5346128
15. Quanjer, P., Tammeling, G., Cotes, J., Pedersen, O., Peslin, R., Yernault, J.C.: Lung volumes and forced ventilatory flows. European Respiratory Journal **6**(Suppl 16), 5–40 (1993). https://doi.org/10.1183/09041950.005s1693, http://erj.ersjournals.com/content/erj/6/Suppl%5F16/5
16. Schuetz, I., Fiehler, K.: Eye tracking in virtual reality: Vive pro eye spatial accuracy, precision, and calibration reliability. J Eye Mov Res **15**(3) (Sep 2022)
17. Shirtcliff, E.A., Finseth, T.T., Winer, E.H., Glahn, D.C., Conrady, R.A., Drury, S.S.: Virtual stressors with real impact: what virtual reality-based biobehavioral research can teach us about typical and atypical stress responsivity. Translational Psychiatry **14**(1), 441 (Oct 2024). https://doi.org/10.1038/s41398-024-03129-x, https://doi.org/10.1038/s41398-024-03129-x
18. Thayer, J.F., Lane, R.D.: A model of neurovisceral integration in emotion regulation and dysregulation. Journal of Affective Disorders **61**(3), 201–216 (2000). https://doi.org/https://doi.org/10.1016/S0165-0327(00)00338-4, https://www.sciencedirect.com/science/article/pii/S0165032700003384, arousal in Anxiety

Dariusz Mikułowski
ORCID: 0000-0002-4567-7846

University of Siedlce
Faculty of Exact and Natural Sciences
Institute of Computer Science
ul. 3 Maja 54, 08-110 Siedlce, Poland

dariusz.mikulowski@uws.edu.pl

Automatic Creation of Sound Binaural Scenes To Develop Spatial Orientation Skills of Blind Individuals

Abstract. This article introduces an approach to generating automatic binaural sounds to create audio scenes that describe environments for visually impaired individuals. Such environmental descriptions can be beneficial for developing spatial orientation skills. Artificially generated binaural scenes can simulate situations where a visually impaired person navigates open spaces, such as a city, or large buildings like train stations, airports, exhibition halls, universities, or hospitals. Practising these skills in a simulated environment can be less stressful than real-world training. This work describes the development of a specialised application that facilitates the creation of binaural sound scenes from simple monophonic sounds. Several scenes were subsequently generated and then tested by three individuals who are visually impaired. This testing aimed to determine whether entirely artificially generated sound scenes are a sufficiently robust approach for constructing more complex simulation environments.

Keywords: Binaural sounds, Spatial orientation of the blind, Automatic sound generation.

1 Introduction

Independent movement is crucial for a blind person's daily life and rehabilitation. However, real-world training can be stressful and requires significant resources, including trained instructors and specialised infrastructure like spatial orientation parks [11].

We propose a virtual simulation environment as an alternative, using **Virtual Reality** (VR) and **Artificial Intelligence** (AI). In this setup, VR can replace physical infrastructure, and AI can act as a virtual instructor. A key element of this simulation is modelling the user's sound environment. We use **binaural sounds**, which create a 360-degree auditory experience, making it possible to simulate complex soundscapes. These sounds can be generated from simple monophonic sounds using, for example, the **Head-Related Transfer Function** (HRTF). The HRTF filter imitates how sound waves are reflected by a person's head and body, altering sound frequency, volume, and arrival time to create the illusion of direction and movement. By processing individual monophonic sounds, we can create acoustic binaural scenes that realistically model situations a blind person might encounter. These scenes include scenarios like walking along a street, crossing an intersection with traffic and a sound-enabled traffic light, or being inside a public vehicle. Our method also allows for the simulation of sounds like passing cars, pedestrians' footsteps, and the tapping of a white cane.

The core of our work is an application that generates these detailed binaural sound scenes from several input monophonic sounds, arranging them in a chronological sequence to create immersive training scenarios. An application was developed as part of a project within the framework of an engineering thesis of Adrian Wysocki [12].

2 The state of art

Research is underway to improve space orientation and navigation for individuals with visual impairments by using a combination of auditory, haptic, and virtual reality (VR) technologies. Auditory systems are a significant focus, especially those using binaural sound to create a Virtual Spatial Reality (VSR). Mikulowski and Szuba's work [9] on a VSR game utilises 3D binaural audio to enhance spatial perception, which can be delivered through bone conduction headphones, allowing users to hear both virtual and real-world sounds simultaneously. Earlier research by the same group [8] proposed an "Ontologically-Based Object Map" that uses binaural sounds to create a simplified audio map for navigation and pre-trip planning.

Binaural guidance is also central to an autonomous navigation system by Hajebi et al [3]. This system uses mobile devices to combine real-time positioning with 3D binaural audio, providing features for journey planning and real-time guidance. To avoid interfering with ambient sounds, Fujita et al. [2] developed a navigation device that uses bone conduction to deliver binaural sound, leaving the ear canals open. This method has been shown to reduce navigation time and improve localisation accuracy. The American Foundation for the Blind's paper, "Exploring New Horizons: Binaural Soundscapes," [1] highlights the importance of **Interaural Time Difference (ITD)** and **Interaural Level Difference (ILD)** in creating immersive 3D soundscapes. This research extends the use of binaural sound beyond navigation to education and recreation, emphasising the need to involve visually impaired individuals in the design process to ensure the technology is truly user-centred.

Another approach is the use of haptic feedback, which provides a tactile dimension to spatial orientation through vibrations and physical sensations. Kleinberg et al. [5] developed a portable haptic system that utilises computer vision, ultrasonic sensors, and machine learning to provide real-time tactile feedback, alerting users to obstacles and promoting safer navigation. Similarly, Kaplan et al. [4] investigated using standard gamepads with haptic feedback and audio descriptions of landmarks to help users with route learning and environmental preparation before they physically enter a space.

VR environments and simulators provide a safe and controlled space for individuals with visual impairments to learn and practice spatial skills. Lahav's review [6] highlights the potential of VR as a tool for navigation planning, training, and as a diagnostic aid for Orientation and Mobility specialists.

Several software libraries exist for creating realistic binaural sounds. **OpenAL** [7] is a cross-platform 3D audio API that models sound sources relative to a listener, supporting various effects crucial for realistic binaural audio. **OpenTK** [10] is a C# binding that allows .NET programs to access OpenAL's functions. Another solution is a **Slab3D**, which is a real-time virtual environment rendering system designed explicitly for 3D sound processing and includes Head-Related Transfer Function (HRTF) databases for accurate binaural rendering. Finally, the Unity Audio Spatializer SDK is also noted for its ability to generate binaural sound.

3 Designing binaural scenes

3.1 Binaural scene generator

The system proposed in this work was implemented as a monolithic desktop application in .NET and C#. The program features a simple graphical interface that enables the creation of binaural sound scenes from input monophonic sounds. OpenTK and OpenAL libraries were used to work with the sound content.

After starting the application, a form is presented to add individual sounds to the newly created binaural scene. In addition to the possibility of adding the sound file itself, it is necessary to provide many parameters that describe the entire binaural scene, both in the context of its component files and time dependencies. The idea is to enable playing the sounds at a specific time, in sequence or in parallel, as well as determine the direction of sound movement towards the listener. Therefore, in the main application form, there are parameters defining the behaviour of the selected sound in the scene and its features, such as the speed of movement or looping of the sound. The following fields are available here:

- The initial position of the sound source relative to the listener, determined by coordinates x, y and Z The listener is located at point 0, 0, 0 in space, and it is directed "face deep into the screen", i.e. in the direction of negative values on the Z axis.
- a direction in which the source will move, conditioned by giving the value of x, y and z,
- the speed of the source movement in the three-dimensional space, determined by the value x, y and z. The values in the speed vector determine how fast the sound moves in each of the axes.
- GAIN - sound volume, the default value is 1 - representing 100
- sound pitch, directly proportional to the speed of its playback, the default value is 1 - representing 100As the sound reproduction increases, its height also increases.
- the time of starting sound playback, expressed in seconds from the beginning of the scene, the default value is 0,
- looping sound - a double selection button that determines whether sound playback is to be repeated.

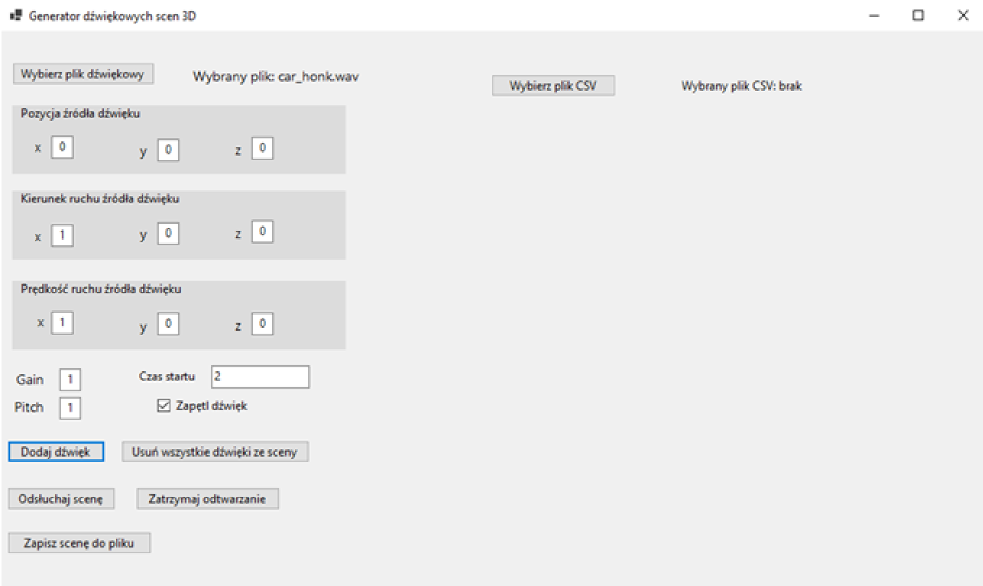


Figure 1: The graphic interface of the scene generator after selecting the .Wav file, providing additional guidelines and adding sound to the stage. Source: own study.

The main form of the application is presented in figure 1.

With the method described above, by completing subsequent forms and choosing files, many sounds can be added to the scene. However, the situation may occur when the user does not want to manually enter the parameters of each sound separately. In this case, there is an option to create a file with the CSV extension, for example, using Excel. The user can thus enter parameters for all sounds at once according to a set scheme, and then load the data from this file into the program. The structure of the CSV file corresponds to the above-mentioned fields in the form. Thus, in the subsequent columns, user can give: the path to the file, three coordinates of the starting point of the sound, three coordinates of the sound direction, three coordinates of the sound speed vector, sound volume, pitch sound, the time of its launch expressed in decimal parts of a second and the value of 1 OUB 0 indicating whether the sound is to be looped. In Figure 2, the application interface is presented after selecting a CSV file.

After creating the whole scene, the user can listen to it. It is recommended to listen to the scenes using headphones, because in this case, the user can experience the full quality of the binaural scene. To listen to the scene or stop it, the appropriate buttons can be used. The last button is used to save the entire scene to a separate sound file, recorded using the binaural technique.

3.2 Audio scene development

The binaural scene generator was used to produce three examples of scenes representing simple situations that a BI may encounter in the city. These were the following scenes:

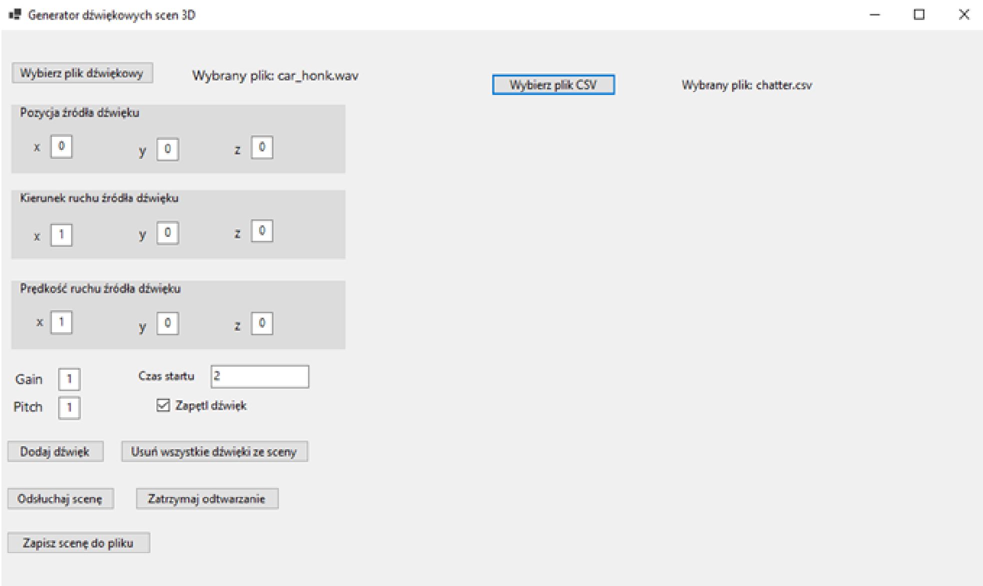


Figure 2: The graphic interface of the scene generator after selecting the CSV file and loading the data into it. Source: own study

- Scene S_1: The user goes along the street. She/He may listen to the sounds of cars driving from the back to the front and from the front to the back on her/his left side. Additionally, the sound of passersby in front of the user may sometimes be emitted. The user also hears the sound of his own steps and the sound of hitting a white cane on the pavement.
- Scene S_2: The user is approaching the street bend. On his left and in front of him, the user hears the sound of driving cars. Cars move from back to the front of the left side of the listener and then turn and move from left to right in front of the listener. Other cars move from right to left in front of the listener and then change direction and move from the front to the back on the left side of the listener. As in the case of a confessing scene, the user also hears the sound of his own steps and the sound of hitting a white cane against the pavement.
- Scene S_3: The user is in front of the pedestrian crossing through the intersection. As in the S_1 scene, the user hears the sounds of cars driving on his left and from the back to the front. In addition, he hears the sound of traffic lights (quick taps) in front of him, meaning that a green light is currently shining, which allows him to cross the pedestrian crossing. In the case of this scene, no sounds of footsteps are heard.

4 Evaluation

The main goal of the evaluation of the created solution was to answer the following questions:

- Are automatically generated binaural scenes give the impression of a real environment, so close to realism

– Can such scenes be used for constructing simulation environments supporting the training of mobility skills and spatial orientation of the blind individuals?

To answer these questions, three blind users were presented with ready-made sound scenes. The blind testers were experienced people who, daily, themselves move in the opening and have trained skills in the field of spatial orientation. They could listen to each stage five times and then answer a short survey consisting of several questions.

– ”What particular sounds do you hear within the presented scene?”
– ”What situation from the real world did you listen to in this scene?”
– ”Please rate on a scale of 1 to 10, how much the scene presented is, in your opinion, similar to the real situation in the environment?”
– ”What do you suggest to improve in the presented scene to make it closer to realism?”

Additionally, users could provide general comments on the survey.

All users easily recognised the type of sounds contained in the scene, such as the sound of a driving car, the sound of steps or the sound of hitting a cane against the pavement. They also had no problem recognising the situation presented by a given scene, although two testers had comments as to the lifts and a light signalling. Slightly worse (usually at 4 out of 10 points), users rated the similarity of such automatically generated scenes to reality. Such scenes are rather suitable for training basic skills to recognise the environment before moving to scenes closer to the realism of those recorded by nature.

5 Discussion and conclusions

In this work, a solution for automatically creating binaural sound scenes is proposed. These scenes can then be used to construct simulated environments that support developing spatial orientation skills for individuals with visual impairments. Our specific approach focused on creating simple binaural scenes from basic monophonic sounds. To achieve this, an application was developed as part of a project within the framework of an engineering thesis of Adrian Wysocki [12]. Using this application, several binaural scenes were generated, which were then tested by three blind users. The results of the experiments indicate that automatically generated specific sounds in these scenes are more distinct and less easily drowned out by ambient noise or additional sounds, which can be an issue with scenes recorded from real environments. For a beginner user who is just starting to learn spatial recognition through hearing, this approach can serve as an excellent foundational tool. Therefore, this technique is well-suited for creating initial sound scenes until users are ready for more advanced scenes created through recording. In the case of generating binaural scenes with an application, it is problematic that the creation of such a scene turned out to be time-consuming. This activity requires many attempts, tuning the parameters of each sound until the desired result is achieved. So we can conclude that recording scenes by nature is a faster method. However, in the future, we plan to enhance the application to test more scenes. We also aim to explore separate applications where this method might be more suitable than other solutions, such as recording scenes from an environment or using libraries designed for computer games.

References

1. American Foundation for the Blind: Exploring new horizons: Binaural soundscapes. American Foundation for the Blind Blog Post (Jan 2024), https://www.afb.org/blog/afb-blog/exploring-new-horizons-binaural-soundscapes, [Online: Access 2025-07-17].
2. Fujita, K., Sato, T., Inooka, T., Minato, T.: Bone conduction auditory navigation device for blind people. Applied Sciences **11**(8), 3356 (2021). https://doi.org/10.3390/app11083356, https://www.mdpi.com/2076-3417/11/8/3356
3. Hajebi, A., Asghari, F., Fereydounian, A.: Autonomous navigation based on binaural guidance for people with visual impairment. ResearchGate (Pre-print/Conference) (Feb 2025), https://www.researchgate.net/publication/380720498_Autonomous_Navigation_based_on_Binaural_Guidance_for_People_with_Visual_Impairment, likely presented at a conference; available on ResearchGate
4. Kaplan, H., Pyayt, A.: Fully digital audio haptic maps for individuals with blindness. Disabilities (2024)
5. Kleinberg, D., Yozevitch, R., Abekasis, I., Israel, Y., Holdengreber, E.: A haptic feedback system for spatial orientation in the visually impaired: A comprehensive approach. IEEE Sensors Letters **7**, 1–4 (2023)
6. Lahav, O.: Virtual reality systems as an orientation aid for people who are blind to acquire new spatial information. Sensors (Basel, Switzerland) **22** (2022)
7. Madmoni, L., Ben-Hur, Z., Donley, J., Tourbabin, V., Rafaely, B.: Design and analysis of binaural signal matching with arbitrary microphone arrays and listener head rotations. EURASIP J. Audio Speech Music. Process. **2025**, 11 (2024)
8. Mikułowski, D., Pilski, M.: An approach for supporting space orientation of the blind using ontologically-based object map. In: Proceedings of the International Conference on Human Aspects of IT for the Aged Population (2015), specific conference details (e.g., volume, pages) may vary based on publication
9. Mikułowski, D., Szuba, B.: Supporting the development of spatial orientation skills of blind people using binaural sounds in VSR game. Studia Informatica. System and Information Technology **31**(2), 49–60 (2025)
10. Ratnarajah, A., Manocha, D.: Listen2scene: Interactive material-aware binaural sound propagation for reconstructed 3d scenes. 2024 IEEE Conference Virtual Reality and 3D User Interfaces (VR) pp. 254–264 (2023)
11. Specjalny Ośrodek Szkolno Wychowawczy dla Dzieci Niewidomych w Owińskach: Park orientacji przestrzennej, https://www.niewidomi.edu.pl/park-orientacji-przestrzennej/, online; [Access: 03 Sep. 2025]
12. Wysocki, A.S.: Projekt i implementacja aplikacji do tworzenia dźwiękowych scen 3D dla potrzeb wspomagania nauki orientacji przestrzennej osób niewidomych zrealizowana z użyciem biblioteki OpenAL. Master thesis (2025)

Stefan Svetsky[1]
ORCID: 0000-0002-2186-5957

Dariusz Mikułowski[2]
ORCID: 0000-0002-4567-7846

Zuzana Cervenanska[1]
ORCID: 0000-0002-6380-4634

Wojciech Nabiałek[2]
ORCID: 0000-0002-8361-2816

[1] Slovak University of Technology,
Bratislava, Slovakia

[2] University of Siedlce
Faculty of Exact and Natural Sciences
Institute of Computer Science
ul. 3 Maja 54, 08-110 Siedlce, Poland

stesvet11@gmail.com, dariusz.mikulowski@uws.edu.pl,
zuzana.cervenenska@stuba.sk, wojciech.nabialek@uws.edu.pl

WPad – Personal User Adapted System as a Supporting Assistive Technology For BVI Users

Abstract. Despite the availability of various assistive tools, BVI teachers still face significant challenges in tasks such as preparing teaching materials, conducting classes, and carrying out research. These activities often require switching between multiple applications, reducing efficiency. To address this, we have been developing an integrated support system consisting of the WPad desktop application, the PIKS web modules (communication channels), and a virtual machine with Windows Server 2022 (as a shared remote computer) enabling team communication. This system works as a Personal User Adapted System also supporting Assistive Technology. It is tailored to the needs of both sighted and BVI users, educators, and researchers, and has been successfully validated in both educational and research activities.

Keywords: User-adapted system, assistive technology, blind and visually impaired, non-relational database paradigm, communication channel, educational packs

1 Introduction

Although many IT tools exist to support Blind and Visually Impaired (BVI) users, performing complex tasks-such as preparing structured documents, teaching, or conducting research-remains challenging. These tasks often require switching between multiple applications, each with its own interface and keyboard shortcuts, which can reduce efficiency and increase cognitive load. Screen readers are essential for BVI users, converting on-screen text into synthetic speech or Braille output [10]. While modern screen readers are available across platforms: JAWS [1], NVDA [2], Narrator (Windows), TalkBack (Android) [4], VoiceOver (Apple) [5], and Orca (Linux) [3]-they still struggle with graphical content and inconsistent accessibility across applications. Users often rely on multiple tools for file management, browsing, and text editing, including OpenOffice, Microsoft Office, and LaTeX [12].

Recent research highlights broader accessibility concerns. Lee and Ashok [11] show how emojis and slang hinder screen reader comprehension on social media, while Pandey et al. [13] reveal usability gaps in mobile apps for visually impaired students. Devi and Kumar [7] critique DRM restrictions that block assistive technologies, and Diaz et al. [8] propose methods to improve accessibility in educational repositories. Other studies offer insights into inclusive system design. Cao [6] identifies trust and usability as key factors in software adoption, and Gori et al. [9] present adaptive file management tools that could benefit BVI users. However, longterm adoption and integration into daily routines remain underexplored.

To address these gaps, we developed a personal user-adapted system tailored to both BVI and Deaf users. It combines WPad as the Windows desktop application, the PIKS web communication modules (Personal Information and Knowledge System as the PHP/MySQL online version), and a virtual machine for shared remote team communication. This solution has been successfully applied in both educational and research contexts, supporting BVI users, educators, and researchers in performing complex tasks more efficiently. The following section explains how the support is implemented, based on an invented specific non-relational and natural text marking paradigm.

2 Methodology and Research Approach

As for System Description and Personalization Layer, WPad-BVI is a personalized, user-adapted system built on a three-component architecture designed to meet the needs of BVI users. These components include:

WPad-BVI Desktop Application WPad-BVI is Developed in Visual FoxPro 9 database platform (VFP9), the WPad-BVI desktop application features a table-like interface with seven short-text columns and one memo column for extended content. Unlike conventional databases, WPad tables are flexible: column names are not fixed, allowing storage of diverse data structures. Specific columns serve roles such as table identification, navigation between tables, categorization of record content, and storage of hypertext, code fragments (HTML, JavaScript, C++, C#, PHP, command line, any programming language), or any plain text. The table also functions as a framework for knowledge representation, similar to AI, as a model of virtual knowledge. It simulates human knowledge and its flow-how a user inserts tacit knowledge into the table. Simply said, virtual knowledge is the content stored in the memo field, and other columns indicate it as metadata or meta-information. Therefore, it is both a human- and machine-readable way of processing knowledge-based activities, such as educational processes.

2.1 Overall Architecture

Figure 1 shows the hybrid layout of WPad?BVI.

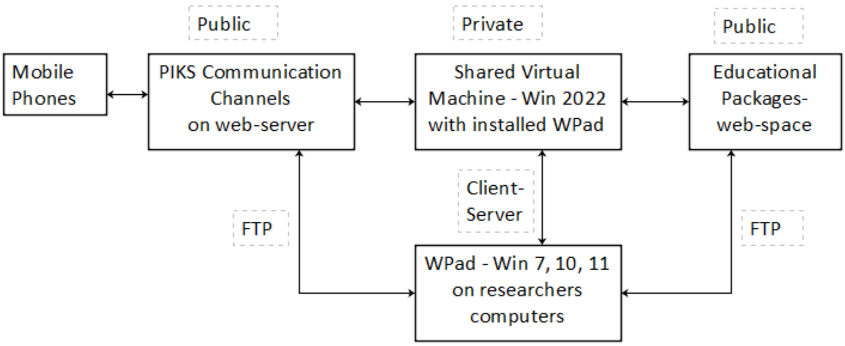


Figure 1: System architecture: (1) WPad desktop on Win7-11, (2) PIKS web channels (public), (3) Windows Server 2022 VM (private), (4) Web learning space (public), plus mobile phones connected to PIKS via FTP and client-server links.

This universal system supports operations like sorting by date or category, keyboard navigation, and HTML generation. All functions are accessible via a simplified menu and customizable hotkeys, tailored for BVI users. Dedicated accessibility features include over hundred possibilities, such as:

- Creating hybrid navigation tables
- Generating multi-row tables from personal know-how manually or automatically
- Creating educational packages from folders or current tables (Ctrl+F1)
- Linking to offline or online sources (e.g., folders, Google Docs, cloud platforms)
- Hybrid search across internet and local systems
- TTS support (e.g., Speak row - F2) with audio automation
- PDF packager for navigation tables to PDF libraries
- FTP transfer to PIKS or external databases
- Image automation: labeling, slide generation (in testing phase)

PIKS Online Channels PIKS (Personal Knowledge Interchange System) is a set of independent online channels accessible via standard web browsers. PIKS channels allow users to: chat, post, and retrieve information, collaborate in groups and transfer data between desktop WPad tables and online channels This ensures seamless integration between local and remote environments.

Virtual Machine Environment A Windows Server 2022 virtual machine hosts WPad, enabling remote access for team members. It allows for real-time collaboration and shared table editing. It allows also for WPad BVI design testing between a BVI IT teacher and the WPad designer, assisted by NVDA and Narrator This hybrid infrastructure supports both individual autonomy and group cooperation.

3 Results - Educational and Research Activities

Two primary activity types were carried out by a blind educator and researcher using WPad-BVI:

3.1 Didactic Activities

Preparation of teaching materials in LaTeX, PowerPoint, and Word, exported to PDF creation of educational packages using WPad tables and folder-based content Packages were published as websites using WPad and deployed via FTP. Courses were conducted in both Polish and English, across multiple educational levels. PIKS channels were used for classroom communication: Individual channels for each group of students, a collective channel for the teacher, uploading materials and assignments, and receiving student submissions, monitoring, and managing group activities centrally.

3.2 Collaborative Research

PIKS channels supported research teams by: Sharing publications, drafts, and resources Organizing subgroup-specific channels Enabling browser-based access without software installation The virtual machine allowed deeper collaboration: Remote access to WPad tables Editing and uploading materials Joint development of research papers This workflow led to the successful production of multiple academic publications.

3.3 Evaluation by verification of action research activities

WPad-BVI was validated through action research, emphasizing practical functionality over statistical metrics. Testing involved over 100 symbolic knowledge tables across diverse domains. Key validated features include:

- HTML generation, language translation, text-to-speech
- Keyboard-only navigation, hotkey activation, and screen reader compatibility
- FTP deployment to live servers

– Cognitive support via automation of repetitive tasks

Educational content processed included:

– University syllabi
– Multilingual course materials
– E-learning repositories
– audio and video tutorials
– educational multiformat outputs

Outputs were successfully deployed to websites and virtual environments.

3.4 HTML-Table Learning Package

Figure 2 illustrates a navigational HTML frameset generated by WPad. The WPad system

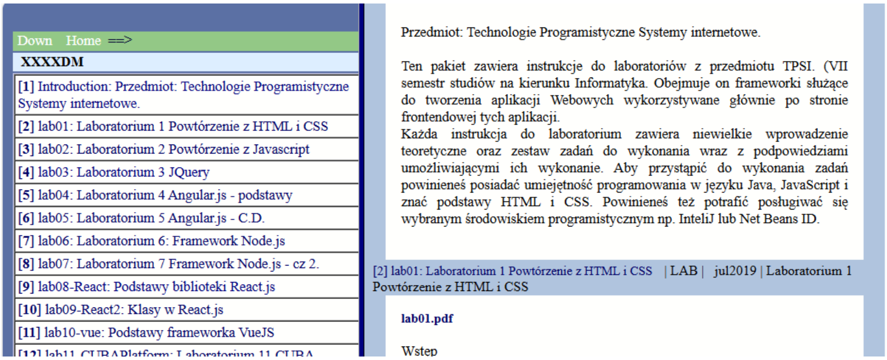


Figure 2: Example of a learning package as a navigational HTML table. The left pane lists topics (click a row to open the PDF in the right pane).

was developed within two research projects under the International Visegrad Fund (IVF) from 2018 to 2024 and continues as non-project research by teams from Slovakia, Poland, the Czech Republic, Hungary, and Ukraine. Earlier development at the Slovak University of Technology supported around 2,000 undergraduates and several diploma theses. Within the IVF project, around 100 students assisted in PIKS design and testing. The following table presents the activities of the IT BVI teacher as evidence that BVI support works in real teaching practice.

4 Discussion

The development and deployment of WPad-BVI highlight several themes in inclusive digital design. The WPad "Navigation" tables let users navigate to any environment, e.g.,creating multirow tables for GoogleDocs (used during COVID), video tutorials or for leisure purposes(YouTube or SoundCloud navigation). They can benefit from the added values of the WPad hybrid system, such as:

Table 1: BVI Teacher Outputs

Course	Lang.	Level	Stud.	Out.
Web tech.	PL	I grade	30	WPad
Distrib. apps	PL	II grade	15	WPad
Univ. Design	PL	I grade	8	WPad
Tech for BVI	PL	Teachers	15	WPad
Web tech.	EN	I grade	20	WPad
Music	PL	Elementary	22	PIKS
English lang.	PL	Elementary	10	PIKS
Mathematics	PL	Elementary	11	PIKS
Total: 8 courses, 2 languages, 4 levels, 151 students, 2 output types				

Accessibility as a Design Paradigm By embedding keyboard navigation, speech synthesis, and customizable menus at its core, WPad affirms that inclusive design benefits all users, not just those with disabilities.

Learning Curve and Cognitive Load Symbolic tables and extensive hotkeys require effort to master?akin to learning a musical instrument. Future work could introduce the following.

– Guided tutorials and contextual help
– Adaptive shortcut mapping
– Integrated training modules

Technical Integration and Screen Reader Compatibility NVDA vs. Narrator discrepancies in memo field handling reveal the fragility of accessibility across platforms and underscore the need for standardized APIs.

Collaborative Knowledge Work PIKS channels plus Virtual Machine access transform WPad into a team-oriented knowledge infrastructure-supporting asynchronous posting and real-time collaboration.

Future Development and Inclusive Assistive Technology Leveraging native Windows SAPI, WPad implements text-to-speech function in tiny VFP objects (10-250 chars), and visual alerts for deaf users, all with minimal overhead-outperforming bulkier C# approaches, including better integrating of Speech Recognition function of Windows (it works in PIKS channels when using Mobile Phones).

Toward a Universal Knowledge Interface WPadś dual-layer architecture (VFP9 tables + HTML tables) and symbolic logic model offer a future-proof, resilient interface paradigm beyond GUIs.

5 Conclusion

WPad functions as a scientifically grounded, practically validated system for inclusive digital environments. Its accessibility-first design and symbolic architecture offer a transparent, modifiable, and effective solution for knowledge-driven education. However, challenges such as incompatibility of NVDA screen reader with memo fields and the cognitive load of mastering symbolic navigation are remain. These are mitigated by streamlined menus and Narratorś integration. Summarising we can compare Using WPad to learning an instrument-that is initially complex, ultimately empowering. What sets WPad apart is its enduring compatibility. Its future is not constrained by technological shifts-it adapts through symbolic logic and use-defined extensions, functioning as a classic Windows component like Notepad, WordPad, or Command Line.

References

1. JAWS for Windows. Software, https://www.freedomscientific.com/products/software/jaws/, accessed on April 2, 2026
2. NVDA (nonvisual desktop access). Software, https://www.nvaccess.org/, accessed on April 2, 2026
3. Orca screen reader. Software, https://help.gnome.org/users/orca/stable/, accessed on April 2, 2026
4. TalkBack. Software, https://support.google.com/accessibility/android/answer/6007100?hl=pl, built-in Android accessibility feature. Accessed on April 2, 2026
5. VoiceOver. Software, https://www.apple.com/accessibility/mac/vision/, built-in macOS and iOS accessibility feature. Accessed on April 2, 2026
6. Cao, K.: Predicting the behavior intention of intelligent office software. SHS Web of Conferences (2024), https://api.semanticscholar.org/CorpusID:267083833
7. Devi, R., Kumar, S.: Digital right management and accessibility of libraries electronic resources for blind and visually impaired users. DESIDOC Journal of Library & Information Technology (2023), https://api.semanticscholar.org/CorpusID:260524858
8. Díaz, F.J., Schiavoni, A., Osorio, A., Amadeo, A.P., Harari, I., Ray, D.: Fixing accessibility issues in open-source teaching repositories. CLEI Electron. J. **20** (2017), https://api.semanticscholar.org/CorpusID:65173462
9. Gori, J., Han, H.L., Beaudouin-Lafon, M.: Fileweaver: Flexible file management with automatic dependency tracking. Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology (2020), https://api.semanticscholar.org/CorpusID:221897046
10. Kipke, S.: New concepts of pc interaction for blind users based on braille displays with atc technology (active tactile control). In: International Conference on Computers for Handicapped Persons (2010), https://api.semanticscholar.org/CorpusID:38410203
11. Lee, H.N., Ashok, V.G.: Impact of out-of-vocabulary words on the twitter experience of blind users. Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (2022), https://api.semanticscholar.org/CorpusID:248420202
12. Leslie Lamport: LaTeX. Software System, https://www.latex-project.org/, accessed on April 2, 2026
13. Pandey, Y.R., Lee, J., Banda, D.R., Griffin-Shirley, N., Nguyen, T., Othuon, V.O.: A survey of mobile app use among university students with visual impairment in india. British Journal of Visual Impairment **41**, 662 – 674 (2022), https://api.semanticscholar.org/CorpusID:260808475

Boopathi Palanisamy[1]
ORCID: 0000-0003-2646-501X

Dariusz Mikułowski[2]
ORCID: 0000-0002-4567-7846

[1] Central University of Tamil Nadu,
Department of English Studies,
Thiruvarur – 610 005, Tamil Nadu, India

[2] University of Siedlce
Faculty of Exact and Natural Sciences
Institute of Computer Science
ul. 3 Maja 54, 08-110 Siedlce, Poland

pboopathi@cutn.ac.in, dariusz.mikulowski@uws.edu.pl

Uncovering Posthuman Interdependence: Accessible Computing and Transformative Potential of Artificial Intelligence

Abstract. The inexorable space of posthumanism that we inhabit in our everyday activities by getting our tasks done through AI, digital devices and technology, arguably, makes us ponder about the interdependent nature of human beings. With the arrival of assistive technology and screen-readers from the late 1990s onwards, people with disabilities began relying on computer technologies for their everyday life, education, employment, etc. This paper will delineate how the rapid development in assistive technology has led to the idea of accessible computing, a phrase became very popular globally since the last decade in the context of disability inclusion. Further, it will also discuss the role of AI and its transformative potential in the contemporary times, exemplifying its indispensable place in all walks of lives. While highlighting several transformative potentials of AI, it will shed light on the counter effect it is bound to have in lifestyle, education, employment, research, etc. in the future. The paper will end by throwing light on how accessible computing and AI contribute to posthuman condition of interdependence in the lives of both disabled and able-bodied people alike.

Keywords: Assistive Technology, Accessible Computing, Transformative Potential of AI, Posthuman Interdependence, Disability Inclusion

1 Introduction

Artificial Intelligence (AI) creates a colossal impact in the lives of people with all abilities, marking a new beginning in the history of human life. While the history of human race is fraught with several revolutions, and many socio-cultural, economic and technological conditions, the life in Twenty-first Century, notably enough, is replete with the domination of computer technology and the recent development of AI largely. Came as a great enabler to human beings, AI has seeped into all walks of human life, making considerable intervention in essential areas, such as education, healthcare, information technology, content creation, among others. Evidently, in the lives of people with disabilities, AI is often touted to be a great transformer, for it substantially reduces their dependency on others in carrying out their daily routine. Besides significantly reducing the dependency, it enables people with disabilities to perform many tasks on computer and mobile phone, which were hitherto impossible for them due to inaccessibility of softwares and mobile applications. The idea of Assistive Technology, a phrase quite often associated with people with disabilities, in fact, got a new meaning and great potential after the introduction of AI in transforming their lives drastically.

Emerged as a major critique of Humanism and its core idea of independence in 1990s, placing able-bodied human as the centre of universe, Posthumanism, a radical theoretical framework, tries to maintain a synthesis between human and non-human entities like technology, nature, animals, among others. Developed as a branch of Critical Posthuman Studies, Posthuman Disability Studies, started by Goodley et all in their article "Posthuman Disability Studies," [7] attempts to advance the idea that disability is a posthuman condition as it deconstructs the idea of independent human life. Further, the notion of Assistive Technology for people with disabilities, notably, falls within the ambit of Posthumanism, for it aids them to rely on technology, mobility aids, communicational devices, guide dogs and others for their everyday activities, thereby drastically reducing their dependence on human beings. This paper discusses how the rise of assistive technology has popularized the concept of accessible computing in the last decade, particularly in relation to disability inclusion. It also examines the transformative role of AI in contemporary times for people with all abilities, while discussing its potential counter effects in areas like lifestyle, education, and employment. Finally, the paper establishes the argument that both accessible computing and AI contribute to a posthuman condition of interdependence for everyone, regardless of ability.

2 Introduction to Posthumanism and the Idea of Interdependence

Emerged as a result of Postmodern emphasis on diversity and decentring of single idea of human, Posthumanism promotes a new way of thinking about human, wherein the intersection of human, non-human entities and technology is widely acknowledged as a new mode of life. It is a strong critique of humanism, a human centric ideology, often approached through Social Studies of Technology, Disability Studies and Cultural Studies broadly. A firm philosophical and cultural movement, Posthumanism seeks to uncover the relationship between human beings and non-human entities, such as environment, nature, technology and animals by fostering a sense of inclusion and interdependence. As Donna Haraway argues in her off quoted book, A Cyborg Manifesto (1985) [9], human beings are essentially a cyborg, a fusion of biology and technology, completely rejecting the idea of only biological aspects of human identity. The entanglement of human and technology as termed by Haraway as cyborg, in fact, deconstructs the binary power relation between human VS non-human, culture VS nature; mind VS body; male VS female; etc., which are central to human lives. Building on this idea, Posthumanism, as stated earlier, questions the notion of anthropocentrism which treats human as the center of universe and thus inherently endowed with the right to dominate nature and other non-human entities. Posthumanism seeks to create ethical and sustainable relationship with non-human entities and strongly opposes the idea of speciesism (superiority of human) that promotes discriminations against non-human beings. In the contemporary times, the singular notion of major social identities, such as gender, sexuality, social relations, family and disability have been widely understood through the framework of Posthumanism. A founding scholar of Posthumanism, Rosi Braidotti [2] advocates for non-hierarchical relationship among human beings, trying to maintain equality and respect among all citizens. Further, Katherine Hayles[10], yet another scholar of Posthumanism, in her book, How We Became Posthuman (1999), states that, after the onset of Posthumanism, our understanding of our selves has shifted from embodied beings to that of pattern of information. Arguably, Posthumanism while promoting AI in carrying out tasks related to education, employment, research and daily routine, makes people aware of the ethical question in connection to law, human life and social order in using AI. Evidently, In the lives of people with disabilities, Posthumanism plays a vital role in realizing the idea of interdependence through the interaction of human and technology. This act of interdependence, we argue, is further promoted by AI through enabling people with disabilities to effectively carry out their everyday tasks without relying on others. In Posthuman condition, Human beings are not isolated according to the traditional thinking pattern, rather, they are placed within complex networks and social systems that condition their existence. The society that is constituted out of Posthuman condition thus inherently inclusive, one which promotes the idea of interdependence of human beings. Thus, the shift, in fact, signals the move from independence to interdependence of human lives, significantly promoting a new way of life. In the context of people with disabilities, as stated earlier, Posthuman condition, argued by Dan Goodley et al [7], has always been part of their existence, making disabled lives as interdependence. People with disabilities, excluded by Ableist ideology of promoting independence, constitute their lives as interdependent by relying on animals, assistive aids and technology. With the notion of new-normal, an idea became popular during Covid pandemic, the idea of interdependence with the help of advance technology and AI is widely promoted even among able-bodied people in the post-Covid period throughout the world. The idea of posthuman interdependence advanced through AI in the recent times for people with all abilities becomes thus a life in new-normal, relying on technology and AI. The idea of accessible computing gains greater significance in the era of AI as it provides opportunity particularly for people with disabilities to use AI more effectively with the help of computer technology for their everyday and professional usages.

3 Accessible Computing in the Age of Posthumanism

Accessible Computing refers to designing technologies based on universal standard that can be effectively and inclusively used by both people with different abilities and disabilities [17]. It denotes making computer applications, technological devices, websites and digital spaces accessible, thereby making the experience of using such things less challenging for everyone, specifically people with disabilities. Major accessibility features include: text-to-speech, closed-captioning, contrast color and large fonts on digital spaces, web-accessibility tools, keyboard shortcuts, etc. Notably, accessible computing is more challenging for people with cognitive or learning disabilities, [23] severe physical disabilities, blindness and hearing impaired [11] due to lack of features on the websites of many organizations and certain needed facilities in the common technological devices available in the market.

In the legal realm, section 508 of American Rehabilitation Act of 1973 and its revised version in 2018 [18] require that all electronic content generated by US Federal agencies must conform to Level A and Level AA success criteria in WCAG 2.0 in order to enable people with disabilities to use computer technology without any barrier. The European Union regulations, such as Directive (EU) 2016/2102 [5] and the European Accessibility Act (Directive 2019/882), [6], establish the legal framework for ensuring digital accessibility. While the former focuses on public sector bodies, the latter, however, expands these requirements to products and services in the private sector. Both directives are based on the international WCAG 2.1 guidelines [22], which define standards for perceivability, operability, understandability, and robustness. The implementation of these regulations was carried out in EU countries to varying degrees and manner. For example, in Poland it was implemented through two regulations namely act of 2019 on the availability of digital public entities, [19] and the Act on ensuring the availability of goods and services from June 2025 [20]. In India, the most significant and comprehensive legislation on disability rights, which mandates accessibility in the physical environment, transportation, information, communication, and other facilities and services is "The Rights of Persons with Disabilities Act, 2016 [16]. There are also specific practical documents recommended for Indian entities such as Harmonised Guidelines and Standards for Universal Accessibility in India, 2021 [15] and IS 17802: Accessibility for ICT Products and Services Standard: published by Bureau of Indian Standards (BIS) [3]. Further, the accessibility of information and communication technology is ensured through the wider movement of Accessible India Campaign 2015 [1] by the Government of India, making all government and public websides accessible for people with disabilities. All these regulations aim to ensure and guarantee that applications and websites are fully accessible for all users, including those with disabilities. Equal opportunities for people with disabilities provided by disability legislations throughout the world has necessitated accessible computing is a basic right of every disabled person. Accordingly, International Standards, such as iso9241-171 [13], provides ergonomics guidance and specifications for the design of accessible software for public use.

In the world of interdependence created by Posthuman conditions, Accessible Computing plays a major role in enabling people with different abilities and disabilities to use digital spaces and technological devices without any barriers. With the help of computer technology, people with disabilities and people with less abilities, however, navigate digital spaces and carry out their daily works without any human assistance. Evidently, technologies like AI, AR/VR, and assistive devices can remove these barriers, enabling individuals to interact with the world more fully. Inclusive design practices, such as adaptive interfaces and personalized user models, are essential to make these technologies accessible to a broader audience regardless of abilities or disabilities. For people with blindness, AI is helping them drastically to carry out their daily works through mobile applications. Many AI applications describe the image and explain the surrounding more precisely. For instance, the mobile applications such as Microsoft's Seeing AI or Be My Eyes are able to read text aloud, identify the products, recognize the people, identifies the color of dress and currency notes, among others. Additionally, the screen-readers for mobile devices such as Voice Over or Talkback have been well enhanced with advance features with the support of AI. For people with deafness, AI is greatly helping them to get real-time transcription and live captioning of online lectures, online meetings on Gmeet, Zoom and other platforms. One of the popular tool, which can be effectively used by the deaf community is Otter.AI. Interestingly, for them, AI provides sound recognition of various sounds like fire alarm, a doorbell and a baby crying. For people with learning disabilities, AI can create personalized educational content and simplify complex text, particularly making it more digestible for people with dyslexia or other learning disabilities [23].

4 Transformative Potential of Artificial Intelligence

Sundar Pichai, CEO of Google and Alphabet views artificial intelligence (AI) as a transformative technology of greater importance to humanity. He believes AI is more significant than foundational discoveries like fire and electricity. While he acknowledges its potential for tremendous positive benefits, he also warns that it carries substantial negative consequences that need to be considered. The transformative potential of AI is often identified with "human-level AI" (HLAI); "high-level machine intelligence" (HLMI); and "artificial general intelligence" (AGI). As many scholars have argued, AI does not describe any one specific application, but rather it refers to a set of computational techniques with the objective of enabling machines to behave intelligently. While Facebook took years to reach major user milestones, ChatGPT which is an important AI technology achieved them in a matter of months.

The following can be surmised as the major breakthrough with the support of AI technology: Healthcare: AI has the potential to drastically reduce or even eliminate medical errors, such as misdiagnosed illnesses or prescription mistakes. For example, a tool called DXplain can help prevent the quarter-million deaths that result from medical errors in the U.S. each year. Law: In the legal field, AI uses predictive analytics to analyze past cases and outcomes, helping lawyers anticipate the likely results of a new case. Retail: Retailers are already using AI for everything from predicting sales trends to optimizing inventory. Major companies like Kroger and Walmart are also investing in self-driving vehicles, which have been shown to cut logistics costs by up to 30%.

In the field of Education, AI helps for Personalized Learning based on the level of students. It automates routine tasks such as grading, attendance tracking and scheduling, thereby reducing the administrative burden on teachers. Furthermore, it provides real-time, objective feedback to students, helping them understand their mistakes and improve their learning strategies. AI-driven assistive technologies aid students with disabilities by offering customized resources and interventions, promoting an inclusive educational setting. In research, AI analyzes large datasets to uncover patterns and trends, facilitating more accurate and timely discoveries across various disciplines. AI tool such as Google NotebookLM helps researchers to summarize their research works, helps for notetaking, convert their research works into podcast for wider reach, etc. AI can assist not only with simple tasks, but also supports engineers working in the exact sciences. By providing appropriate prompts with Chat GPT or Gemini, in the form of scripts (e.g., in Python), a big data analyst can run machine learning models or perform data aggregation. Similarly, programmers can use tools code for programmers such as Chat for Code and others [4] like IDE extensions, such as those for Visual Studio or copilot [21] to generate code snippets, debug, and verify code correctness. For researchers with disabilities, particularly with blindness, AI tools help read the printed and handwritten materials in the Realtime and translate them into any language of their choice. Examples of such solutions are the already mentioned Be my Eyes which also works in the Windows version or an Instant Translate add-on for the NVDA screenreader.

On ethical level, in the field of education and research, AI systems can perpetuate existing biases if trained on skewed data, leading to unfair outcomes for marginalized communities. The use of AI in education involves the collection of sensitive student data, thereby paving way for misuse. Notably enough, more dependence on AI tools may reduce human interaction in educational settings, potentially hindering the development of critical thinking, problem solving ability and social skills among students. Further, in research, AI has the potential to generate false data, leading to wrong findings [14]. It provides misinformation of titles and authors of research works, thus resulting in wrong citations. It has the high potential to create plagiarism, in many cases, which sadly cannot be deducted through available tools. The urge to use AI excessively in research, however, leads to copyright infringement, which has already been cautioned by experts.

4.1 Use of AI in the Lives of People with Blindness

AI has radically transformed the lives of people with blindness by providing better accessibility and efficiency. Developments in robotics have led to the creation of AI-powered guide robots that assist visually impaired individuals in navigating complex environments like airports, hospitals, and malls [8]. These robots analyze surroundings, avoid obstacles, and efficiently guide users to their destinations. AI-powered speech-to-text software i.e. build into smart glasses [12] allows individuals with blindness to write, study, and communicate socially by converting natural speech into text with high accuracy. AI-powered standard smart home devices, such as smart thermostats, security cameras, and lighting systems, can be controlled through voice commands, allowing blind individuals to manage their living spaces independently.

5 Conclusion

The integration of diverse aspects such as technology, non-human objects, etc. promoted by the Posthuman condition makes people with different abilities and disabilities live interdependently. The marked shift in the lives of human beings from independence to interdependence signals the era of Posthumanism, a life condition thrusts emphasis on cyborg. The tools like Accessible Computing and AI play a major role in the Posthuman world, making human lives more interdependent on technology. In the inexorable entanglement of computer technology and AI, it is crucial for people with all abilities and disabilities to be well-integrated into the technologically-driven world of Posthumanism. It can only be possible, if we realize the meaning of inclusion not only in the built environment, but also in the digital spaces, common technological devices, and softwares and tools used for education, research, works spaces and every day works. Accessibility of computer technology and tools of AI, at the time when millions of people use excessively, is not a choice for people with disabilities, rather it is a basic right to be a human.

References

1. Accessible India Campaign: Accessible India Campaign (Sugamya Bharat Abhiyan). `https://depwd.gov.in/en/accessible-india-campaign/` (2025), last access: 10 sep. 2025
2. Braidotti, R.: The Posthuman. Polity Press, Cambridge (2013)
3. Bureau of Indian Standards: Accessibility for ICT Products and Services (2021)
4. Chen, V., Zhu, A., Zhao, S., Mozannar, H., Sontag, D., Talwalkar, A.: Need help? designing proactive ai assistants for programming. Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (2024), `https://api.semanticscholar.org/CorpusID:273186766`
5. European Union: Directive (eu) 2016/2102 of the european parliament and of the council on the accessibility of the websites and mobile applications of public sector bodies. Official Journal of the European Union, L 327/1 (2016), `https://eur-lex.europa.eu/eli/dir/2016/2102/oj`
6. European Union: Directive (eu) 2019/882 of the european parliament and of the council of 17 april 2019 on the accessibility requirements for products and services. Official Journal of the European Union, L 151/70 (2019), `https://eur-lex.europa.eu/eli/dir/2019/882/oj`
7. Goodley, D., Lawthomb, R., Runswick Coleb, K.: Posthuman disability studies. Subjectivity pp. 1–20 (2014)
8. Guerreiro, J.P.V., Sato, D., Asakawa, S., Dong, H., Kitani, K.M., Asakawa, C.: Cabot: Designing and evaluating an autonomous navigation robot for blind people. Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility (2019), `https://api.semanticscholar.org/CorpusID:204917508`
9. Haraway, D.J.: A cyborg manifesto: Science, technology, and socialist-feminism in the late twentieth century. In: Simians, Cyborgs, and Women: The Reinvention of Nature, pp. 149–81. Routledge, New York (1991)
10. Hayles, N.K.: How We Became Posthuman. Virtual Bodies in Cybernetics, Literature, and Informatics. University of Chicago Press, Illinois (1999)
11. Huenerfauth, M.: Human-computer interaction and automatic text simplification: Understanding the perspective of deaf and hard of hearing users. CTTS@SEPLN (2021), abstract
12. indianexpress.com: Ally Solos: These smart glasses use AI to help low-vision users see the world around them, `https://indianexpress.com/article/technology/ally-solos-these-smart-glasses-use-ai-to-help-low-vision-users-see-the-world-around-them-10196149/`
13. International Organization for Standardization: Ergonomics of human-system interaction – part 171: Guidance on software accessibility (2008), `https://www.iso.org/standard/38079.html`
14. Kooli, C.: Chatbots in Education and Research: A Critical Examination of Ethical Implications and Solutions. Sustainability **15**(7), 5614 (2023). `https://doi.org/10.3390/su15075614`
15. Ministry of Housing and Urban Affairs: Harmonised Guidelines and Standards for Universal Accessibility in India (2021)

16. The Rights of Persons with Disabilities Act, 2016. Gazette of India, Extraordinary, Part II, Section 1, No. 49 (2016), Indie
17. Storer, K.M., Branham, S.M.: Deinstitutionalizing independence: Discourses of disability and housing in accessible computing. Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility (2021), https://api.semanticscholar.org/CorpusID:239011858
18. U.S. Congress: Section 508 of the rehabilitation act (1973), https://www.section508.gov/manage/laws-and-policies/section-508-law/, as amended, specifically by the Workforce Investment Act of 1998
19. Ustawa z dnia 4 kwietnia 2019 r. o dostępności cyfrowej stron internetowych i aplikacji mobilnych podmiotów publicznych. Dz. U. 2019 poz. 848 (2019), Polska
20. Ustawa z dnia 26 kwietnia 2024 r. o zapewnianiu spełniania wymagañ dostępności niektórych produktów i usług przez podmioty gospodarcze. Dz. U. 2024 poz. 731 (2024), Polska
21. Vaithilingam, P., Glassman, E.L., Groenwegen, P., Gulwani, S., Henley, A.Z., Malpani, R., Pugh, D., Radhakrishna, A., Soares, G., Wang, J.Y., Yim, A.: Towards more effective ai-assisted programming: A systematic design exploration to improve visual studio intellicode's user experience. 2023 IEEE/ACM 45th International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP) pp. 185–195 (2023), https://api.semanticscholar.org/CorpusID:259290004
22. W3C Web Accessibility Initiative (WAI): Web content accessibility guidelines (wcag) 2.1. W3C Recommendation (Jun 2018), https://www.w3.org/TR/WCAG21/, https://www.w3.org/TR/WCAG21/
23. Waggoner, T., Jose, J.A., Nair, A.S., Manikandan, S.: Inclusive design: Accessibility settings for people with cognitive disabilities. ArXiv **abs/2110.05688** (2021), https://api.semanticscholar.org/CorpusID:238634761

Tomas Sladcik
ORCID: 0009-0009-2131-3755

Michal Janosek
ORCID: 0000-0001-9439-4706

University of Ostrava
Dept. of Informatics and Computers
st. 30. dubna 22, Ostrava, Czechia

tomas.sladcik@osu.cz, michal.janosek@osu.cz

Agent-Based Reinforcement Learning for Smart Heating

Abstract. This paper presents a reinforcement learning-based heating control system for smart homes, utilizing Q-learning to optimize gas boiler operation. The system employs a single agent that evaluates multiple room states independently and determines appropriate heating actions. Each room state is represented by a vector composed of the target, current indoor, and outdoor temperatures. The agent uses an epsilon-greedy strategy to balance exploration and exploitation, with a five-minute delay to retrospectively assess the effect of each action. Rewards or penalties are assigned based on heating efficiency, temperature trends, and proximity to the target temperature. The system is implemented entirely in Node-RED without external machine learning libraries, ensuring transparency and low complexity. The experimental setup includes three thermostatic heads and one outdoor sensor, integrated via Home Assistant. A supervisory component aggregates the individual decisions and determines whether the gas boiler should be activated based on weighted priorities. The architecture is scalable with respect to the number of rooms, lightweight in terms of system requirements, and shows potential for integration into real-world smart home environments.

Keywords: Smart heating, Agent, Home Assistant, Node-RED, RaspberryPi, reinforcement learning.

1 Introduction

The effective control of gas boilers in smart homes remains a challenge despite the increasing adoption of smart thermostats and zone control systems. Although current solutions use advanced sensors, predictive models, and fuzzy logic, decision-making mechanisms often operate in isolation - without a comprehensive overview of the global state of the home.

A typical example is when individual thermostats generate heating demands based on local deviations from the target temperature, without considering outdoor conditions, energy priorities, or the current state of other zones. This can lead to frequent and inefficient boiler switching, excessive cycling, and unnecessary energy consumption [1]. In recent years, approaches have emerged that use fuzzy control within multi-agent systems [1], model predictive control (MPC) for output temperature stabilization [2], machine learning for consumption prediction and operation control [3], or dynamic matrix control (DMC) integrated in an IoT architecture controlled by a microcontroller [4]. However, there is still potential for extending these systems with adaptive behavior, the ability to learn from previous states and decisions.

The solution promotes the deployment of a reactive agent embedded in the home environment and the autonomous adjustment of environmental conditions according to predefined goals [5].

The aim of this work is to design and implement a reactive agent using Q-learning, which will be able to switch a simple gas boiler controlled by two states (On/Off) and optimize energy consumption while maintaining thermal comfort. The Q-learning method is probably the most widely used reinforcement learning technique. The advantage of this method is its ability to be deployed immediately without any pre-trained data. The agent initially has zero value functions and chooses actions at random (trial and error). However, by continuous learning and reward in the form of reward or punishment, its precision gradually improves with each passing pass [5].

2 Experimental part

This section describes both the hardware and software parts of the agent for receiving data from sensors, processing them, and then selecting the output (logical state Off/On). The aim of the experiment is to use commonly available resources that an ordinary user with basic technical knowledge can deploy in their newly created or existing smart home.

For clarity, this chapter is divided into three subchapters. The first presents the hardware used. The second focuses on software architecture. Finally, the third introduces the algorithm based on reactive agent learning.

2.1 Hardware

The experimental setup consists of the following components:

- Raspberry Pi 5 microcomputer (4 GB of RAM), serving as the agent's computing unit.
- The Aqara Radiator Thermostat E1, used to measure both the current and the target indoor temperature.
- SwitchBot Thermo/Hygrometer (Indoor/Outdoor), used to measure outdoor temperature and humidity.

One of the main advantages of these thermostatic heads is their built-in firmware, which autonomously regulates the flow of hot water according to the difference between the current and the target temperature.

In addition, they communicate with the central unit through a Zigbee network and provide valuable data to the system - such as current and target temperature, heating status (on/off), and more - which can be further processed and used to create various automation scenarios.

The SwitchBot Thermo/Hygrometer was selected for its simplicity and suitability for outdoor placement. In addition to measuring temperature, it also monitors humidity, which could in the future help more accurately determine whether the weather is dry or rainy.

2.2 Software architecture

The open-source Home Assistant (HA) platform was installed on RaspberryPi, which provides integration of sensors and actuators via standard protocols (Zigbee, Z-Wave, Wi-Fi, Bluetooth).

The agent's decision-making logic is provided by a Node-RED tool, which is widely used in home and industrial automation. It makes it easy to create scenarios using visual programming based on the principle of data flows (flow-based programming).

The agent itself is responsible solely for deciding whether to activate the gas boiler. It does not control the water flow through individual radiators, as this is handled autonomously by the thermostatic heads.

2.3 Reactive agent learning

This section outlines the architecture of the decision-making process, including sensor data aggregation, state representation, learning strategy, and action selection.

The experiment is based on a reactive agent that interacts with the environment in the house through thermostatic heads and the outdoor temperature. Monitors the environment and, based on the data obtained, decides to turn on the gas boiler. For each selected action, it receives a reward or punishment.

The agent uses reinforcement learning (Q-learning) to refine its performance. The Q-values are updated using the standard Bellman update rule:

$$Q(s,a) = Q(s,a) + alpha[r + gamma * \max(Q(s',a')) - Q(s,a)] \quad (1)$$

where $Q(s,a)$ denotes the expected reward for taking action a in state s, alpha is the learning rate, gamma is the discount factor, and r is the delayed reward. In this implementation, the reward is calculated retrospectively based on real-time temperature measurements and heating effectiveness (see Section 3). This method was chosen for its simplicity and ease of implementation in Node-RED, without relying on any external libraries.

The agent stores each action in a state vector formatted as

$$target\_temp - actual\_temp - out\_temp \quad (2)$$

along with an evaluation indicating whether the action led to a positive (reward) or negative (punishment) outcome.

The Q-table vector format is designed for a single heating unit. When multiple units are present, the table structure remains unchanged, but decisions are made separately for each thermostatic head.

In this experiment, a total of three thermostatic heads are used, resulting in three independent decisions that feed into the control of a single boiler.

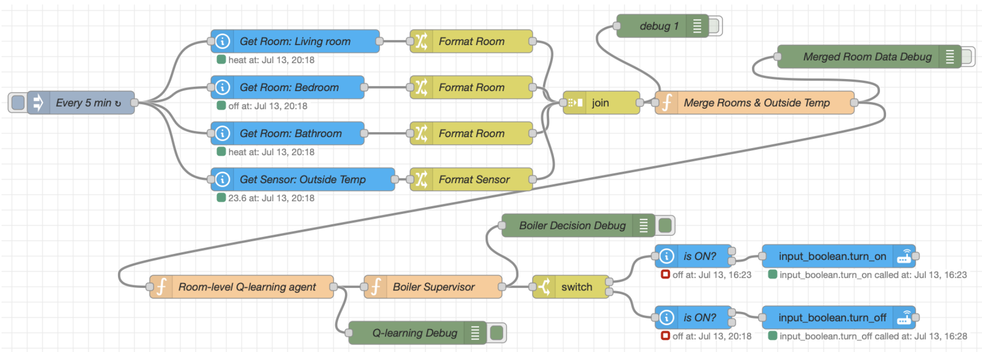


Figure 1: Node-RED flow diagram

The agent is triggered every 5 minutes via an Inject node. In the first step, it reads the current states from all input sensors:

- Three thermostatic heads of the Aqara Radiator Thermostat E1 (current and desired indoor temperature).
- One Switchbot Thermo/Hygrometer Indoor/Outdoor sensor (current outdoor temperature).

In the Format Room node, relevant data - including current and target temperature, room name, and priority - are extracted from the sensors mentioned above. To simplify data handling and ensure scalability, all inputs are joined into an array of objects using the Join node. This array also includes the current outdoor temperature. The Merge Rooms & Outside Temp function node then assigns the current outdoor temperature to each individual room object. The result is an array of enriched room objects, ready for further processing by the agent. The Room-level Q-learning agent node functions as an adaptive decision-making component that selects the On or Off action based on the current state of each room - defined by the target temperature, current indoor temperature and outdoor temperature.

For every room, it constructs a unique state vector and applies a epsilon-greedy strategy to choose between exploration (random selection) and exploitation (selecting the action with the highest Q-value). The decision-making process is further influenced by the temperature trend, which is derived by linear regression from recent measurements. The selected action is then logged in the history for future evaluation.

The learning is performed retrospectively with a delay of five minutes. Once the defined interval has elapsed since the action was executed, its impact on the room temperature is assessed, and a corresponding reward or penalty is calculated.

```
if (pastEntry.action === 'on') {
    // Penalty for overheating
    if (pastEntry.temp_in > pastEntry.target + 0.4) {
        reward -= 10;
    }
    // Heating efficiency rating
    if (tempChange > 0.2) {
        reward += 10 * (tempChange / timeDiffMinutes);
    } else if (tempChange < -0.1) {
        reward -= 8; // marné topení
    }
} else { // action === 'off'
    // Reward for correct shutdown in case of overheating
    if (pastEntry.temp_in > pastEntry.target + 0.5) {
        reward += 3;
    }
    // Evaluation of correct heating
    if (tempChange < -0.1) {
        reward += 5; // správné netopení
    } else if (tempChange > 0.2) {
        reward -= 3; // mělo se topit
    }
}
```

Figure 2: Reward system

The final node, Boiler Supervisor, evaluates all decisions made by the room-level agents. Calculate a weighted score based on the priority of each room and determine whether the gas boiler should be activated.

```
//  Go through all the results and calculate the weighted score
for (let result of qLearningResults) {
    let priority = result.priority;
    let weight = priority; // directly use priority as weight

    // Initialize priority statistics if they do not exist
    if (!priorityStats[priority]) {
        priorityStats[priority] = { on: 0, off: 0 };
    }

    // Add to weighted score
    if (result.action === 'on') {
        weightedScore += weight;
        priorityStats[priority].on++;
    } else {
        weightedScore -= weight;
        priorityStats[priority].off++;
    }
}
```

Figure 3: Weight score of Boiler Supervisor

3 Results

The experiment is currently in test mode. The agent was trained in three phases: basic scenarios, daily simulations, and extreme situations, using synthetic temperature data (see Section 2.3).

Because the system was developed in summer, when heating is normally inactive, a virtual switch in Home Assistant was used to simulate boiler control.

Preliminary results show that heating stays off most of the time. It is switched on only if at least two of three thermostatic heads detect a temperature 0.5 °C below the target. If only one head reports this difference, heating remains off—except in high-priority rooms, where it is activated.

4 Conclusions and future work

This paper presented a lightweight and transparent reinforcement-learning approach to smart heating control. Using a single Node-RED agent, the system is able to make room-level decisions based on simple state vectors derived from temperature readings. The architecture requires no pre-trained models or external libraries, making it suitable for low-power smart home environments.

Although the system was tested under non-heating conditions during the summer, training and evaluation were based on simulated scenarios that covered a broad range of temperature dynamics. These scenarios allowed the agent to learn how to respond to various heating demands in a controlled manner. Furthermore, same evaluation metrics are suitable for real-world deployment as well. Where available, variables related to concomitant systems and activities may augment the state vector. As a result, the learned Q-values are expected to generalize to real-world heating situations once deployed during the actual heating season. The supervisory decision scales linearly with the number of rooms, so scaling to larger setups should not add significant overhead.

Future work will focus on real-world deployment during the heating season, dynamic adjustment of reward parameters, comparison with classical rule-based control strategies, and the implementation of the system as a Home Assistant add-on to improve usability.

References

1. Jimenez-Bravo, D.M., Lozano Murciego, A., De La Iglesia, D.H., De Paz, J.F., Villarrubia Gonzalez, G.: Central heating cost optimization for smart-homes with fuzzy logic and a multi-agent architecture. *Applied Sciences* **10**(12), 4057 (2020). `https://doi.org/10.3390/app10124057`
2. Conceicao, C., Quinta, A., Ferreira, J.A.F., Martins, N., Santos, M.P.S.D.: Embedded model predictive control of tankless gas water heaters to enhance users' comfort. *Machines* **11**(10), 951 (2023). `https://doi.org/10.3390/machines11100951`
3. Tsoumalis, G.I., Bampos, Z.N., Chatzis, G.V., Biskas, P.N.: Overview of natural gas boiler optimization technologies and potential applications on gas load balancing services. *Energies* **15**(22), 8461 (2022). `https://doi.org/10.3390/en15228461`
4. Su, Y.: An intelligent heating system based on the Internet of Things and STM32 microcontroller. *Energy Informatics* **7**(1), 22 (2024). `https://doi.org/10.1186/s42162-024-00326-2`
5. Kubik, A.: *Intelligent Agents: Software Development Based on Multi-Agent Systems*. 1st edn. Computer Press, Brno (2004)

Marek Pilski
ORCID: 0000-0003-2362-0416

University of Siedlce
Faculty of Exact and Natural Sciences
Institute of Computer Science
ul. 3 Maja 54, 08-110 Siedlce, Poland

marek.pilski@uws.edu.pl

A Comparative Analysis of Key Features of WCAG-based Website Digital Assessment Methodologies

Abstract. A website accessibility assessment methodology is a process that guides an auditor through a series of steps to determine whether a website complies with the guidelines (WCAG). This article presents a comparative analysis of key features of three prominent web accessibility evaluation methodologies: WCAG-EM (Website Accessibility Conformance Evaluation Methodology) developed by W3C, RGAA (Référentiel Général d'Amélioration de l'Accessibilité) it is mandatory for public institutions in France, and the Polish Web Accessibility Evaluation Methodology designed for the Polish public sector. Each of these frameworks offers a distinct approach to assessing the digital accessibility of websites, reflecting different legal, technical, and organizational contexts.The article compares these methodologies in terms of scope, structure, legal integration, level of detail, testing procedures, and practical applicability. The analysis highlights their respective strengths and limitations, offering insights into how evaluation strategies can be tailored to national contexts while maintaining alignment with international accessibility standards. Finally, a comparative analysis of selected features of the discussed methodologies was performed.

Keywords: WCAG, WCAG-EM, RGAA, Polish Web Accessibility Evaluation Methodology.

1 Introduction

Digital accessibility refers to the design and development of websites, applications, and other digital tools in a way that ensures usability for individuals with various impairments, including visual, auditory, motor, and cognitive limitations. It involves creating digital solutions that are clear, understandable, and operable by the majority of users, while also supporting assistive technologies used by individuals with disabilities.

This concept encompasses all components of a digital product - source code, user interface, and content. The barriers users encounter typically stem not from their impairments but from poorly implemented accessibility standards WCAG (Web Content Accessibility Guidelines) [2]. Importantly, digital accessibility, while essential for some, offers significant benefits for all users. For example, captions enhance video accessibility in noisy environments, high-contrast designs improve readability for the elderly, and voice control facilitates hands-free operation for drivers.

In Poland, public institutions are legally mandated to ensure digital accessibility, as stipulated by the 2019 Act on the Digital Accessibility of Public Entities' Websites and Mobile Applications [1]. Like Poland, other countries are also implementing their own legal acts that impose the obligation to make digital products available. Implementing accessibility measures broadens the reach of public services and enhances user experience across devices. However, compliance with technical guidelines such as WCAG does not, in itself, guarantee accessibility. Rather, digital accessibility should be understood as an ongoing process that requires strategic planning, institutional commitment, cross-functional collaboration and its reliable assessment.

The evaluation of digital accessibility of websites is typically conducted by an auditor. This assessment should be based on content compliance guidelines aligned with recognized standards. The Web Content Accessibility Guidelines (WCAG), published by the World Wide Web Consortium (W3C), constitute the most widely cited set of digital accessibility standards. However, demonstrating WCAG compliance requires independent verification. Such evaluation is often complex and time-consuming, primarily due to the structural and functional complexity of modern websites. These platforms may include a wide array of sophisticated features, layouts, hierarchies, styles, dynamically changing content, forms, and more. Furthermore, from a technical standpoint, any website may claim compliance with WCAG while failing to meet its actual checkpoints, as auditors may employ different methodologies for content assessment. To address this issue, various organizations have undertaken efforts to develop standardized approaches for reliable evaluation of digital accessibility. These efforts have led to the emergence of digital accessibility evaluation methodologies, which are discussed in this article.

Such methodologies are valuable resources for initiating structured assessments of digital accessibility. They typically define a consistent process that auditors should follow, including the target compliance level (e.g., WCAG A, AA, or AAA), the scope of the evaluation, and the selection of web pages to be reviewed (either all or only key pages). These frameworks often incorporate checklists, guidance on conducting automated and manual tests, best practices, and recommendations. They also suggest the use of automated tools to streamline the otherwise labor-intensive evaluation process, as well as the involvement of persons with disabilities in the assessment process. Additionally, they provide templates for documentation, reporting, and accessibility declarations.

2 Methodologies for assessing the digital accessibility of websites

The following section presents three significant methodologies for evaluating website accessibility: WCAG-EM, RGAA, and the Polish Web Accessibility Evaluation Methodology.

2.1 WCAG-EM

One of the most significant methodologies for evaluating the digital accessibility of websites is the Website Accessibility Conformance Evaluation Methodology (WCAG-EM) [3]. This methodology was developed by the World Wide Web Consortium (W3C) as part of the Web Accessibility Initiative (WAI) [4]. WCAG-EM represents the official approach to assessing the conformance of websites with the WCAG.

Purpose and Scope. WCAG-EM was designed to provide consistent and structured guidance for evaluating the accessibility of entire websites, rather than individual pages [5]. Its main objective is to define the structure and procedures by which accessibility evaluations should be conducted. Importantly, WCAG-EM does not prescribe specific auditors or automated tools for conducting evaluations.
This methodology is particularly relevant for: organizations required to report the accessibility level of their digital services, accessibility auditors, designers and IT teams preparing conformance documentation, public institutions subject to legal regulations (e.g., the European Accessibility Act or the Web Accessibility Directive).
WCAG-EM allows for flexible adaptation of the scope and depth of evaluations depending on the goal of the assessment. It can be applied in the context of full certification audits, internal quality control, comparative studies, or evaluations of pilot projects.

Methodological Structure. The WCAG-EM methodology consists of five main phases:

1. *Define the Evaluation Goal*. At this stage, the purpose of the evaluation is specified, such as confirming conformance for an audit, identifying barriers from the end-user perspective, or monitoring changes over time.
2. *Define the Evaluation Scope*. This involves delineating the boundaries of the website or application under review, including the set of URLs, sections, functions, or resources to be included. All relevant components of the user interface, navigation elements, forms, and embedded resources must be considered.
3. *Explore the Website and Identify Components*. The evaluator explores the website to understand its structure, navigation logic, functionalities, and the technologies used in its construction. This phase also involves identifying different types of pages (e.g., forms, search interfaces, galleries, login pages, etc.).
4. *Select a Representative Sample of Web Pages*. Due to the complexity of modern websites, it is not feasible to analyze every single page. WCAG-EM recommends selecting a representative sample that reflects the typical functions, content types, and layouts of the website. This sample should include main pages, subpages, forms, various templates, and dynamic content.

5. *Evaluate Conformance and Report the Findings*. The final stage involves both manual and automated assessments of the selected components in relation to WCAG standards. WCAG-EM does not mandate the use of specific tools but encourages a hybrid approach combining automated methods with expert evaluation. The results are then aggregated and presented in a conformance report, which may include: level of conformance (A, AA, or AAA), specific errors and warnings, recommendations, scope of limitations.

Principles of Evaluation. WCAG-EM emphasizes the importance of transparency and reproducibility in the evaluation process. Each phase should be fully documented, from the justification of the chosen scope and page sample to the detailed description of applied methods and tools. For each evaluated unit, success criteria should be clearly identified, along with the reported conformance level and any limitations.
A distinctive feature of WCAG-EM is its flexibility. It can be scaled according to the needs, resources, and organizational context of the auditing team.

Significance in Practice and Policy. WCAG-EM is currently the recommended methodology by W3C and is widely used by accessibility auditors worldwide. In Europe, it has become the foundation for implementing requirements under the Web Accessibility Directive for public sector websites and mobile applications.
Numerous institutions also adopt WCAG-EM as a reference for developing their own digital accessibility policies. The methodology serves as a basis for designing automated evaluation solutions and hybrid approaches that combine automation with user testing.

Limitations. Despite its many strengths, WCAG-EM is not without limitations. A key challenge lies in the subjectivity of interpreting certain WCAG success criteria, which can lead to discrepancies between different evaluators. Additionally, the methodology relies heavily on manual analysis, which can be time-consuming and resource-intensive, especially for large-scale websites.
Another significant limitation is that WCAG-EM does not explicitly incorporate user experience (UX) or testing with users with disabilities, both of which are crucial complements to traditional conformance audits.

Conclusion. WCAG-EM is a comprehensive, flexible, and highly recommended methodology for digital accessibility evaluation, establishing a robust framework for assessing conformance with WCAG 2.x standards. While it does not replace automated tools or usability testing, it provides the foundation for a systematic, well-documented, and comparable evaluation process. Its importance is expected to grow in parallel with increasing legal and societal demands for digital inclusion.

2.2 RGAA

One of the most advanced and institutionally supported approaches to web accessibility evaluation in Europe is the French Référentiel Général d'Amélioration de l'Accessibilité

(RGAA) [6] [7]. This framework serves as the official digital accessibility standard in France and was developed to align national digital infrastructure with the international WCAG. Beyond being a practical auditing tool, RGAA functions as a legally binding document, distinguishing it from many internationally adopted but non-enforceable methodologies.

Legal and Institutional Context. RGAA was developed in response to France's 2005 Law on Equal Access to the Information Society, with its first version published by the Ministry for Digital Transformation and the Civil Service in 2009. Since its inception, RGAA has undergone several revisions to stay aligned with newer iterations of the WCAG. The current version is based on WCAG 2.1, corresponding to the AA conformance level.
The standard is mandatory for all public sector entities and organizations performing public missions in France. Its enforcement is overseen by regulatory bodies such as DINSIC and CNIL, and compliance reports are required to be publicly accessible. Thus, RGAA integrates both technical and regulatory dimensions within a unified framework.

Structure and Characteristics of RGAA. RGAA is characterized by precisely formulated and unambiguous test criteria, enabling replicable and measurable evaluation of user interface components. The framework consists of:

- a thematically organized list of accessibility criteria (e.g., images, tables, forms, multimedia),
- descriptions of test conditions, including specific questions and test cases,
- code examples and result interpretations, which help clarify ambiguous scenarios,
- mappings to WCAG, facilitating comparison with international guidelines.

In total, RGAA defines over 100 detailed test criteria, each mapped to corresponding WCAG success criteria and accompanied by a standardized testing procedure. This structured format significantly enhances its usability in real-world audits.

Evaluation Methodology. The RGAA audit process is divided into three key stages:

1. *Website Exploration and Functional Component Identification*. The audit team analyzes the site's structure, front-end technologies, and the variety of user interface elements.
2. *Conformance Testing Against RGAA Criteria*. Conducted by certified auditors, using a combination of automated tools and manual verification. The well-defined test scenarios are central to ensuring consistency and reliability of results.
3. *Reporting and Documentation*. Each audit concludes with a compliance report that includes error listings, percentage of conformance, and recommendations for remediation. French public institutions are required to publish an accessibility statement based on RGAA findings.

Unlike methodologies such as WCAG-EM, RGAA mandates full coverage of the evaluated application or website rather than relying on a representative sample. While this ensures a higher degree of accuracy, it also entails greater resource investment.

Integration with Digital Policy and Practice. RGAA plays a central role in shaping and implementing digital accessibility policy in France. Major public and private-sector organizations - such as SNCF, Orange, and EDF - regularly conduct audits using this standard. It serves as the foundation for certification, public project evaluation, and IT procurement processes. The strength of RGAA also lies in its robust implementation ecosystem, which includes: official documentation templates, dedicated testing tools (e.g., Tanaguru, Asqatasun), auditor certification programs and community platforms for sharing best practices. These support mechanisms facilitate widespread adoption and continuous professional development.

Strengths and Limitations. RGAA is distinguished by its high level of precision and formalization. Its standardized testing scenarios significantly reduce subjectivity and ensure high audit consistency across different evaluators - offering a clear advantage over more interpretive methodologies such as WCAG-EM.
However, this formalization can also be a barrier for smaller organizations or development teams lacking dedicated accessibility resources. Moreover, RGAA focuses primarily on technical conformance, with limited attention to usability or the real-world experiences of users with disabilities.

Conclusion. RGAA represents an advanced national standard for digital accessibility evaluation, integrating technical, organizational, and legal requirements into a cohesive conformance framework based on WCAG. It is one of the most comprehensive accessibility assessment tools in Europe and offers a valuable model for the development of similar standards in other countries. Its success is rooted in both the clarity and measurability of its criteria and its institutional embedding within the French legislative system. As legal obligations and expectations regarding digital inclusion continue to grow across Europe, RGAA serves as a benchmark example of how accessibility policy and practice can be effectively aligned.

2.3 The Polish Web Accessibility Evaluation Methodology

Among national initiatives aimed at the systematic assessment of digital accessibility, a notable contribution is the Polish web accessibility evaluation methodology, developed by the Widzialni Foundation [8] in collaboration with the University of Silesia. This is one of the most mature and comprehensive Polish frameworks, integrating normative (WCAG 2.1-based), practical (audit-focused), and educational components. It serves as the foundation for evaluating the accessibility of websites across the public and private sectors, as well as supporting academic research and comparative studies.

Development Context The methodology was developed in response to growing legal requirements following the implementation of Directive (EU) 2016/2102 and the Polish Act on the Digital Accessibility of Public Sector Websites and Mobile Applications (April 4, 2019). The Widzialni Foundation, a national leader in digital accessibility, has long been involved in expert, educational, and implementation efforts. The methodology was created in partnership with researchers from the University of Silesia in Katowice.

Purpose and Scope. The methodology aims to adapt international WCAG standards to the realities of Polish public administration websites -especially those of municipalities, schools, cultural institutions, and other local entities. It takes into account the specifics of the Polish language, cultural context, the technological landscape of the public sector, and the organizational limitations of local governments.

Methodological Foundations and Evaluation Scope. The framework is based on WCAG 2.1 Level AA criteria, in line with legal requirements, but it goes further by providing detailed testing scenarios and contextual interpretations tailored to Polish websites. The audit structure employs a mixed-methods approach, combining:

- automated analysis using tools such as Axe, WAVE, TAW, and SortSite,
- manual expert audits conducted by certified auditors from the Widzialni Foundation,
- functionality and interface usability testing,
- in some cases, user testing involving individuals with disabilities.

The audit encompasses technical elements (HTML structure, ARIA attributes, contrast ratios, semantic markup) and functional components (forms, navigation menus, multimedia, keyboard focus, screen reader compatibility, etc.).

Audit Procedure. The audit process, as defined by the Widzialni Foundation and the University of Silesia, follows a clearly structured sequence:

1. *Website Exploration and Functional Component Identification*. This involves analyzing front-end technologies, CMS platforms, navigation structures, and content organization.
2. *Sample Selection*. Representative subpages are selected for testing—typically 10 to 15 key pages featuring forms, multimedia, tables, document structures, and interactive elements.
3. *Automated and Manual Testing*. Each test is documented using a standardized audit form including error descriptions, location, user impact, and repair recommendations.
4. *Reporting*. A detailed accessibility report is generated, presenting: a percentage-based conformance score, classification of errors (critical, major, moderate, minor), recommendations for both developers and content editors and a summary suitable for publication as a formal accessibility statement.

Unlike WCAG-EM, this methodology often includes user testing, offering a more comprehensive, user-centered evaluation.

Practical Implementation. The methodology has been widely adopted in practice, most notably in the Digital Accessibility Ranking of Local Government Units, which evaluates hundreds of websites from municipalities, counties, and regional authorities in Poland. This enables nationwide monitoring of digital accessibility, identification of common barriers, and the formulation of systemic recommendations.
Moreover, the methodology is used as a foundation for auditor training and certification, and as educational material in fields such as UX, interface design, social informatics, and e-government.

Contribution to National Accessibility Policy. The Polish methodology is interdisciplinary, pragmatic, and educational in nature. It combines international best practices with local technological and institutional realities, serving as a valuable complement to global approaches such as WCAG-EM or ACT Rules. Its design is especially suited to the needs of local public sector entities, many of which lack dedicated accessibility teams.
As such, the methodology has played a key role in building awareness and competencies in digital accessibility across Poland.

Strengths and Limitations. Key strengths of the methodology include: adaptation to the Polish legal and technological context, practicality and measurability, inclusion of user testing, setting it apart from many other approaches, promotion of transparency and reproducibility.
Limitations include: focus on WCAG Level AA, without native support for Level AAA, limited testing for documents and mobile applications, the cost and complexity of expert and user-involved audits may pose challenges for smaller institutions

Conclusion. The Polish web accessibility methodology represents a mature and pragmatic approach to evaluating digital accessibility in Poland. It combines the rigor of WCAG 2.1 AA with user testing, training, and nationwide monitoring. Compared to WCAG-EM and RGAA, it is distinguished by its strong alignment with the operational realities of Polish public institutions and its emphasis on the participation of people with disabilities in the audit process.

3 Comparative of Digital Accessibility Evaluation Methodologies

Table 1 summarizes the key characteristics of the methodologies discussed in this study. Among them, WCAG-EM provides the most universal and flexible methodological framework, although it does not include specific test procedures. It requires expert knowledge and must be integrated with appropriate tools to be effectively applied. RGAA and the Polish accessibility methodology represent national implementations of WCAG. The RGAA offers a comprehensive set of test procedures and is deeply embedded in legal regulations, making it a robust reference standard in the French context. The Polish methodology, in contrast, is a simplified, practice-oriented model, adapted to the specific needs and constraints of the local context. It proves particularly effective in accessibility audits and training programs, offering a pragmatic approach suitable for institutions seeking to improve accessibility with limited resources.

Table 1: Comparative of Digital Accessibility Evaluation Methodologies

Criterion ↓	**WCAG-EM (W3C)**	**RGAA (France)**	**Polish Methodology**
Institution	W3C/WAI	French government (DINUM, formerly DINSIC)	Fundacja Widzialni and University of Silesia
Year of Development (Version)	2014 (WCAG-EM 1.0)	2009 (currently RGAA 4.1, updated in 2023)	2013 (with updates in 2020 and 2023)
Normative Basis	WCAG 2.0 / 2.1 / 2.2	WCAG 2.1 (partially 2.2)	WCAG 2.1, levels A and AA
Structure	Methodological framework without specific test cases	106 technical tests + glossary + appendices	21 accessibility tests with definitions and scoring sheets
Test Coverage	Methodological structure supporting any WCAG-based test	13 main accessibility areas covered via specific tests	Focus on 21 essential accessibility criteria relevant for Polish websites
Test Type	Manual + supported by automated tools	Manual, semi-automated, expert-based	Mostly manual expert audits, supported by automated tools
Sampling Procedure	Defined procedure: selection of representative pages and user paths	Not specified – tests are applied to individual pages or applications	No strict procedure, test scope depends on website type
Reporting Format	Report structure recommended, flexible output format	Predefined forms (Excel/HTML) required for legal compliance	Audit sheets (PDF or Excel) + summary table + technical explanation
Practical Use	Adopted across many EU countries, used for audits and conformity assessments	Mandatory for French public sector institutions (by law)	Widely used by public institutions, NGOs, universities in Poland
Legal Status	Non-binding but influential reference model	Legally binding for French public digital services	Non-binding, recommended in context of the Polish Digital Accessibility Act (2019)
Documentation Availability	Official W3C site (EN)	Official RGAA site (FR), selected translations available	Free documentation (PL) by Fundacja Widzialni
User Involvement	Optional user testing, encouraged but not required	Purely technical focus; no user involvement	User testing not mandatory, but possible for advanced audits
Update Frequency	No regular update schedule; latest in 2014	Frequent updates (latest in 2023)	Irregular updates, latest major update in 2023
Complexity Level	High – designed for accessibility professionals	High – requires technical expertise in HTML, ARIA, WCAG	Medium – suitable for trained staff and institutions
Target Users	Accessibility experts, auditors, consultants	Public sector auditors, developers, accessibility professionals	Public institutions, NGOs, educators, certified trainers
Tool Integration	Compatible with tools like Axe, WAVE, TAW	Integrated with specific tools like Tanaguru, Asqatasun	Works with support tools (e.g., WAVE, ANDI, Axe)

4 Summary

The purpose of this article was to analyze and compare three selected methodologies for evaluating the digital accessibility of websites: WCAG-EM (developed by W3C), RGAA (the French governmental methodology), and the Polish methodology developed by Fundacja Widzialni in collaboration with the University of Silesia. Each of these methodologies constitutes a distinct response to the challenges of implementing digital accessibility standards, all drawing upon the Web Content Accessibility Guidelines (WCAG), yet operationalizing them through differing technical, organizational, and legal frameworks.

WCAG-EM provides a flexible methodological framework that supports systematic evaluation through the selection of representative samples and the combination of manual and semi-automated techniques. While internationally recognized and highly adaptable, it lacks a predefined set of tests and requires a high level of expertise, which may limit its accessibility to less experienced organizations or evaluators.

In contrast, RGAA offers a nationally standardized model mandated by law for the French public sector. Its strengths lie in the comprehensive and detailed list of technical tests, clarity of interpretation, and official backing. The methodology is further reinforced by supporting tools such as Tanaguru and Asqatasun, which facilitate its integration into institutional workflows and audits.

The Polish methodology, developed by Fundacja Widzialni and the University of Silesia, is a practical tool tailored to the local administrative and educational context. It stands out for its relative simplicity, accessibility to non-specialists, and the provision of ready-to-use audit forms and scoring sheets. Although it does not hold legal authority, the methodology has been widely adopted in practice across public institutions, NGOs, and higher education institutions in Poland.

The comparative analysis reveals not only technical differences but also contrasts in functional design and institutional integration. While WCAG-EM serves as a universal, framework-oriented approach, RGAA and the Polish model are more localized and operationally focused. Despite these differences, all three methodologies share a common objective: enhancing the accessibility of digital content for people with disabilities through structured evaluation processes.

The analysis indicates that each of the methodologies presented in the article has been developed to suit its specific context, evaluation purpose, the competence level of auditors, and the expectations of end users. This suggests that future research should consider emphasizing the development of hybrid approaches - combining standardized testing procedures with user-centered perspectives - while also adapting tools to the diverse needs of both public and private sector institutions.

References

1. Journal of Law 2019 poz. 1696 (2019). Ustawa z dnia 19 lipca 2019 r. o zapewnianiu dostępności osobom ze szczególnymi potrzebami (in English: Act of 19 July 2019 on ensuring accessibility for people with special needs), Sejm RP, Warszawa 2019.
2. Web Content Accessibility Guidelines (WCAG) 2.2. World Wide Web Consortium. W3C. https://www.w3.org/TR/WCAG22/. Last accessed 1 July 2025.

3. WCAG-EM Overview: Website Accessibility Conformance Evaluation Methodology. W3C Web Accessibility Initiative (WAI), https://www.w3.org/WAI/test-evaluate/conformance/wcag-em/. Last accessed 1 July 2025.
4. Website Accessibility Conformance Evaluation Methodology (WCAG-EM) 1.0. W3C Working Group Note 10 July 2014 https://www.w3.org/TR/WCAG-EM/. Last accessed 1 July 2025.
5. P. Acosta-Vargas, S. Luján-Mora and L. Salvador-Ullauri. Evaluation of the web accessibility of higher-education websites. 15th International Conference on Information Technology Based Higher Education and Training (ITHET), pp.1-6, Istanbul, Turkey (2016). https://doi.org/10.1109/ITHET.2016.7760703.
6. Référentiel général d'amélioration de l'accessibilité - RGAA. (in English: General Accessibility Improvement Framework - RGAA) https://accessibilite.numerique.gouv.fr/ Last accessed 1 July 2025.
7. Moiz Yamani. Understanding RGAA: France's Accessibility Law and What It Means for You. February 3, 2025. https://www.barrierbreak.com/understanding-rgaa-frances-accessibility-law-and-what-it-means-for-you/ Last accessed 1 July 2025.
8. Przemysław Marcinkowski, Marcin Luboń, Izabela Mrochen, Monika Szczygielska, Artur Marcinkowski. Metoda oceny dostępności cyfrowej strony internetowej zgodnie z zasadami WCAG 2.1 (in English: A method for assessing the digital accessibility of a website in accordance with WCAG 2.1 principles). Fundacja Widzialni. Częstochowa (2020)

Przemyslaw Simiński

ORCID: 0000-0002-2323-3152

University of Siedlce
Faculty of Exact and Natural Sciences
Institute of Computer Science
ul. 3 Maja 54, 08-110 Siedlce, Poland

przemyslaw.siminski@uws.edu.pl

Unmanned Ground Vehicles (UGVs) on the Modern Battlefield: Threats in Light of the Russia–Ukraine War Experience

Abstract. This paper examines the operational employment of Unmanned Ground Vehicles (UGVs) in contemporary urban warfare, drawing primarily on experiences from the Russia–Ukraine conflict. As modern battlefields increasingly shift towards densely built-up environments, UGVs emerge as a critical capability for mitigating risks to personnel while enabling reconnaissance, resupply, and casualty evacuation in contested urban terrain. The analysis focuses on six principal threat categories affecting UGV operations: kinetic threats from conventional and asymmetric weapons systems, electronic warfare countermeasures including GPS spoofing and communication jamming, cyber intrusions targeting control systems and firmware, information warfare campaigns aimed at undermining operator confidence, logistical constraints particularly battery performance in extreme conditions, and normative-ethical challenges surrounding autonomous lethal decision-making. Based on documented battlefield incidents from Bakhmut and Avdiivka, the study demonstrates that effective UGV deployment requires robust mesh networking architectures, enhanced autonomous navigation capabilities, and comprehensive electromagnetic compatibility solutions. The findings underscore that without resilient communications and cyber-hardening measures, unmanned platforms become vulnerable assets rather than force multipliers on the 21st-century battlefield.

Keywords: Unmanned Ground Vehicles, Urban Warfare, Electronic Warfare, Russia-Ukraine War, Autonomous Systems, Cyber Threats, Military Robotics

1 Introduction

In recent decades, the battlefield has undergone a process of steady urbanisation. Cities, once merely areas at risk, have become the central arenas of military operations. An analysis of combat operations in Ukraine demonstrates that regardless of the phase of the conflict Russian and Ukrainian forces consistently gravitated towards urban areas where the most intense clashes occurred: Kyiv, Bakhmut, and Mariupol. The underlying rationale is straightforward: in a context of limited human and material resources, urban environments serve simultaneously as strategic objectives and natural fortresses. As noted in the literature [1,3], today's numerically smaller armed forces are often unable to encircle or isolate urban agglomerations completely. Instead, they become entangled in a series of micro-sieges, targeting individual buildings, streets, and districts. Dense urban structures impede reconnaissance, restrict the effective range of conventional weapon systems, and amplify the effect of "dead ground". This, in turn, incentivises solutions that mitigate risks to personnel and enable observation and precision strikes within the urban labyrinth. Given the permanence of urbanised warfare and its demands for mobility, situational awareness, and force protection, Unmanned Ground Vehicles (UGVs) emerge as a logical response to three critical challenges. First, UGVs can assume the most hazardous tasks [2], including short-range reconnaissance, supply delivery, casualty evacuation, and the neutralisation of improvised explosive devices (IEDs). Second, due to their compact dimensions and agility, ground robots can operate in environments inaccessible to tanks or armoured vehicles such as stairwells, narrow alleyways, and underground passages. Third, network-centric integration of UGVs with aerial drones and stationary sensors enables a dominant and multi-dimensional understanding of the battlespace within densely built-up areas.

2 Operational Environment

The battle for Bakhmut illustrated the evolving nature of the Russia–Ukraine war, characterised more by a succession of miniaturised sieges than by large-scale frontal assaults. Instead of rapid breakthroughs, we observe the methodical attrition of city blocks, intersections, and individual buildings. For defenders, the city becomes a fortress; for attackers, a maze of ruins in which manoeuvring heavy vehicles is virtually impossible [9].

The combination of dense construction and widespread rubble creates an intricate, multi-tiered terrain of basements, upper storeys, and corridors, where small and manoeuvrable UGVs can penetrate areas inaccessible to tanks or Infantry Fighting Vehicles (IFVs). Sustaining combat contact in such an environment requires continuous ammunition resupply and casualty evacuation operations where UGVs significantly reduce risk to human crews.

In February 2024, Russian forces employed a strategy of mass deployment of KAB guided aerial bombs (30–60 per day) to breach Ukrainian defences in Avdiivka. Only after weeks of relentless aerial and artillery bombardment were they able to enter the ruins of the city. The consequences were devastating: the total destruction of critical infrastructure, the formation of vast belts of burnt-out debris impeding wheeled movement, and the high saturation of unexploded ordnance and improvised mines [4,9].

Such heavily contaminated battlefields favour UGVs equipped with combat engineering modules. The destruction of road infrastructure necessitates high-terrain-capability platforms such as those using hybrid wheel-track systems or active suspension. The adversary's advantage in aerial strike capabilities further underscores the need for UGVs with reduced thermal and acoustic signatures.

Both Russian and Ukrainian forces have deployed extensive arrays of electronic warfare systems (e.g., Krasukha-4, Tirada-2, Nota-S, Bukovel-AD), capable of disrupting GNSS, GSM, and UHF frequencies at ranges of several kilometres. Conventional radio equipment is routinely jammed, compelling operators to utilise directional antennas, frequency-hopping techniques, or wired communications.

This dense electromagnetic interference necessitates several technical adaptations for UGVs: Higher levels of autonomy (e.g., inertial and optical navigation), Redundant communication links (e.g., mesh networks, optical laser communication), Electromagnetic concealment modules, to resist detection by adversarial SIGINT systems. Ukrainian research highlights that, in most cases, stable radio communication without specialist shielding measures rarely exceeds 4–6 km from command nodes.

3 Threats to Unmanned Ground Vehicles (UGVs)

3.1 Kinetic Threats

Standard small arms calibres (5.45–7.62 mm) easily perforate the light armour of most UGVs. Inexpensive anti-tank guided missiles (ATGMs) with shaped charges such as the Lancet or Scalpel are capable of destroying entire vehicles. Smart mines, equipped with seismic or magnetic sensors, can be triggered by the minimal mass of small robots. In 2024 alone, Russian and Ukrainian FPV drone units conducted over 3,000 attacks per month, many of which targeted light evacuation robots. During demining operations in Zaporizhzhia, UGVs triggered anti-side mines responsive to low-level magnetic signatures [10].

3.2 Electronic Threats

Russian jamming systems such as Krasukha-4, Tirada-2, and Bukovel-AD can suppress GPS and UHF frequencies at distances exceeding 10 km. Spoofing techniques introduce false coordinates or intercept video streams. Electromagnetic pulse (EMP) munitions (e.g., 3M-EMP rockets) can temporarily reset onboard controllers. Russian forces have also used AeroScope systems to identify drones and direct artillery strikes at their operators. Following the jamming of Starlink signals over Crimea, Ukrainian UGVs lost connectivity with their C2 centres.

Inertial navigation systems, supplemented with LiDAR, are therefore essential. Ukrainian findings show that, without advanced electronic protection, reliable radio connectivity rarely extends beyond 4–6 km from the command node. This limitation has significant implications for the design and deployment of unmanned systems on the contemporary battlefield [11].

3.3 Cyber Threats

Hostile cyber operations represent an escalating threat to unmanned ground systems. Techniques such as malicious packet injection within mesh networks can result in the redirection of a robot's trajectory or its complete immobilisation. Particularly concerning are incidents of reverse engineering of captured UGVs, which enable the installation of compromised firmware embedded with hidden backdoors. In 2024, Russian laboratories in Taganrog demonstrated Ukrainian Uran-IX platforms that had been captured, reprogrammed, and reconfigured to operate exclusively under the control of Russian command systems. This case, documented in a report by CNA [12], illustrates the scale of the cyber threat and underscores the urgent need for the implementation of robust cryptographic security protocols to ensure operational integrity.

3.4 Informational Threats

The information environment surrounding unmanned systems has become increasingly sophisticated and contested. Disinformation campaigns often involve the deliberate dissemination of fabricated video content, falsely portraying UGV malfunctions with the objective of undermining operator confidence and public trust in the technology. A notable example includes the Russian Telegram channel "Rybar", which circulated doctored footage allegedly showing Ukrainian robots firing upon civilians. These efforts aim not only to demoralise adversary forces but also to influence public perception and policymaking in allied states [13].

3.5 Logistical Threats

Logistical challenges significantly impact UGV performance and mission reliability. Lithium-ion batteries experience a loss of over 30% capacity at temperatures below –10°C, substantially reducing operational endurance. During the winter of 2023–2024, small UGVs and FPV drones deployed by the Ukrainian "Rubizh" Brigade exhibited an average autonomy of only 12 minutes, despite official specifications claiming up to 25 minutes. Loss of C2 links may result in UGVs becoming inoperative in the field or falling into enemy hands. To address these constraints, hybrid propulsion systems such as those using JP-8 fuel range extenders, inductive charging stations located within fortified positions, and "return-to-base" algorithms with automated self-destruction after 15 minutes of signal loss are increasingly required.

3.6 Normative and Ethical Challenges

Perhaps the most complex and controversial issue concerns the use of lethal force by UGVs. On one hand, such decisions must be made rapidly and proportionately to the threat; on the other, there is a clear risk of unauthorised or unjustified engagement. The deployment of autonomous ground platforms equipped with AI-based targeting systems capable of using force without human input raises profound questions regarding their compliance with international humanitarian law (IHL). In the chaotic and uncertain conditions of real-world combat zones, the reliability of decision-making algorithms diminishes significantly, increasing the risk of tragic errors in target identification.

3.7 Key Observations from the Ukrainian Theatre

Teleoperation, direct control of the platform by a human operator via joystick remains the dominant control model for light evacuation platforms and EOD robots. The operator, typically situated in a command vehicle or fortified shelter, receives a high-resolution video feed from the platform and issues real-time control commands. This model offers several undeniable advantages. Chief among them is the immediacy of human situational awareness, which ensures high precision during complex manipulative tasks or weapons deployment. Furthermore, the compatibility of teleoperation with existing operational procedures minimises the need for intensive personnel retraining. However, Ukrainian battlefield experience has laid bare the critical vulnerabilities of this approach. Russian electronic warfare systems such as Krasukha-4and Ukrainian equivalents like Bukovel-AD can effectively disrupt UHF and L-band transmissions over distances of up to 10 kilometres. Loss of telemetry in this model results in the immediate loss of control over the platform. Additionally, disruptions to the Starlink satellite network over Crimea revealed the dangers of over-reliance on a single communications provider, which may become a single point of failure under concerted jamming efforts

4 Conclusions

Operational insights from Ukraine affirm a stark truth: A robot without communications is merely a high-tech wreck, and a robot lacking cyber resilience is a weapon in enemy hands. Accordingly, it is imperative for the Polish Armed Forces to invest in: Native mesh networking architectures, Autonomy-enhancing technologies, Robust counter-jamming and electromagnetic compatibility (EMC) solutions. Only by doing so can future UGV battalions truly enhance soldier survivability and reduce fire response times in the highly dynamic conditions of the 21st-century battlefield.

References

1. Kagan, F.W., Barros, G., Stepanenko, K.: Russia's War in Ukraine: Operational Assessment 2022–2024. Institute for the Study of War, Washington DC (2024)
2. King, A.: Urban Warfare in the Twenty-First Century. Cambridge University Press, Cambridge (2023)
3. Locher, M.: The Urban Siege Paradigm: Small-Unit Combat in Dense Terrain. RAND Corporation, Santa Monica (2023)
4. Silvus Technologies: StreamCaster MN-MIMO Radio Family Datasheet. https://silvustechnologies.com/products/streamcaster. Last accessed 15 May 2025
5. NATO Standardization Office: STANAG 4586 (Ed. 5) – Standard Interfaces of UAV Control System (UCS) for NATO UAV Interoperability. Brussels (2022)
6. Mari, G., Virovac, D., Gazzoni, M.: Real-Time LIDAR-SLAM in Dust-Obscured Environments. Robotics 14(2), 110–129 (2025). https://doi.org/10.3390/robotics14020110
7. Khramov, D., Ukolov, A.: EMP Vulnerabilities of Unmanned Ground Platforms. In: Smirnov, P., Manko, S. (eds.) MIL-EM 2024 – 14th Int. Conf. on Military Electromagnetics, LNCS, vol. 14567, pp. 55–70. Springer, Cham (2025). https://doi.org/10.1007/978-3-031-56789-4_6
8. Хурин И.О., Матвеев Л.І. (2023). Мультиагентна модель пункту управління групою безпілотних літальних та наземних апаратів при виконанні спеціальних місій Тестинг анд Цертифицатион, 1 (1), 39-46.
9. Vlasov A.V., Zhukov I.A., Matiushchenko O.G., Rieznikov Yu.V. (2024). Method for preliminary assessment of the techno-economic level of mechanical systems for humanitarian demining. Testing and Certification, 2 (4), 14-21.

10. Zasjadko A.A., Kuznetsov V.O., Lytovchenko V.V., Husak M.Yu., Tsuria E.I. (2024). Failure mode and effects analysis of mechanized humanitarian demining systems at the testing stage. Testing and Certification, 2 (4), 41-50.
11. Андрушко М.В., Аркушенко П.Л., Кузмич О.Є., Гузии Є.О., Флорин О.П. (2024). Обґрунтування типовоu методики проведення випробувань виробів броньованоu техніки з оцінювання рівня захисту від minноu загрози із використанням сучасних інформаційно-вимірювальних систем. Тестинг анд Цертифицатион, 2 (4), 7-13.
12. Орлов С.В., Звиглянич С.М., Червотока О.В., Херащенко М.О. (2024). Аналіз методів лабораторних вібраційних випробувань літальних апаратів безпілотних авіаційних комплексів. Тестинг анд Цертифицатион, 4 (6), 7-12. 36-99-ПБ.
13. Артюх О.М., Слюсаров О.С., Рубан Д.П., Руснак В.М., Калиниченко О.С. (2025). Особливості проєктування амфібійних машин: ключові виклики та інженерні рішення. Тестинг анд Цертифицатион, 1 (7), 36-48.
14. Simiński P., Wyzwania dla systemów informatyczno-elektronicznych pojazdów wojskowych), Wydawnictwo UwS. Siedlce 2024.

Ewa Szczepanik[1]
ORCID: 0000-0002-1320-3275

Szymon Szczepanik[2]

[1] University of Siedlce
Faculty of Exact and Natural Sciences
Institute of Computer Science
ul. 3 Maja 54, 08-110 Siedlce, Poland

[2] Warsaw University of Technology
Faculty of Electronics and Information Technology
ul. Nowowiejska 15/19, 00-665 Warszawa, Poland

ewa.szczepanik@uws.edu.pl, szymon.szczepanik.stud@pw.edu.pl

The Use of Artificial Intelligence in Teaching Robotics at School Using LEGO Mindstorms

Abstract. This article explores how artificial intelligence (AI) can be integrated into school-level robotics education through LEGO Mindstorms EV3 kits. It highlights the educational benefits of combining AI with robotics to enhance students' digital competencies, foster problem solving skills, and introduce real-world applications of emerging technologies in a hands-on and engaging way. The paper also discusses challenges associated with implementation, teacher training, and data privacy.

Keywords: Artificial Intelligence, Robotics, LEGO Mindstorms, Educational Technology, Programming, Machine Learning, STEM Education, Project-Based Learning

1 Introduction

In recent years, especially following the pandemic, we have witnessed rapid technological advancement that has significantly impacted everyday life and education. This is especially evident in technical education, including robotics instruction. Traditional teaching methods are increasingly being replaced by innovative approaches that emphasize interactivity, experiential learning, and the use of modern digital tools.

One such tool is the LEGO Mindstorms EV3 kit, which has long served as an engaging educational platform for teaching the fundamentals of robotics, particularly algorithmic thinking and programming, to children and adolescents.

The accelerated pace of technological progress has also fueled the growing interest in artificial intelligence (AI). In recent years, AI has gained increasing attention in various educational contexts. The integration of robotics and AI offers entirely new opportunities for both students and educators. This article aims to demonstrate how AI can be incorporated into teaching practices, particularly in robotics education using LEGO Mindstorms kits, and to examine the potential benefits and challenges of this approach.

2 LEGO Mindstorms as an Educational Tool

It is widely recognized that building with blocks fosters creativity among children and adolescents. LEGO Mindstorms kits, especially the EV3 version and the newer LEGO Mindstorms Robot Inventor, encourage creative thinking while also supporting the development of logical and algorithmic thinking. They introduce learners to programming fundamentals and help them develop problem-solving skills.

These kits include not only bricks, but also sensors, motors, and programmable control units ('bricks') that allow users to build and program a wide range of robots. Thanks to their intuitive interface and compatibility with programming languages such as Scratch and Python, LEGO Mindstorms kits are used across various educational levels, from primary schools to universities.

By incorporating LEGO Mindstorms into robotics instruction, educators can boost student engagement, spark interest in programming through immediate, visible results ("I can see how my code works"), and improve students' logical reasoning, communication skills (particularly through teamwork), and other soft skills.

3 Artificial Intelligence in Education and Robotics

Artificial Intelligence (AI) encompasses technologies that enable machines to exhibit abilities similar to those of humans, such as reasoning, learning, decision-making, inference, and creativity. These systems can analyze their surroundings and make informed decisions—essentially mimicking aspects of human cognition.

In education, AI can support personalized learning, track and analyze student progress, respond to individual learning needs, and assist teachers in developing customized teaching materials. From this perspective, AI not only facilitates learning but also presents new challenges for educators and education systems.

AI can support teachers in the following ways:

Data analysis – such as test scores and student grades, and help identify areas that require more attention.

Individualized learning – individualized instruction that takes into account the learning style of each student and academic progress.

Personalization of educational materials – making learning more engaging and effective, for example, virtual tutoring, adapting test questions, and creating student-specific learning paths.

Automation of administrative tasks – such as grading assignments and record-keeping.

Supporting online learning – making learning more interactive and engaging, and providing students with more opportunities to interact with teachers and other students.

Providing support for students with disabilities – such as dyslexia or autism spectrum disorder.

Automatic language translation – allowing students to easily access educational materials from around the world.

However, in the field of robotics, artificial intelligence is used to increase machine autonomy. Applications include image processing (e.g., obstacle detection, object recognition), real-time decision-making (e.g., path planning, collision avoidance), machine learning (e.g., learning behaviors from data), and voice interaction (e.g., speech recognition and processing).

Integrating AI into **educational robotics** allows students to explore advanced computer science concepts in a practical context and to understand how AI operates in real-world scenarios. In addition, it supports student cognitive development and improves teacher instructional abilities.

Currently, we have a wide selection of educational robots. Examples include LEGO Education (WeDo, BOOST, SPIKE, MINDSTORMS), Photon, Ozobot, Makeblock mTiny and Codey Rocky, Bee-Bot/Blue-Bot, and robotic spheres like the Sphero SPRK+. Choosing a specific robot depends on many factors, including the child's age and the tasks it will perform.

However, it is essential to strike a healthy balance between technological innovation and traditional teaching methods, ensuring that educational robots serve as support tools, not replacements for teachers.

4 Possibilities for Integrating AI with LEGO Mindstorms

The integration of AI elements into LEGO Mindstorms projects can take various forms, depending on students' skill levels and the available resources. Below are several possibilities:

4.1 Machine Learning and Data Classification

Using Python libraries (e.g., scikit-learn) and external platforms such as Google Teachable Machine, students can build models that classify sensor input (e.g., sound, light, motion). These models can be implemented in LEGO robots to support real-time decision-making (e.g., "If the sound is loud and the light is red, then stop.").

4.2 Image and Video Recognition

With USB cameras or smartphones, and libraries such as OpenCV, students can equip robots with visual capabilities, enabling them to detect colors, lines, shapes, and even recognize objects. In simpler cases, a robot might follow a line; in more advanced applications, it could recognize road signs or human faces. For instance, students might build a robot that learns to identify colors under varying lighting conditions. Using a color sensor and basic classification algorithms, the robot could navigate a board, stopping at red and accelerating at green. AI enables the robot to “learn” how colors appear in different environments.

4.3 Voice Interaction

Through integration with speech recognition systems such as Google Speech API or Microsoft Azure, students can program robots to respond to voice commands. This provides an effective introduction to natural language processing (NLP). A LEGO robot equipped with a speaker and microphone (e.g., via Raspberry Pi) could be programmed to follow voice instructions such as “turn left,” “say hello,” or “sing a song.”

4.4 Artificial Neural Networks

At more advanced levels, students can experiment with basic artificial neural networks for pattern recognition. For example, a robot could learn to predict the presence of obstacles based on previous measurements of distance and speed. In such projects, students program a robot to analyze data from an ultrasonic sensor and camera to autonomously avoid obstacles. Using basic reinforcement learning algorithms, the robot can learn from mistakes and iteratively refine its behavior.

5 Conclusion

Integrating artificial intelligence into robotics education using LEGO Mindstorms presents an exciting opportunity to bridge theoretical knowledge with hands-on application. It strengthens students’ digital competencies and familiarizes them with the technologies shaping the future.

From a teacher’s perspective, educational robots can serve as effective tools—not only helping students grasp curriculum content but also enabling personalized instruction, boosting motivation, and fostering collaboration and communication skills.

However, despite these benefits, incorporating AI into educational robotics also poses several challenges. One major hurdle is the need for ongoing teacher training to improve digital competencies. Additionally, there is a need for practical, ready-to-use lesson plans that integrate AI concepts in a way that teachers can confidently apply.

The introduction of AI in education also raises concerns about student data privacy, requiring schools to uphold the highest standards of protection. Another limitation lies in the hardware capabilities of LEGO Mindstorms kits, especially the need for integration with external platforms (e.g., Raspberry Pi), as well as the relatively high cost of hardware and software licenses.

References

1. Russell, S., Norvig, P.:Artificial Intelligence: A Modern Approach (4th ed.), Pearson, (2020)
2. AlQarzaie, K.N., AlEnezi, S. A.: Using LEGO MINDSTORMS in Primary Schools: Perspective of Educational Sector International Journal of Online and Biomedical Engineering (iJOE) **18**(01),139–147(2022) https://doi.org/10.3991/ijoe.v18i01.27579

3. Fombona, J., Sáez, J., Sánchez S.: Artificial intelligence and robotics in education: Advances, challenges, and future perspectives. Social Sciences and Humanities Open, 11, (2025) `https://doi.org/org/10.1016/j.ssaho.2025.101533`
4. Parsons, S. , Sklar, E.:Teaching AI using LEGO Mindstorms, Accessible Hands-on AI and Robotics Education, Papers from the 2004 AAAI Spring Symposium
5. Valkanova, V., Jähn, M., Andre, E. Learning by Teaching with LEGO Mindstorms in Artificial Intelligence Education. Proceedings of the IEEE EDUCON Conference. (2021)

ISIT Intelligent Systems and Information Technologies

University of Siedlce

Wydawnictwo Naukowe UwS

ISBN: 978-83-68754-07-0