WIIS2025

Proceedings of Workshop on Intelligent Information Systems

Vladimir Andrunachievici Institute of
Mathematics and Computer Science,
Moldova State University
October 16-18, 2025, Chisinau

CZU 004+004.8(082)
P 93

VLADIMIR ANDRUNACHIEVICI INSTITUTE OF MATHEMATICS AND COMPUTER SCIENCE, MOLDOVA STATE UNIVERSITY
5, Academiei str., Chisinau, Republic of Moldova, MD 2028
Tel: (373 22) 72-59-82, Fax: (373 22) 73-80-27
E-mail: imam@math.md
WEB address: http://www.math.md

Editors: Dr. I. Titchiev, Prof. S. Cojocaru, Prof. C. Gaindric

Authors are fully responsible for the content of their papers.

Descrierea CIP a Camerei Naţionale a Cărţii din Republica Moldova

Proceedings of Workshop on Intelligent Information Systems : WIIS 2025, October 16-18, 2025 Chisinau / editors: I. Titchiev [et al.]. – Chişinău : [S. n.], 2025 (Valinex). – 378 p. : fig. color, tab.

Antetit.: Vladimir Andrunachievici Institute of Mathematics and Computer Science, Moldova State University. – Referinţe bibliogr. la sfârşitul art. – 100 ex.

ISBN 978-9975-68-536-8.

004+004.8(082)
P 93

This issue is supported by the research project 011301, "Information Systems based on Artificial Intelligence".

COMMITTEES

Program Committee

Inga Ţiţchiev (chair),	Serghei Ivanov (France),
Svetlana Cojocaru (co-chair),	Serghei Verlan (France),
Constantin Gaindric (co-chair),	Mircea Petic (RM),
Nataliya Pankratova (Ukraine),	Constantin Ciubotaru (RM),
Stefan Kramer (Germany),	Artiom Alhazov (RM),
Stefan Pickl (Germany),	Alexandru Colesnicov (RM),
Horia Teodorescu (Romania),	Liudmila Burţeva (RM),
Florin Filip (Romania),	Olesea Caftanatov (RM),
Vasile Apopei (Romania),	Iulian Secrieru (RM),
Adrian Iftene (Romania),	Elena Guţuleac (RM),
Lenuţa Alboaie (Romania),	Ina Ciobanu (RM),
Andrei Păun (Romania),	Corina Negara (RM),
Marinela Ianculescu (Romania)	Tudor Bumbu (RM),
Diana Inkpen (Canada),	Andrei Rusu (RM)
Panayotis Yannakopoulos (Greece),	

Organizing Committee

Inga Ţiţchiev,	Liudmila Burteva,
Svetlana Cojocaru,	Tatiana Verlan,
Constantin Gaindric,	Elena Gutuleac,
Mircea Petic,	Irina Gheorghita,
Tudor Bumbu,	Alexandru Parahonico,
Andrei Rusu,	Nichita Degteariov,
Olesea Caftanatov,	Vlada Colesnicova,
Artiom Alhazov,	Daniela Caganovschi,
Iulian Secrieru,	Argentina Fron,
Olga Popcova,	Matei Octavian Turcan
Ludmila Malahov,	

Proceedings of Workshop on Intelligent Information Systems
WIIS2025, October 16-18, 2025, Chisinau, Republic of Moldova

Towards a computational map of similarities and changes in the Romanian language

Invited paper

Liviu P. Dinu

Abstract

Natural languages are living ecosystems; they are constantly in contact and, by consequence, they change continuously. Traditionally, the main Historical Linguistics problems (How are languages related? How do languages change across space and time?) have been investigated with comparative linguistics instruments. We propose here (with a focus on Romanian, in the more general frame of Romance languages), computer-assisted methods for the main problems in HL (related words discrimination, protoword reconstruction, languages similarity, semantic divergence, etc.). Our studies on Romance languages rely on a digital resource for HL that we constructed and published (RoBoCoP – ROmance BOrrowing COgnate Package) containing a comprehensive and reliable database of Romance cognates and borrowings based on the etymological information provided by publicly available dictionaries in five languages: Spanish, Italian, French, Portuguese, and Romanian (the largest known database of this kind). In this presentation, we will focus on the Romanian language. To answer the first question, we are interested not only in the phylogenetic classification of natural languages but also in the degree of similarity between two languages. Via various techniques and metrics, we offer an answer at three levels: phonetic, lexical, and syntactic. For the second question, based on RoBoCoP dataset we were able to perform the most extensive experiments up to date on a series of HL tasks, including cognates identification, cognate-borrowing discrimination, automatic protoword reconstruction, semantic divergence, etc. for Romance languages, showing that computer-assisted methods where computational methods are integrated with linguistic knowledge is a viable direction for tackling these problems in historical linguistics.

Liviu P. Dinu

ORCID: https://orcid.org/0000-0002-7559-6756
University of Bucharest,
Faculty of Mathematics and Computer Science
Human Language Technologies Research Center
E–mail: ldinu@fmi.unibuc.ro

Proceedings of Workshop on Intelligent Information Systems
WIIS2025, October 16-18, 2025, Chisinau, Republic of Moldova

Next-Generation e-Health Systems: Predictive and Context-Aware Monitoring for Ageing Populations

Invited paper

Marilena Ianculescu

Abstract

Healthcare services need to be reconsidered in the light of the demographic shift towards an ageing population and the increase in the prevalence of neurodegenerative diseases. Conventional methods, with their reactive and piecemeal nature, have not been effective with a quick response. To support predictive and contextual monitoring, the next stage of development of e-Health ecosystems requires the implementation of hybrid cloud-edge structures, artificial intelligence (AI), and Internet of Things (IoT) services. They are regarded not as technological elements but as socio-technical infrastructures in support of healthcare providers, patients with empowerment, as well as regulatory and ethical compliance.

This vision has been the output of decades of medical informatics research at the National Institute for Research & Development in Informatics – ICI Bucharest. The research we conducted demonstrates the dynamism of combining physiological, cognitive, emotional, nutritional, and environmental data sources into reliable and compatible systems. Predictive models on neural networks and hybrid machine learning techniques help in the early detection of age-related risk factors, mobility abnormalities, and mild cognitive impairment. Our monitoring systems can provide comprehensive health profiles and useful recommendations

by taking into account contextual factors such as emotion, stress, and lifestyle.

The architecture of these systems is based on deployments that are distributed across cloud and edge nodes using microservices-based, modular deployments. Scalability, reliability, resilience, and compliance with the European AI Act and the GDPR are obtained based on privacy-by-design, robust consent handling, and comprehensible results. On top of the privacy of health data, security and reliability are essential in ensuring acceptance among patients and physicians.

Decision support is the key component of the future e-Health solutions. Advanced analytics pipelines convert raw multimodal data into information with clinical significance. The outcome is converted into risk scores, practitioner alerts, as well as longitudinal dashboards. The ease-of-use interfaces allow collaboration in decision-making and self-management of patients and caregivers. The technologies are facilitated by a proper user-friendly design that makes them accessible and accommodates the problem of digital literacy when targeting the ageing demographics.

The success of this strategy can be illustrated by the reliable and effective solutions provided by ICI Bucharest that consist of integrated monitoring environments of healthy ageing, neurodegenerative disease prediction modules, emotion and stress recognition modules, as well as nutrition-based decision support. Although it addresses different areas of health, these instances share a common factor, that is, secure and scalable architectures that are interrelated with predictive and context-aware analytics. They show how difficult and how possible changing the research prototypes into solutions with practical applications can be.

There are a number of challenges that still have to be dealt with. Clinical validation is required to ensure that prediction models are flexible. Federated learning and explainable AI methods must be deployed since the heterogeneity of data sets leads to non-IID and inaccurate distribution. Widespread adoption requires interoperability with current healthcare systems, long-term reimbursement pathways, and compatibility with caregiver practices. In the future, the capabilities of e-Health systems will be enhanced with the use of digital twins, adaptive decision support, and real-time risk stratification. These strategies align with

European goals of preventing dementia, active and healthy ageing, as well as long-term transformation of healthcare.

Overall, IoT, AI, cloud computing, and secure data governance converge toward the creation of predictive and context-aware infrastructure in the next-generation e-Health systems. It has been proven to revolutionize cognitive health care and ageing with the solutions that are built on decades of medical informatics research and help to encourage autonomy, support clinical decision-making, and reduce the burden on healthcare systems. To be effective, successful, and accessible, they have to be a balance between clinical evidence, user trust, technological innovation, and societal demands.

Marilena Ianculescu

ORCID: https://orcid.org/0009-0006-8196-0047
National Institute for Research and Development in Informatics,
Bucharest, Romania
E–mail: marilena.ianculescu@ici.ro

Proceedings of Workshop on Intelligent Information Systems
WIIS2025, October 16-18, 2025, Chisinau, Republic of Moldova

Quantification of diversity in natural language processing

Invited paper

Agata Savary

Abstract

The concept of diversity has received increased consideration in Natural Language Processing (NLP) in recent years. This is due to various motivations like promoting equity and inclusion, approximating human linguistic behavior, and increasing systems' performance. Diversity has, however, often been addressed in an ad hoc manner in NLP, and with few explicit links to other domains where this notion is better theorized. We survey articles in the ACL Anthology from the past 6 years, with "diversity" or "diverse" in their title. We find a wide range of settings in which diversity is quantified, often highly specialized and using inconsistent terminology. We put forward a unified taxonomy of why, what on, where, and how diversity is measured in NLP. Diversity measures are cast upon a unified framework from ecology and economy [1], with 3 dimensions of diversity: variety, balance, and disparity. We discuss the trends that emerge due to this systematized approach. We believe that this study paves the way towards a better formalization of diversity in NLP, which should bring a better understanding of this notion and a better comparability between various approaches.

This is a joint work with Louis Estève (Université Paris-Saclay, France), Marie-Catherine de Marneffe (Université Catholique de Louvain, Belgium), Nurit Melnik (The Open University of Israel) and Olha Kanishcheva (Jena University in Germany and the STEP University in Ukraine), within the framework of the UniDive COST Action on Universality, Diversity and Idiosyncrasy in Language Technology.

References

[1] A. Stirling, “A general framework for analysing diversity in science, technology and society,” *J R Soc Interface*, vol. 4, no. 15, pp. 707–719, 2007, DOI: 10.1098/rsif.2007.0213.

Agata Savary

ORCID: `https://orcid.org/0000-0002-6473-6477`
Université Paris-Saclay, CNRS, LISN, France
E–mail: `agata.savary@universite-paris-saclay.fr`

Proceedings of Workshop on Intelligent Information Systems
WIIS2025, October 16-18, 2025, Chisinau, Republic of Moldova

Information measures for digital humanities: a tribute to Solomon Marcus

Invited paper

Andrea Sgarro

Abstract

In his celebrated book *Poetica matematică* (*Mathematical Poetics*) [1], Solomon Marcus, one of the founders of modern computational linguistics, has made use of information measures as are the entropy of Claude Shannon or the informational energy of Octav Onicescu. More precisely, he used them to evaluate aspects of three masterpieces by the Romanian poet Mihai Eminescu, *Somnoroase păsărele*, *La mijloc de codru*, and *Se bate miezul nopții* (*Sleepy little birds*, *In the middle of the wood*, and *The bell strikes midnight*).

Shannon's entropy is a measure of uncertainty and average surprise (uncertainty before an experiment is performed and surprise when the actual result is observed), while informational energy is a measure of redundancy (of lack of surprise). Of the two, Shannon entropy is by far the better known, as it is undisputably the best known information measure anyhow. It has been largely used in what nowadays would be called *digital humanities*: after an initial excitement which led to testing the new tool in literature and psychology, the response has often been negative: Shannon entropy would sometimes give answers which were in sharp contrast with common sense (and good taste in literature).

Actually, a basic feature had been ignored or undervalued: Shannon entropy measures only *statistical* uncertainty, while it ignores other facets of this extremely rich and complex entity, say the pragmatic facet. Needless to say, Solomon Marcus, as he

explicitly stresses, was perfectly conscious of what he was measuring and of what he was not measuring, making it clear that he was dealing only with statistical aspects of Eminescu's poems.

Nowadays the domain called *management of incomplete knowledge* tries to cover all the aspects of incomplete information, where the incompleteness can be due, say, to uncertainty, ignorance or even contradiction – human reasoning can deal successfully with contradicting information, and so should be able to do Artificial Intelligence; the list of facets would be quite long.

Many new non-statistical information measures have correspondingly been introduced; we shall cover in detail the case of the *possibilistic entropy* [2], [3], a sound counterpart to Shannon entropy, since it can be validated both operationally and axiomatically, exactly as is the case for Shannon entropy. Shannon entropy is validated operationally in coding theory, as it gives a measure of how much redundancy is present in a text and how much this text can be compressed without losing information.

As for the axiomatic validation, one lists desirable properties (*axioms*) which a measure of statistical uncertainty should have and then proves that only *one* mathematical formula satisfies all of the axioms; one can use, for example, the axioms chosen by Alexandr J. Hinčin. Instead, possibilistic ignorance pertains to fuzzy logic and multivalued logic rather than probability and statistics, and its mathematical expression is in terms of logical values rather than probability values.

As for the facets of incomplete knowledge, one deals with *ignorance* rather than uncertainty. A list of axioms, which a measure of possibilistic ignorance should satisfy, can be given such that the possibilistic entropy turns out to be the single solution; analogously, a coding theorem can be proven, which supports the operational meaning of the possibilistic entropy in text compression.

Thus, Marcus's experiments could be re-taken so as to gauge other aspects of poems like the ones by Eminescu, which are dealt with in *Poetica matematică*. Information measures might have been disappointing for people who expected too much of Shannon entropy, but the tools that are now available do offer the opportunity to delve deeper and take further the work so brilliantly initiated by Solomon Marcus, as we shall argue.

References

[1] S. Marcus, *Mathematical Poetics*, București: Editura Academiei Republicii Socialiste România, 1970, 404 p. (in Romanian).

[2] A. Sgarro, "Possibilistic information theory: a coding theoretic approach," *Fuzzy Sets and Systems*, vol. 132, no. 1, pp. 11–32, 2002, DOI: https://doi.org/10.1016/S0165-0114(01)00245-7.

[3] A. Sgarro, "An axiomatic derivation of the coding-theoretic possibilistic entropy," *Fuzzy Sets and Systems*, vol. 143, no. 3, pp. 335–353, 2004, DOI: https://doi.org/10.1016/S0165-0114(03)00123-4.

Andrea Sgarro

ORCID: https://orcid.org/0000-0002-7048-1466
University of Trieste, Italy
E–mail: sgarro@units.it

Proceedings of Workshop on Intelligent Information Systems
WIIS2025, October 16-18, 2025, Chisinau, Republic of Moldova

Infogravity Theory: Information–Gravity Dualism and the Emergent Quantum-Gravitational Fabric of Space–Time in a Computational Fractal Universe

Veaceslav Albu

Abstract

The proposed article offers a comprehensive introduction to Infogravity Theory as a scientific cosmological theory that treats the Universe as a computational process whose fractal dynamics and memory are inseparable. The Infogravity is named for its two coupled theoretical pillars – Information (Embedded and Evolutionary [Plirophoria]) Dualism and Gravity (Newtonian/Einsteinian and 4D quaternionic evolutionary−information curvature) Dualism – each grounded in the Strong Conservation of Information in the Universe Postulate: no information ever created can be lost (Strong CIU). The Universe organized as a fractal holarchy acts as a Universal Quantum Computer (UQC) – a distributed, asynchronous deep Boltzmann architecture with atoms and Standard-Model particles serving as input/output heads, executing a self-learning, self-calibrating computation across all scales. The Information–Gravity Dualism couples informational processes to geometric and inertial structure, yielding emergent space-time as Plirophoria – the ongoing storage and propagation of evolutionary information. Landauer's bound ($E \geq K_b \times T \times \ln 2$) and Einstein's $E = m \times c^2$ establish a contextual accounting channel $m = (K_b \times T \times \ln 2)/c^2$ for irreversible record-formation when energy dissipation cannot thermalize in the ordinary way (Einstein-Landauer bridge). We formalize the tripartite holonic fractal architecture of the UQC Universe (baryoholons, cyberholons, and pliroholons, anchored in the previous two), where laws of nature emerge as fixed-points of recursive self-learning – with a fractality whose effective dimension tracks the Density of Emerging Information (with D between 3 and 4), a morphology whose self-similarity coincides with the pattern of Life. Postulates of the Infogravity theory and fundamental properties of evolutionary information (Plirophoria) are concisely and compactly formulated. We articulate falsifiable predictions, near-term room-temperature experiments in living systems, and a dark-sector interpretation in which light Plirophoria (unanchored from baryoholons) maps onto Dark Energy and robust, complex Plirophoria (yielded from cyberholons) maps onto Dark Matter.

Keywords: Infogravity; Informational and Gravity Dualism; Einstein–Landauer bridge; Evolutionary Information-Plirophoria; Density of Emergent Information DEI; DEI quaternionic axis; spacetime emergence.

1 Introduction

The aim of Infogravity theory is to integrate three persistent motifs of modern science – information, computation, and emergence – into a single framework that treats information with the same ontological status as matter and energy. The central empirical intuition is simple: nature persists, and persistence requires selection over possible transitions based on a twofold principle – on occurred previous successful quantum state collapse under similar environmental conditions, and that can guarantee monotonic local growth of the Universe's fractal dimension as Density of Emergent Information. Successful selections are written into the world as locally embedded matter information, and previous selections are simultaneously written into the Universe's neural networks type memory as experience. We elevate this intuition from metaphor to mechanism: space as a neural network evolving structure contains the trace of actual informational embedding in all the Universe's 10^{80} atoms, and the arrow of time is the monotonic growth of that embedding.

Historically, physics has engulfed two ontologies of law: a regularist view, in which laws are summaries of observed regularities, and a necessitarian view, in which laws constrain what can happen. Both readings have motivated successful research programs, but neither, standing alone, explicates how laws act with immediacy across a 3D world or how they co-compose. The analogy with virtual worlds [1] hits the goal: in a simulator, laws are algorithms that update states. They do not travel through space; they compute state transitions wherever relevant data reside. By analogy, Infogravity places the Universe's lawful regularities in an informational dimension that is simultaneously active upon the 3D world.

1.1 Core Infogravity Theory Postulates

The main hypothesis of the Infogravity theory - Universal Quantum Computer (UQC) hypothesis, describes the Universe as a distributed, asynchronous quantum deep Boltzmann machine with atoms as input/output heads and topologically encoded evolutionary information as its emergent recurrent inner layers.

Postulate 1 – Strong Conservation of Information (SCI)

All information ever created in the evolving Universe remains at the Universe's disposal and cannot be destroyed. Apparent loss is re-encoding, export, or redistribution across potential, emergent, and embedded capacities. This pressure toward robustness favors topological encodings and motivates the Information and Gravity dualisms.

Postulate 2 – Information Dualism

The dual nature of Information by natural kinds manifests in reality from one side, as the classical (corpuscular, collapsed) state of information embedded in the Universe's matter during the evolution of all classical objects and phenomena of the real world. From the other side, it manifests as unchangeable by its content and physical volume, gravitationally and topologically encoded quantum evolutionary information (or Plirophoria) about the history of all collapsed quantum states during the evolution of each classical object of any scale in the Universe, as the totality of all emerged classical objects during the Universe's lifespan.

Postulate 3 – Gravity Dualism

From the first two Postulates follows the Gravity Dualism. On one side, the standard 3D Newton-Einstein curvature (force), responsible for the embedded information encoded in matter/energy interactions. On the other side, 4D quaternionic curvature, as it emerges robustly, topologically encodes previous quantum states as an evolutionary information of a given qudit of the Universe. The second side of Gravity Dualism rests on the mass-energy Einstein–Landauer bridge, which reflects the gravitational nature and cost of emergent space-time [16]. That Dualism imposes scale scale-dependent appearance of gravity. Within galaxy classical objects, it appears as a force, driven by space-time curvature. Intergalactic space and voids are created by its complementary part – robustly topologically encoded evolutionary information. Density of Emergent Information is maximally valued at all fractal scales within a galaxy, and it takes its minimal values in intergalactic voids. That also explains the filament structure of the Universe on large scales.

Postulate 4 – Einstein–Landauer Coupling

Logical irreversibility of the evolving Universe as a Quantum Computer by natural kinds [15] dissipates heat (Landauer, $Q_{min} = Kb \times T \times \ln 2$), and energy and mass are equivalent (Einstein, $E = m \times c^2$). Hence, sustained information processing can contribute to mass-energy accounting wherever a durable organization (ΔU) is written and kept – mass-energy Einstein – Landauer bridge. To say instead that information "has" mass is totally different and does not describe reality.

Postulate 5 – Generated Space

Space is Emergent from the Informational Topological Structure of encoded evolutionary information about irreversible changes of quantum states of each of 10^{80} atoms of the Universe and particles of the Standard Model as input/output layers under the UQC hypothesis. Every interaction as a collapse event that encodes Plirophoria into the Universe leaves a gravitational imprint, effectively making the fabric of spacetime a memory field. In other words, Space represents the durable ledger written by successful transitions – minimally as pliroknots; collectively as plirobraids and pliroholons, that according to infogravity ontology, represent recurrent inner neural networked layers of UQC as a Boltzmann machine. Statistical laws that resemble Bayesian inference and Matthew effect guide successful transitions in this plirophoric realm.

Postulate 6 – Holarchic Dynamical Fractal

The Universe evolves as a tripartite holarchy organized as a dynamical fractal whose self-similarity echoes Life's patterning and its effective fractal dimension covaries with Density of Emergent Information (DEI) and lies between 3 and 4 dimensions. Tripartite holarchy consists of Baryoholons (mass–energy carriers), Cyberholons (living/cybernetic agents), and Pliroholons (informational structures, acting as Markov blankets for the first two types of holons) co-act to compute natural-kind dynamics of the Universe. Each class has distinct constraints and contributes to DEI.

Postulate 7 – Quantum Border of Fields (3D ↔ 4D interface)

The quantum realm functions as an effective border between the 3D classical baryonic world and a 4D quaternionic Plirophoria field (the informational substrate of stored evolutionary information). At this border, advanced/retarded symmetry (Wheeler–Feynman) encodes constraint injection (from the absorber) and state confirmation (to emitter I/O heads), enabling DEI-guided selection and curvature/memory updates (Cramer). Quantum exchanges are routed by offers and confirmations across holonic boundaries. The environment's plirophoria absorbs advanced proposals; confirmations bias retarded guidance toward collapses that maximize DEI under homeostatic and energetic budgets. This postulate selects the channel; the Einstein–Landauer coupling prices the commit (heat vs ΔU vs non-thermal). The Strong Conservation of Information law holds – no energy creation.

1.2 Key Concepts & Properties.

1.2.1 Plirophoria

Plirophoria is evolutionary information, as a complementary part of embedded in matter information that follows from the Information Dualism Postulate. It represents a robust recording of previous states of the embedded information. State 'NOW' of the embedded locally in matter information as its quantum states become state "PAST", for the actual system, and a new "NOW" state represents a collapsed state of the system. Simultaneously, the "PAST" state of the embedded information becomes actualized and robustly memorized as evolutionary information about that system. It acts as a record, resource, and regulator.

1.2.2 DEI (Density of Emergent Information)

DEI operationalizes the rate at which selection-relevant information is produced per unit structure/time. Practical proxies include: (i) compression-based complexity, (ii) multi-scale redundancy (mutual information; persistent homology), (iii) controllability/observability metrics, and (iv) resilience indicators. Under the holarchic-fractal postulate, effective fractal dimension and DEI co-vary [4].

1.2.3 Pliroknots & Plirobraids

A pliroknot is a minimal durable record; plirobraids are entangled chains that protect and propagate information across distance and scale. They are analogues – not identities – of topological protection; their presence should bias motion and signaling in error-suppression experiments.

1.2.4 Quaternionic Channels

Three orthogonal recursion families plus a scalar route define four channels whose compositions support concurrent law-learning and error correction. This is bookkeeping for how multiple organizational modes co-evolve without destructive interference (quaternionic decomposition is a modeling convenience, not an ontic claim).

1.2.5 Derived Time ↔ Generated Space

Time tracks $\Delta_{info} \geq 0$ – the monotonic accumulation of embedded information; space is the durable ledger of successful transitions. Together, they form a coupled index: recorded memory thickens geometry while Δ_{info} paces temporal order.

1.2.6 Law Learning (derivative principle)

Regularities in plirophoric recurrent inner layers we call laws and are DEI attractors – policies that optimize emergent information under stability costs. Where DEI is high, learned laws persist; where DEI is low, they drift. This derives from Postulates 1–7 and follows the autodidactic Universe principles [23], [15].

1.3 Evidence & Predictions (selected)

- Landauer's principle (1961; 1991) [2, 3] – Irreversible operations dissipate $= k_B T \ln 2$ heat/bit. Informational processes are caloric; Φ_{info} cannot be free.
- Einstein's Energy – Mass equivalence (1905) [4] – Combined with Landauer, durable information histories can be gravitationally relevant in context.
- Deutsch's universal computation (1985) [5] – If the world supports universal computation, 'laws' can be read as learned programs shaped by feasibility and selection.
- Wiener's cybernetics (1961) [6] – Feedback carves islands of stability; those islands are records that shape later dynamics.
- Mandelbrot's fractals (1982) [7] – Multiscale morphology implies measurable redundancy; DEI should correlate with fractal features.

- Verlinde's emergent gravity (2010; 2016) [8, 9] – Microstate/entropic gradients can mimic gravity; Infogravity supplies a memory ontology and SCI constraints.
- Computable universes [10, 11] – Laws may be emergent; Infogravity adds falsifiers and measurement pipelines.
- Biological information density [12, 13, 14] – Biospheres are high-DEI engines, ideal for probing information–energy couplings.

2 The Universe as UQC: a distributed Boltzmann architecture with atomic I/O heads

Premise. The Universal Quantum Computer (UQC) hypothesis describes the Universe as a distributed, asynchronous quantum deep Boltzmann machine with recurrent layers, operating not on an external clock but on locally available physical transitions [15]. In this view, atoms and particles constitute the input/output heads of a universal computation: each collision, emission, absorption, scattering, and binding quantum event is both (i) state propagation in Hilbert space and (ii) message passing through the fabric that we retro-describe as geometry.

What does "Boltzmann" mean here? "Boltzmann" emphasizes thermodynamic learning: state distributions relax to energy-weighted equilibria that encode and compress informational regularities [15]. The deep aspect points to hierarchies of interaction scales – from quarks and leptons up to galaxies – through which latent variables ("hidden layers") become effective laws when stabilized by recurrent feedback.

Asynchrony. No global clock. The Universe updates via local interactions driven by free energy gradients and selection of dissipative channels; global coherence emerges from dense local couplings [15, 16, 17].

Atoms as I/O heads. In this UQC, atoms are not mere passive labels; they are ports through which messages (photons, phonons, excitations) are written, read, and validated. The selection rules and line strengths become routing constraints of the cosmic computation [15]. When a state is measured (coarse-grained by an environment), the output is written into accessible channels (radiation, matter rearrangements), while the input is the ensemble of microstates that seeded the transition. One can imagine this by drawing atomic heads along UQC layers, with recurrent informational feedback running through thermodynamic flows [17].

Why is this architecture testable? In a UQC Universe, information is not epiphenomenal; it is operative – it constrains energy pathways and geometry via Information and Gravity dualisms. This implies observable budgets, rate limits, and energy/gravity bridge signatures whenever logically irreversible updates occur. In this model, the Universe functions as a distributed, asynchronous quantum deep Boltzmann machine, composed of nested recurrent informational networks that evolve by natural kinds.

- Its architecture is not engineered, but emergent: Baryoholons, Cyberholons, from Pliroholons to Pliroknots, Plirostrands, Pliro-braids;
- The UQC model includes the following conceptual features:
  - Quantum coherence supports nonlocal informational flow.
  - Deep neural structure enables self-similar recursion.
  - Asynchronous evolution mirrors the causally open nature of the cosmos.
- Each layer of reality – from quantum fields to biological systems – acts as a computational holarchy: a recursive layer that processes, topologically encodes, and transmits the evolutionary information – Plirophoria.

3 Information–Gravity Dualism from Strong Conservation of Information

Strong Conservation of Information (SCI). No information ever created in the Universe can be lost. Information may be copied, dispersed, encrypted, coarse-grained, decohered, or entangled – but not annihilated. This is not a metaphor; it is a governance law: SCI is the axiom from which Infogravity's dualisms follow [18, 1, 19].

Dualism statement. If SCI holds, then the energetic cost of logically irreversible updates (erasure, coarse-graining, record consolidation) must be paid somewhere, tracked, and preserved in the global ledger. The Landauer bound fixes a minimum heat per erased bit at temperature T:

$$Q_{min} = k_B \times T \times \ln 2.$$

Einstein's relation, $E = m \times c^2$, guarantees that energy retained (as durable structure, binding work, curvature) contributes to inertial/gravitational accounts. Therefore, whenever a system writes durable information, the Universe must allocate the Landauer-mandated cost to (i) immediate heat, (ii) internal energy of organization (ΔU), and (iii) curvature/inertial budget.

This yields the Information–Gravity Dualism: changes in informational organization – especially durable records – co-determine gravitational/inertial structure. In practice, the dualism is channel-wise: we examine how much cost appears as heat versus organizational energy versus non-thermal escapes (e.g., neutrinos), under SCI constraints.

- In the Infogravity framework, gravity is not a primitive force, but an emergent informational field generated through the recursive computation of evolutionary information – Plirophoria.
- Every interaction that encodes Plirophoria into the Universe leaves a gravitational imprint, effectively making the fabric of spacetime a memory field.

- Collapse Events in QM as Irreversible Computations by Natural Kinds.
- Absorber and Transactional Frameworks as Precedents: Wheeler–Feynman absorber theory [20] & Cramer's transactional interpretation [21].
- The Einstein-Landauer Mass-Energy-Information Equivalence Principle ($M = kT \ln 2/c^2$)

4 Plirophoria: emergent space–time and storage of evolutionary information

Emergent space–time. In Infogravity, what we call the fabric of space-time is the macroscopic expression of emergent quantum gravity that tracks and stabilizes Plirophoria [17]. Curvature becomes the shadow of information gradients: cohesive, low-entropy records pull on the ledger via $E = m \times c^2$, while flows of free energy represent informational work that shapes geometry over time.

Storage and export. Plirophoria stores within anchored pliroholons (e.g., genomes, neural connectomes, crystalline memories, encoded fields) and exports through unanchored pliroholons (e.g., radiative messages, chemical signals) that weave extended Markov blankets binding holons into ensembles. Infogravity explicitly ties this to evolutionary information as a co-cause of structure and to emergent geometry through dualism.

4.1 Plirophoria: Definition and Properties

Plirophoria ($\pi\lambda\eta\rho o\varphi o\rho\iota\alpha$) is the cumulative evolutionary information accumulated by natural kinds within an object from inception to end. It is quantum, gravitationally/topologically encoded, and unchangeable in both content and physical volume once stabilized [15, 22]. Methodologically, Plirophoria instructs us to follow the writes across scales: cyberholons learn; pliroholons remember; space-time carries the memory forward (P1–P5). Plirophoria names the ever-expanding body of evolutionary information the Universe stores as it computes itself. Plirophoria is both content and carrier: a dynamic holon field of pliroholons – informational entities with anchoring to specific physical substrates (baryonic or cybernetic).

4.2 Plirophoria's Seven properties [17, 18, 19]

Objectivity. Universal information participates causally and is retrievable across scales, enabling budgeting akin to energy and mass.
Duality. Stability (records) and generativity (novelty) co-produce cumulative complexity; suppress one, and evolution stalls.
Fractality. Organization exhibits self-similar structure across scales consistent with holarchic architectures, self-similarity with the pattern of phenomenon of Life in the Universe, and with fractal dimension as Density of Emergent Information with effective dimension between 3 and 4.
Homeostaticity. Information underwrites control loops that maintain viability while permitting adaptation; homeostasis is dynamic regulation, not stasis.
Imaginarity. Complex/quaternionic formalisms capture phase-like relations and concurrent channels in information dynamics.
Extradimensionality. Effective informational coupling can mimic extra degrees of freedom in emergent dynamics, presenting alongside 3D gravitational effects.
Evolutionarity. Usable records accumulate monotonically ($\Delta info \geq 0$), while remaining open to novelty; this arrow aligns with derived time in UQC.

5 Tripartite holonic fractal architecture of the UQC Universe

Holarchic triptych. Infogravity models the Universe as a tripartite holonic fractal (baryoholons, cyberholons, and pliroholons), which follows from the same fundamental ontological status of matter, energy, and information in the Universe.

Baryoholons – holons anchored primarily in baryonic matter (particles, atoms, stars, galaxies). The baryoholons realm remains responsible for the creation of diversity within baryonic matter that, in their turn, becomes input blocks for the cyberholonic realm creation and evolution.

Cyberholons – living and life-like adaptive agents (cells, tissues, organisms, ecosystems, cognition), bounded by Markov blankets ("anchored pliroholons").

Pliroholons (unanchored) – informational quanta that ferry updates between holons: photons, phonons, neutrinos, codes; they stitch holons into cyber-physical ensembles.

Dynamical fractal, life-pattern similarity. The holarchy is fractal: self-similar organization repeats across scales; its self-similarity pattern coincides with the pattern of the phenomenon of Life in the Universe [18]. The effective fractal dimension, D_eff, tracks the Density of Emerging Information (DEI) and typically lies between 3 and 4 [1, 15]. Where DEI rises, the direction of evolution (in state space) points to richer holonic nesting and law-emergence via self-learning.

Laws as learned invariants. In the UQC, laws of nature are stable fixed-points of the self-learning recursion – emergent invariants discovered, selected, and retained because they minimize free energy of representation across scales [15, 16]. In this sense, "law" is the name we give to a reliable statistical structure with low description length under repeated re-computation.

Markov blankets in practice. Each holon has a boundary that screens internal from external states while passing sufficient statistics – the holon's anchored pliroholon. In cyberholons, blankets become biological membranes and sensorimotor interfaces; in baryoholons, blankets include gravito-electromagnetic horizons and transport shells (e.g., stellar radiative/convective zones). The membrane-calculus view (Păun) formalizes the nested computation of these blankets [18, 1, 19, 22].

6 Dark Sector as a Partition of Plirophoria

The Infogravity theory proposes a structured split tied to organizational complexity.

Light Plirophoria (LP), yielding by unanchoring baryoholons $\leftrightarrow$ Dark Energy: astronomically numerous, low-complexity pliroholons produced in fusion cycles stream outward and accumulate in outskirts/voids, yielding a tension-like, repulsive-looking effect at large scales.

Complex Plirophoria (CP), unanchoring cyberholons $\leftrightarrow$ Dark Matter: robust, topologically complex pliroholons produced by cyberholons and other long-duration information architectures concentrate inside galaxies and behave as an additional mass component [15, 22].

Predictions. LP density correlates with the integrated star-formation history (SFR), anti-correlates with dense galactic interiors, and enhances in the void outskirts. CP density correlates with complexity proxies – metallicity, exo-planet incidence, chemotactic free-energy density, bio-proxies, or generic DEI. Lensing with morphology-aware priors should reveal residual structure aligned with informational organization, not just baryon density.

Falsifiers. If rotation curves and lensing maps are fully explained by baryons + standard halos + systematics with no residual correlation to complexity/bio-proxies, CP is weakened; if LP shows no outskirts/void gradient against SFR histories after controls, LP is weakened [15, 22].

The methodological demand is preregistration: let information organization compete with standard parameters fairly. In other words, Light Plirophoria – simple, weakly anchored informational organization – behaves as dark-energy-like pressure distributed in outskirts and voids: its pliroholons are light, less holon-bound, and push out on large scales. Complex Plirophoria – robust, holonic architectures (life-like, cybernetic, self-repairing) – aggregates within galactic interiors where cyberholonic formation dominates; its anchored pliroholons are stiffer, adding mass-like loading – dark-matter-like behavior. This allocates information-dynamical roles to galactic morphology without resorting to exotic particles: the dark sector is, in part, how Plirophoria distributes across holons.

7 The Observer Problem in Infogravity

7.1 Observer as holon boundary

In Infogravity, an Observer is not an external homunculus; it is a pliroholon ancored on a given baryo- or cyberholon and effectively resembles and forms its Markov blanket, which separates internal entangled states from external entangled states of the plirophoric realm, while transmitting sufficient statistics (sensory inputs, active outputs) through its inputs/outputs baryonic fractal layer.

"The collapsing active state of each plirohead of each Universe's holon is aimed to maximize the DEI in involved baryomatter and cybermatter with the purpose to produce new material objects of increased complexity, that guarantee new knowledge and data within its evolving and stored plirophoria. For this purpose, any forward incoherent state of each active plirohead within its quantum states of a given matter occurs with the probability that increases the future DEI by embedding a relevant quantum state in a given matter. By its essence, the learned laws of nature represent the Observer within each active plirohead that is connected with the Universe's Control Center, where all laws of nature nest" [15].

Because SCI forbids annihilation, any coarse-graining or collapse (in the effective description) is a physical bookkeeping event: it must deposit its informational cost into heat, ΔU, or non-thermal channels. This is what binds observation to the Einstein–Landauer budget. Infogravity contrasts the "Scientific Physical Reality" vs an Info-Computational Universe, explicitly noting that in the latter, the observer is part of the system observed and information is ontologically fundamental. That stance collapses the old observer/external-world dichotomy into interacting holons participating in the same UQC dynamics.

7.2 Quantum Darwinism, decoherence, and Plirophoria (Infogravity view)

While standard accounts invoke environment-induced decoherence and quantum Darwinism for consensus states, Infogravity reframes the result: pointer states are those whose writes are cheapest for the blanket to keep while maintaining homeostasis and predictive power. They become anchored pliroholons, adding to Plirophoria and stiffening local geometry. This view respects the Infogravity stance that space-time is emergent from informational topology and grounds the Observer as a computational membrane that learns.

8 Transactional/Absorber Mechanism

8.1 Offer–confirmation cycles as routing of writes

Postulate 7 explicitly cites absorber and emitter theory, situating Wheeler–Feynman and Cramer's Transactional Interpretation as conceptual precedents. Infogravity embraces a pragmatic reading: "offers" (forward-in-time proposals) and "confirmations" (selection/handshakes) are real routing constraints on informational exchange between holons. They govern which channels light up, which messages (unanchored pliroholons) are accepted/rejected, and how records become anchored (e.g., memory, structure). This is not a metaphysical adornment; it is a control plane for channel selection in a UQC Universe.

8.2 Observer + Transactional = who selects, who pays

The Observer (holon boundary) selects channels based on internal/external priors and costs.

The Transactional/Absorber mechanism routes and confirms exchanges at scale.

The Einstein-Landauer budget prices the commitment of a record.

In tandem, these "budgeting" mechanisms ensure that writes occur only when paid for and kept (SCI), with routing shaped by offers/confirmations and selection executed at holon blankets.

9 Conclusions

Infogravity advances a computational cosmology where information is constitutive of structure. If the Universe is UQC-by-Natural-Kinds, and if SCI holds without exception, then dualism follows: the cost of irreversible updates is paid somewhere – as heat, as durable internal energy (organization, ΔU), and, by Einstein, on the inertial/gravitational ledger.

At stellar scale, the Ockham-minimal multiplicity $B_{th} = Q/(k_B \times T \times \ln 2)$ is inevitable for thermalized fusion, making channel partition (thermal vs non-thermal) the central astrophysical question. At the life scale, the daily Landauer floor at physiological temperature is large, yet no persistent surplus heat appears beyond metabolic budgeting in controlled settings; the natural reconciliation is that a fraction of the information-mandated cost is retained as ΔU during durable writes – a life-scale Einstein-Landauer bridge.

The tripartite holonic fractal – baryoholons, cyberholons, pliroholons – organizes this picture into a dynamical hierarchy whose self-similarity mirrors Life, and whose fractal dimension tracks Density of Emerging Information. In this sense, the laws of nature are the conserved theorems of a self-learning Universe. In the UQC model, quantum systems evolve as components of a distributed, asynchronous computational architecture – a natural deep Boltzmann machine embedded in the cosmos itself. The Infogravity theory bridges: a) Physics and cybernetics; b) Ontology and systems theory; c) Computation and emergence.

Acknowledgments. The author wishes to thank Svetlana Cojocaru, Vice President of the Academy of Sciences of Moldova, Corr. Memb., Prof., for important and constructive guidance of the research path; Constantin Gaindric, Corr. Memb., Prof., and Professor Averkin A. N. for their time and productive discussions on the above topics. Project SIBIA – 011301, "Information systems based on Artificial Intelligence," has supported part of the research for this paper.

References

[1] V. Albu, "Argument for Existence of the Fourth Dimension of Laws of Nature and Its Informational Unity with the 3D Space of the Real World from Analogy with a Virtual World," BA thesis submitted to the University of Wales Trinity Saint David, Lampeter, UK, 2017.

[2] R. Landauer, "Irreversibility and heat generation in the computing process," *IBM Journal of Research and Development*, vol. 5, no.3, pp. 183–191, 1961.

[3] R. Landauer, "Information is Physical," *Physics Today*, vol. 44, no. 5, pp. 23–29, 1991.

[4] A. Einstein, "Does the inertia of a body depend upon its energy content?" *Annalen der Physik*, vol. 18, pp. 639–641, 1905.

[5] D. Deutsch, "Quantum Theory, the Church–Turing Principle and the Universal Quantum Computer," in *Proceedings of the Royal Society A*, vol. 400, no. 1818, 1985, pp. 97–117.

[6] N. Wiener, *Cybernetics: Or Control and Communication in the Animal and the Machine*, MIT Press, 1961.

[7] B. Mandelbrot, "The Fractal Geometry of Nature," W.H. Freeman, 1982.

[8] E. Verlinde, "On the Origin of Gravity and the Laws of Newton," arXiv:1001.0785, 2010.

[9] E. Verlinde, "Emergent Gravity and the Dark Universe," arXiv:1611.02269, 2016.

[10] R. Penrose and H. Zenil, "A Computable Universe," World Scientific, 2013.

[11] S. Alexander, W. J. Cunningham, J. Lanier, and L. Smolin, "The Autodidactic Universe," arXiv:2104.03902, 2021.

[12] J. S. Torday, "Homeostasis as the Mechanism of Evolution," *Biology*, vol. 4, no. 3, pp. 573–590, 2015.

[13] J. Gaite, "The Fractal Geometry of the Cosmic Web and its Formation," in *Advances in Astronomy*, 2019, DOI: https://doi.org/10.1155/2019/6587138.

[14] Y. M. Bar-On, R. Phillips, and R. Milo, "The biomass distribution on Earth," *PNAS*, vol. 115, no. 25, pp. 6506–6511, 2018.

[15] V. Albu, "The Universe as a Universal Quantum Computer: A Road Map for Its Possible Architecture," in *Proceedings of the Workshop on Intelligent Information Systems WIIS2022*, (Chisinau, Republic of Moldova), October 06-08, 2022, pp. 14–43.

[16] V. Albu, "The Computing Quantum Universe and Einstein-Landauer's Principle of Mass-Energy-Information Equivalence," in *IMCS60*, 2024, pp. 220–226.

[17] V. Albu, "Infogravity – A Unified Framework for Evolutionary Information, Emergent Quantum Gravity, and the Fabric of Space-Time," Poster in *QISS*, 2025.

[18] V. Albu, "An Inquiry into Indispensability of Information in Hylomorphic Unity of the Universe and the Phenomenon of Life," MA thesis submitted to the University of Wales Trinity Saint David, Lampeter, UK, 2018.

[19] V. Albu, "Encapsulating the Concept of Information in Scientific Paradigm of Reality: A Possible Scenario of a Theory of Everything," in *IMCS-55*, 2019, pp. 275–282.

[20] J. A. Wheeler and R. P. Feynman, (1945–1949). Absorber theory papers: *Rev. Mod. Phys.*, vol. 17, 1945, pp. 157–181; *Rev. Mod. Phys.*, vol. 21, 1949, pp. 425–433.

[21] J. G. Cramer, "The Transactional Interpretation of Quantum Mechanics," *Reviews of Modern Physics*, vol. 58, no. 3, pp. 647–688, 1986.

[22] V. Albu, "Infogravity: How the Universe's Evolutionary Information Possibly Explains Emergent Quantum Gravity and Fabric of Space-Time," in *WIIS 2023*, 2023, pp. 24–36.

[23] Stephon Alexander, William J. Cunningham, Jaron Lanier, Lee Smolin, Stefan Stanojevic, Michael W. Toomey, and Dave Wecker, "The Autodidactic Universe," arXiv:2104.03902v1 [hep-th], 29 Mar 2021.

Veaceslav Albu

ORCID: `https://orcid.org/0009-0004-4764-8266`
Moldova State University, "V. Andrunachievici" Institute of Mathematics and Computer Science, Chisinau, Republic of Moldova
E–mail: `veaceslav.albu@math.md`

Proceedings of Workshop on Intelligent Information Systems
WIIS2025, October 16-18, 2025, Chisinau, Republic of Moldova

Adaptive Computing Infrastructure for Research and Education Support

Peter Bogatencov

Abstract

In the article, approaches and solutions for the development of an adaptive distributed computing infrastructure are presented, oriented to adaptive reconfiguration for the operation of different workflows depending on users' needs. The common infrastructure is dedicated to supporting different types of workflows and providing efficient access to various types of computing resources and services for research data analytics and the operation of production services supporting different scientific and educational activities.

Keywords: information technologies, distributed computing, e-infrastructure & services for research and education.

1 Introduction

Real demands of the development of computing infrastructure and services to support the accumulation, storage, and processing of research data are constantly increasing. Development of distributed, High-Throughput and High-Performance computing (HPC) infrastructures for solving complex tasks with specific demands of computing resources, abilities to store and access increasing amounts of research data are actively growing now, including in Moldova [1]. This requires permanent expansion of hardware resources and the use of new tools to automate resource allocation, equipment management, and service delivery. The introduction of modern architectural solutions, tools, and platforms, as well as the installation of new high-performance computing resources, has made it possible to intensify the use of distributed computing infrastructure. The concept of creating an adaptive distributed heterogeneous computing infrastructure has been developed. The proposed adaptive computing infrastructure offers conditions for providing requested computing and storage resources over a high-speed optical network.

End-users are looking for a wide range of services for working with data and computing resources for elaboration and executing complex applications and new services. The created infrastructure is oriented to support open science and be integrated into the European Open Science Cloud Initiative (EOSC), aimed at the accumulation of various scientific information for the organization of open access to research information and services that have a significant impact on the intensification of the use of distributed computing resources. Important areas of work in this direction focused on the deployment of new types of computing infrastructures that will integrate High-Throughput, HPC, and storage computing resources and gain benefits to end users from uniting computational resources of multiprocessor clusters with effective application platforms, users' interfaces, and infrastructure management tools offered by Cloud infrastructure.

2 Deployment of distributed heterogeneous computing infrastructure

Work on the implementation of the distributed computing infrastructure in Moldova started in 2007 when the first Agreement on the creation of the MD-GIRD Joint Research Unit Consortium and accompanying Memorandums of Understanding were signed by seven universities and research institutes of Moldova. Since this time, the works started on the deployment of the national computing infrastructure that included the integration of computing clusters and servers deployed in the main national universities and research institutions. For effective integration of different types of computing resources into the common distributed infrastructure, a high-capacity communication backbone was used, provided by the National Research and Educational Network RENAM [2].

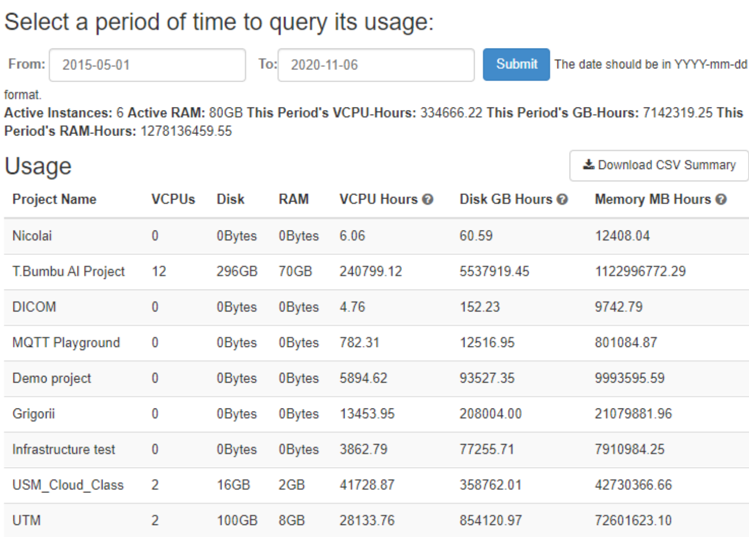
In 2014-2015, as a result of the participation of the Institute of Mathematics and Computer Science (IMCS) and partner organization RENAM in the regional project “Experimental Deployment of an Integrated Grid and Cloud Enabled Environment in BSEC Countries on the Base of gEclipse (BSEC gEclipseGrid)” supported by Black Sea Economic Cooperation Programme (http://www.blacksea-Cloud.net), resources of the distributed computing infrastructure were extended and new technologies were deployed [3].

The experience and results accumulated during this project realization allowed deploying Cloud infrastructure for applications development, evaluation, and testing, based on OpenStack version 13 Mitaka (release 2016), installed and accessible at https://cloud.renam.md/. It was deployed using two computing nodes of the multiprocessor cluster operated in IMCS, which were taken out of the mainstream configuration of this computing facility. The total amount of resources of these two servers was quite modest even for this time - only 16 CPU cores, 32 GB RAM, 750 GB HDD, 1 Gbit/s network. Despite this, the created infrastructure was widely used, both for new applications evaluation purposes, and in several short-term projects, in which it was used to deploy quickly a small virtual infrastructure to test various scenarios and products.

In 2018, the infrastructure was extended by modern high-performance servers with 32*n CPU threads, 128 GB RAM, 4 Tb RAM storage, and 10 GB network cards. This upgrade has opened a new stage in the development of the computing facility at IMCS – additional resources for several institutional projects in the field of Machine Learning and Neural Language Processing were provided and actively used (see Fig.1).

In 2019, the Cloud infrastructure of IMCS began to offer a new service – support of online lectures organization in the Moldova State University (MSU) and the Technical University of Moldova. The experience gained over these years and obtained users’ feedback allowed us to specify clear needs for further development of the computing infrastructure in IMCS. However, a simple increase of resources was no longer an option – the technologies used in the existing OpenStack releases no longer met modern security standards, and the manual installation process applied to upgrade OpenStack created insurmountable difficulties for infrastructure expansion and administration.

To overcome the known limitations of the initial realization of distributed computing infrastructure in Moldova, the Faculty of Mathematics and Computer Science of MSU, IMCS MSU, and RENAM united efforts for developing and providing resources of the innovative federated computing environment that ensures the efficient use of available hardware and software resources, adaptive allocation of resources, accumulation and development of various workflows, software platforms, tools, technologies for research data analytics, research and education services deployment and operation, solving problems that require high-performance computing capacity and processing large amounts of data.



Select a period of time to query its usage:

From: 2015-05-01 To: 2020-11-06 Submit The date should be in YYYY-mm-dd format.

Active Instances: 6 **Active RAM:** 80GB **This Period's VCPU-Hours:** 334666.22 **This Period's GB-Hours:** 7142319.25 **This Period's RAM-Hours:** 1278136459.55

Usage

Download CSV Summary

Project Name	VCPUs	Disk	RAM	VCPU Hours	Disk GB Hours	Memory MB Hours
Nicolai	0	0Bytes	0Bytes	6.06	60.59	12408.04
T.Bumbu AI Project	12	296GB	70GB	240799.12	5537919.45	1122996772.29
DICOM	0	0Bytes	0Bytes	4.76	152.23	9742.79
MQTT Playground	0	0Bytes	0Bytes	782.31	12516.95	801084.87
Demo project	0	0Bytes	0Bytes	5894.62	93527.35	9993595.59
Grigorii	0	0Bytes	0Bytes	13453.95	208004.00	21079881.96
Infrastructure test	0	0Bytes	0Bytes	3862.79	77255.71	7910984.25
USM_Cloud_Class	2	16GB	2GB	41728.87	358762.01	42730366.66
UTM	2	100GB	8GB	28133.76	854120.97	72601623.10

Figure 1. Example of computing resource usage in IMCS MSU

Due to participation of Moldova in the EU-funded project EU4Digital – EaPConnect, computing facilities operated in MSU, RENAM, IMCS, and also in the Institute of Emergency Medicine, were upgraded as follows (Fig. 2): new high-performance servers for computing nodes and storage elements were installed; a 10 Gbps optical network was deployed between the main computing sites; IMCS and RENAM servers' rooms were modernized – modern uninterruptible power supply systems for server equipment and industrial air-cooling systems were installed. Current activities focused on the deployment and development of HPC cloud infrastructure by implementing SaaS, PaaS, and IaaS service models for the joint use of parallel computing clusters integrated into the common distributed computing environment. The available computing resources now comprise more than 400 CPU cores, 2 NVIDIA T4 Tensor Core GPU units, and over 54 TB of RAM storage [4].

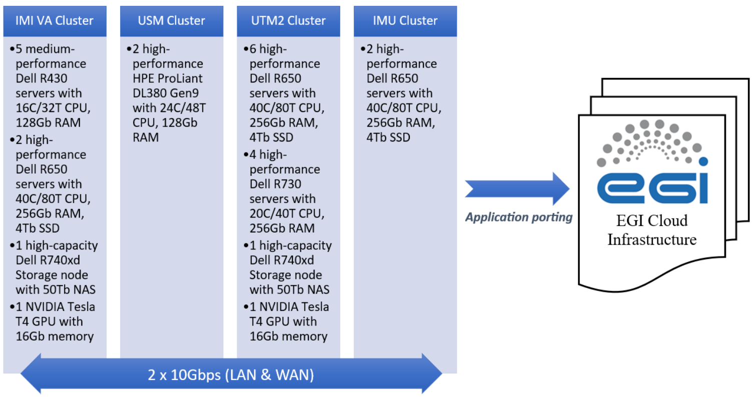


Figure 2. Parameters of the distributed computing infrastructure

Distributed computing infrastructure uses the adaptive execution frameworks that can be adjusted for the execution of different complex applications, specific services, with the possibility to certify and onboard some of them as generic services to EOSC. The research teams from MSU, IMCS MSU, Institute of Applied Physics MSU, and Institute of Chemistry MSU developed several applications that require scalable resources of multiprocessor clusters integrated in the common distributed heterogeneous computing infrastructure.

The adaptive allocation of computing resources is used for deploying the following types of computing infrastructures dedicated to the execution of specific workflows, depending on users' requirements and production operation of a predetermined set of commonly used services and applications:

- Deployment of multiprocessor clusters for computing modeling, complex applications development, and execution;
- Heterogeneous installations that include GPU cards for artificial intelligence (AI) models development and machine learning support;
- Allocation of resources of virtualized and bare metal servers, software platforms for production operation of widely used services and applications;
- Scalable computing resources for scientific cloud infrastructure operation;
- Storage servers' infrastructure.

3 Adaptive Configuration of the available resources

The fact is that the existing heterogeneous infrastructure, especially the deployed OpenStack Cloud infrastructure is a very flexible and, as a result, a very complex product, consisting of hundreds of open-source "bricks" components combined into a single system, which have many dependent on each other services and components, with their own settings, which must be written into the configuration files before executing the commands to "dock" the components into a single working environment. Mistakes when installing and configuring such a complex system are inevitable and sometimes irreversible, and necessitate reinstalling all components from scratch. This approach makes further administration and scaling of the system an absolutely impractical and complicated process.

Installing and configuring heterogeneous infrastructure manually is great for getting familiar with the internals of a system and understanding how its components interact. However, when installing the system on large infrastructures, where clustering of more than 5-10 servers is used, manual installation is an absolutely inappropriate approach and is unpromising in terms of further scaling. To overcome these limitations, increase the reliability and the possibility of further upgrading and scaling of such a sophisticated system, a completely different approach is required.

To solve this problem, it was decided to use a modern approach to the infrastructure configuration and administration, usually applied to cloud systems – using Deployment Tools, which allow creating scripts to automate the installation of the system. It is difficult to imagine a modern IT project without applying such solutions. About a dozen of such kind of tools are already known to automate configuration processes; the main ones and being used more often are Ansimble, Puppet, Chef, Juju, ManageX. To deploy a new or improve the existing computing infrastructure on the basis of modern equipment, we have chosen as a preferred solution a combination of open-source tools, such as MAAS (Metal-As-a-Service), Juju Charms, and also ManageX for perspective.

MAAS is designed to deploy quickly and easily Ubuntu configurations across multiple servers using techniques utilized for cloud platforms. But unlike cloud platforms, resource allocation on such kind of cluster occurs at the level of physical servers, not virtual environments. At the heart of MAAS, there is the simple idea of Preboot eXecution Environment (PXE) booting and a tool for deploying and maintaining Juju environments, which turns the installation and configuration process into an extremely simple task, performed by using two or three commands [5]. It usually takes too much time to configure manually the OS and services on each server node in the cluster, whereas tools like MAAS can deploy an entire cluster in just a few minutes.

Juju is an orchestrator that can be used to declaratively describe the infrastructure configuration of an application: which applications are running, on which machines, in how many copies, and how they are linked to other services. The custom code for configuring individual virtual machines with Juju is called Charm [6].

Relatively recently, there was announced ManageX configuration tool from Terawe Corporation that was studied and tested in the existing heterogeneous computing environment. This tool became popular for managing computing resources in many universities and research centers. ManageX is used to deliver on-demand, self-provisioned research workspaces for compute-intensive, data-driven, and AI-enabled projects. These environments support customized compute configurations, automated data pipelines, HPC workloads, and integrated access to cutting-edge AI models and services – all without the additional burden of infrastructure management.

Having the installation of an automation system, we get not only a gain in man-hours for deploying a ready-made production-ready infrastructure, but also flexibility in its administration and ease of scaling. For example, having a Compute Node setup script written, you can start provisioning any number of new nodes with a single command, using an existing debugged configuration. This allows for minimizing the occurrence of errors when commissioning new system components and carrying out maintenance with minimal delays or even without downtime.

The creation of new, adaptable infrastructure will allow us to eliminate many of the limiting factors of the infrastructure currently operating in our research computing infrastructure. The main advantages of the updated system are: automated configuration of the extended pool of computing resources, block-storage for creating a backup, a new network model that allows users to independently create self-service local networks with local addresses, and use the mapping of floating IP addresses. This, in turn, increases security and significantly reduces the use of limited assets like public IPs.

4 Computing Resources for Production Operation of Services and Applications

Production services and applications require predetermined computing resources. These resources are mainly statically allocated for the support of services and applications operation:

- Research data repositories platforms;
- Videoconferencing services – BigBlueButton and eduMEET integrated with LMS Moodle instances;
- Eduroam and eduGAIN federated services, including the operation of the National IdM Federation LEAF;
- Resources for DICOM Network distributed medical application;
- General content hosting services;
- Resources for Jupyter Notebook – a web-based interactive computing platform.
- Resources for development, deployment, and operation of the National EOSC node as a part of the common European EOSC federation.

It is important to provide research and educational institutions access to widely used platforms oriented to research data and publications collection, which are a convenient instrument for institutional and thematic repositories deployment. As a solution for the creation of universal data repositories, resources were allocated and instances for operation DSpace and EuroCRIS platforms were installed [7, 8].

One of the examples of use of pre-allocated resources is a multi-node distributed video-conferencing system, which provides facilities for the organization of online classes and has been operating since the beginning of the lockdowns caused by the COVID-19 pandemic. Videoconferencing system is offering access to two types of instances powered by the open-source projects, BigBlueButton [9] and eduMEET [10]. The VC system is integrated with the LMS Moodle, creating a self-sufficient distant e-learning platform, and it is now actively used for distant learning by the main universities of Moldova: the Moldova State University, the Academy of Economic Studies of Moldova, as well as by some smaller institutions in Chisinau and in the regions. It allows hosting roughly 1.5-2k of concurrent users daily, with the peaks up to nearly 3k users in about 100 separate virtual rooms distributed among the servers' clusters.

The effective use of the VC system has been achieved by uniting 9 distributed VC instances support nodes in a cluster using the Scalelite project. Scalelite is an open-source load balancer that manages a pool of VC servers. It makes the pool of VC servers appear as a single (very scalable) server. A front-end, such as Moodle or Greenlight, sends standard BigBlueButton and eduMEET API requests to the Scalelite server, which, in turn, distributes those requests to the least loaded VC server in the pool. In addition, Greenlight is used as a meeting management plugin and a pool of three Traversal Using Relay NAT (TURN) servers for relaying the traffic between peers behind the NAT.

Pre-allocated computing resources are also used for the operation of the Authorization and Authentication Infrastructure (AAI) in Moldova: National Identity Management Federation (IdM) – LEAF [11] and the national eduroam infrastructure. National IdM Federation is the main instrument for access to various services, including computing resources. LEAF IdM Federation is linking now nine institutional Identity Management Providers (IdP), six of them are realized and operated by using computing resources allocated to LEAF Federation.

Although some research and educational organizations in Moldova continue to use different Authorization and Authentication mechanisms to access various services in Moldova and internationally, the main proposed approach for providing convenient access to national and international service providers is the integration of all research and educational institutions in the National IdM Federation. LEAF is a part of AAI that is realizing eduGAIN service – a global inter-federation AAI that integrates different types of user identification mechanisms and IdM federations.

The common computing infrastructure is offering resources for the production operation of distributed applications, like the DICOM Network medical information system [12]. The allocated computing infrastructure is used for collecting anonymized medical data extracted from the application databases for data processing using various data analytics tools. Available storage resources are also used for long-term archiving of the collected medical data.

Resources of the created infrastructure are used for providing general information hosting services. This presumes allocating various hosting resources according to users' requests:

- Allocation of VMs with preinstalled hosting platform;
- Installation and operation of a number of selected Content Management Systems (WordPress, Joomla, etc.) and providing users access to these platforms.

Computing resources are used for the permanent operation of Jupyter Notebook – a web-based interactive platform for research data collection and analytics. The Notebook combines live code, equations, narrative text, and visualizations. Jupyter Notebook allows users to compile all aspects of a data project in one place, making it easier to show the entire process of a project to your intended audience. Through this web-based application, users can create data visualizations and other components of a project to share with others via the platform.

In perspective, the allocation of production resources for the creation EOSC National node in Moldova is planned. This node will be a part of a joint European federated infrastructure that unites the Central EU EOSC node, a number of thematic and selected national EOSC nodes [13]. According to the current plan, the node will provide the following possibilities to users:

- Research Data sets repositioning;
- Access to Data services;
- Data workflows for research data analysis and open-source software;

- Data processing services, including High Throughput and High-performance computing;
- Storage services;
- Virtual Research Environments (VRE);
- Collection and cataloguing of scientific applications.

5 Resources for Scientific Cloud Infrastructure Operation

In 2023, the concept of the creation of a heterogeneous computing infrastructure was elaborated, which includes the deployment of multi-zone IaaS based on OpenStack cloud middleware [14]. Works on deployment of the updated Scientific multi-zone IaaS based on the new OpenStack version are progressing now, taking into account continuation of physical computing resources upgrading by installation of new servers in all three main datacenters. As a result, today in IMCS, MSU, and RENAM, OpenStack version 2023.1 Antelope is deployed, which is currently the stable release that will be actively maintained at least for several upcoming years, offering more features, more processing power, and flexibility of operation [15].

In the current distributed cloud understructure, useful and important components – block storage and Virtual eXtensible Local Area Network (VXLAN) traffic tagging are already implemented. These tools will be used for the development of future cloud infrastructures, too. Block storage allows the creation of volumes for organizing persistent storage. The Block storage component, used in the created multi-zone IaaS, is deployed on a separate storage sub-system and allows the creation of block storage devices and mounting them on a virtual machine through special drivers over the network. This type of volume is a persistent storage that can be reused when the virtual machines are deleted. Thus, you can easily move data from one virtual machine to another, or quickly scale up VM performance by creating a virtual machine with larger resources and simply mount volumes to it with all available data for further processing.

VXLAN is an advanced and flexible model of interaction with the network. In the upgraded cloud infrastructure, in addition to the usual "provider network" model, which allocates one real IP address from the pool of provider network addresses to each virtual machine, a self-service network is also available. A self-service network allows each project to create its own local network with Internet access via NAT (Network Address Translation). When using the self-service model, the floating IP technology becomes available, which allows you to temporarily bind the IP address from the provider network to any of the virtual machines in the project, and at any time detach it and reassign it to any other virtual machine of the project. The replacement occurs seamlessly; that is, the address does not change inside the machine, but remains the same – the address is from the internal network of the project. The external IP address remains assigned to the project and can be reused by other machines within the project.

For effective operation of the upgraded computing infrastructure, a new 10G virtual networking segment was deployed that allowed to switch connection of all servers to N x 10G interfaces [16].

6 Allocation of Resources for Multiprocessor Clusters Operation

Many research teams in Moldova are developing scalable complex applications or using specialized software packages that are oriented on parallel data processing, execution in a multiprocessor cluster environment. For these users, specific multiprocessor resources of the common computing infrastructure can be allocated. These resources allow for configuring multiprocessor clusters that unite several servers and allow for deploying a computing environment that provides a significant number of compute cores and volume of high-speed memory for parallel applications development and execution.

Multiprocessor Clusters can be created as a part of a virtualized Scientific cloud infrastructure; also, resources for clusters' operation can be configured using only bare metal servers, or it can be a combination of virtualized and bare metal servers. The process of computing resource allocation and cluster configuration can be managed automatically. The management tools for heterogeneous cloud infrastructure configuration that are described in Section 6 can be applied to multiprocessor cluster configurations too.

For the simplification of multiprocessor clusters' resource utilization, it is important to have configuration tools for the adjustment of the execution framework for running complex applications as part of workflows used for the selected applications execution. Now, a new approach is being developed and realized for the collection and systematization of different types of user workflows for the deployment of libraries with annotated workflows, workspaces, and execution environment descriptions that will allow to re-use of created workflows and, in perspective, to automate the adjustment of execution frameworks according to users' needs. In perspective, the AI-based technologies can be applied for finding appropriate workflows and adjusting for the execution of a target set of applications.

7 Specialized Infrastructure for AI Applications and Machine Learning

Expansion of the use of AI technologies in research and education brings up different infrastructure needs to traditional research computing. Traditional infrastructures are changing according to the two major 'phases' of AI technologies utilization for applications development and use:

- AI Training requires massive computational resources working in perfect synchronization.
- AI Inference is when those trained models answer your questions – lighter computing, but needs to be lightning-fast and globally accessible.

Large volumes and high-quality data are an essential need for AI systems. Modern AI models require vast amounts of high-quality, diverse data to avoid biases and ensure their reliability. Universities and research institutions across the globe are generating valuable datasets, but without coordinated strategies to share this data responsibly, individual institutions face significant limitations in both scale and diversity. This is particularly problematic for AI research, research models based on AI applications, where model quality is directly tied to the breadth and depth of training data. Modern AI models require hundreds or thousands of GPUs working in perfect concert, computational demands that exceed most individual institutions and even national infrastructures possibilities. For the intensification of AI technologies' use in research and educational institutions, the following approaches are recommended for realization [17]:

- Deploy specialised AI clusters: dedicated to GPU resources for research training and inference;
- National Research and Educational Computing Infrastructure Sovereignty: reduced dependency on commercial providers;
- Federated resource sharing: pool capacity across institutions and borders during peak demand;
- Fair resource allocation systems: reservation and scheduling across resources distributed among institutions.

Following the increasing necessity to use specialized, AI-oriented computing resources, a part of the available computing infrastructure is planned to be allocated for supporting AI-based applications and models development. This pool of resources will comprise NVIDIA T4 Tensor Core GPUs integrated with modern servers that will have the ability to offer hundreds of CPU cores. The following tools for supporting AI applications are ready to be preinstalled to simplify the creation of research workspaces for data-intensive AI-enabled projects:

- TensorFlow 2 – an end-to-end open-source machine learning platform.

- Keras: Deep Learning for Humans. Keras is a high-level, deep-learning API developed by Google for implementing neural networks. It makes the implementation of neural networks easy and supports multiple backend neural network computations.

These tools and resources can be a useful facility for Natural Language Processing and Machine Learning applications that are being developed in IMCS USM.

8 Storage Infrastructure

The distributed computing infrastructure comprises dedicated storage sub-systems for a large amount of data archiving and providing resources for the whole distributed infrastructure data backup.

Systemic concepts for organizing data storage in a distributed heterogeneous computing environment were proposed and analyzed, based on the application of the multi-zone and multi-level storage concept using various platforms for organizing data storage and archiving (NextCloud, FreeNAS, other hierarchical memory management systems – including Oracle SAM-QFS, HPE Data Management Framework, Quantum StorNext). Several approaches were realized and appropriate storage management systems deployed for data collection and storing that allow realizing reliable multi-level data archiving and implementing various solutions for the organization of permanent data storage.

Elements of the storage servers' infrastructure are now distributed among three sites for reliable data backup and archiving, which allows storing copies of data at least in two different locations.

Automated back-up system is making multi-level backups that include copying of virtual machine images, regular copies of content from operating software platforms, moving datasets from operative databases to data preservation archives, etc.

9 Conclusion

Heterogeneous computing infrastructures, virtualization, and Cloud technologies give resource providers flexible tools for effective resource manipulation in small to medium computing infrastructures.

The created adaptive computing infrastructure is a successful example of adaptation of new technologies and open-source software platforms for providing computational resources to the scientific and educational community (Fig. 3).

There are perspectives to continue developing the created scientific computing infrastructure and technologies at national, regional, and international levels. The works focused on the adaptation and implementation of new open-source tools for extending the usability of the created federated computing infrastructure. A perspective direction that can enhance the development of the national scientific computing infrastructure and services is cooperation with partners within new activities and projects initiated by GEANT, European Grid Initiative Foundation, and other European e-Infrastructures. There are perspectives to be engaged in projects that comprise activities related to the integration of the national distributed infrastructure in the European Cloud and High Throughput computing infrastructures.

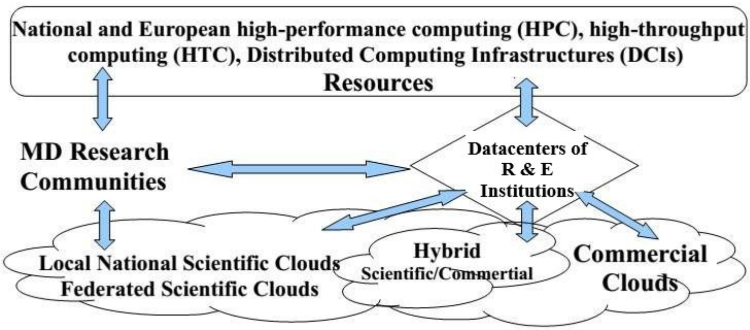


Figure 3. Heterogeneous computing resources for Research Communities in Moldova

We expect a positive outcome from the realization of the concept described above of the adaptive computing infrastructure deployment. This e-Infrastructure would provide a wide range of resources and unified services for national research and educational institutions, which obviously contribute to the support of science and education development in Moldova.

Acknowledgments. This work was supported by the "EaPConnect2" project funded by the EU (grant contract ENI/2019/407-452) and the Institutional project No 011301 SIBIA "Information Systems Based on Artificial Intelligence" (Republic of Moldova).

References

[1] G. Secrieru, P. Bogatencov, and N. Degteariov, "Extension of distributed computing infrastructure and services portfolio for research and educational activities," in *Proceedings of the 10th International Conference "Distributed Computing and Grid Technologies in Science and Education" (GRID'2023), Russia, 3-7 July 2023* (Physics of Particles and Nuclei, 2024, Vol. 55, No. 3), pp. 492–494. DOI: `10.1134/S1063779624030717`. ISSN 1063-7796.

[2] P. Bogatencov, G. Secrieru, N. Degteariov, and N. Iliuha, "Scientific computing infrastructure and services in Moldova," *Journal Physics of Particles and Nuclei Letters LNCS*, vol. 13, no. 5, pp. 685–688, 2016, Springer Link. DOI: `10.1134/S1547477116050125`. `http://link.springer.com/journal/11497/13/5/page/2`.

[3] H. Astsatryan, A. Hayrapetyan, W. Narsesian, P. Bogatencov, N. Iliuha, R. Kvatadze, N. Gaamtsemlidze, F. Florian, G. Neagu, and A. Stanciu, "Deployment of a Federated Cloud Infrastructure in the Black Sea Region," in *Computer Science and Information Technologies. Proceedings of the CSIT Conference, Sep. 23-27*, Erevan, Armenia, 2013, pp. 283–285.

[4] P. P. Bogatencov, G. V. Secrieru, B. Hîncu, and N. Degteariov, "Development of computing infrastructure for support of Open Science in Moldova," in *Workshop on Intelligent Information Systems (WIIS2021) Proceedings*, (Chisinau), IMI, 2021, pp. 34–45. ISBN 978-9975-68-415-6.

[5] MAAS documentation. [Online]. Available: https://maas.io/docs.

[6] Juju documentation. [Online]. Available: https://juju.is/docs.

[7] https://www.dspace.com/en/pub/home/support/supdoc.cfm.

[8] https://eurocris.org/what-eurocris.

[9] BigBlueButton, Open Source Virtual Classroom Software. [Online]. Available: https://bigbluebutton.org/.

[10] https://edumeet.org/.

[11] https://federations.renam.md/.

[12] A. Golubev, P. Bogatencov, and G. Secrieru, "DICOM data processing optimization in medical information systems," *Scalable Computing: Practice and Experience. Scientific International Journal for Parallel and Distributed Computing.* West University of Timisoara, pp. 189–201, 2018. ISSN: 1895-1767. DOI: https://doi.org/10.12694/scpe.v19i2.1399.

[13] EOSC Federation Handbook. Version 1, EOSC Association, 2025, Zenodo. [Online]. Available: https://doi.org/10.5281/zenodo.14999577.

[14] P. P. Bogatencov and N. P. Degteariov, "Integrated e-Infrastructure to support research and educational activities," in *Proceedings of Workshop on Intelligent Information Systems (WIIS2023)*, (Chisinau), VA IMCS, 2023, pp. 43–54. ISBN 978-9975-68-492-7.

[15] P. P. Bogatencov, G. V. Secrieru, R. Buzatu, and N. P. Degteariov, “Distributed computing infrastructure for complex applications development,” in *Proceedings of Workshop on Intelligent Information Systems (WIIS2022)*, (Chisinau), VA IMCS, 2022, pp. 55–65. ISBN 978-9975-68-461-3.

[16] P. P. Bogatencov, G. V. Secrieru, and N. P. Degteariov, “Development of effective access to the distributed scientific and educational e-infrastructure,” in *Proceedings of the 9th International Conference ”Distributed Computing and Grid Technologies in Science and Education” (GRID’2021)*, (Dubna, Russia), 5-9 July 2021. vol. 3041, urn:nbn:de:0074-3041-3, pp. 503–507. DOI: 10.54546/MLIT.2021.21.73.001. ISSN 1613-0073.

[17] GENIAL: When are GENerative AI Tools a Catalyst for Learning? [Online]. Available: https://lse-dsi.github.io/genial/.

Peter Bogatencov

ORCID: https://orcid.org/0000-0003-0664-300X
Moldova State University, “V. Andrunachievici” Institute of Mathematics and Computer Science, Chisinau, Republic of Moldova
E–mail: bogatenc@asm.md

Message segmentation vs. latency in computer networks

Ion Bolun

Abstract

The degree of influence of message segmentation on latency in wide-area computer networks, depending on various factors, is examined. Based on some queueing models and calculations performed, it is shown that the influence in question is not univocal; there are many cases in which segmentation is inopportune even at relatively large message sizes. With the decrease in the number of "node-to-node" hops on the "source-destination" path, the number of cases in which message segmentation is inopportune increases. Cases are identified where non-segmentation of messages can reduce their latency by up to 6 times or more.

Keywords: data transfer channel, latency, pachet, router, user request.

1 Introduction

At relatively high costs and high quality of the requested services, the efficient use of the resources of wide-area computer networks is of great importance. User requests to the network and the network's response to them constitute messages to be transferred. Request messages are transmitted from each user station (SU) or from the user stations of each separate local area network (LAN) (SU/LAN) to certain data centers (DC) of the network. To be transferred, messages are segmented, the resulting segments are encapsulated in data packets, the latter being encapsulated in data frames that are transferred from node to node along the path between source and destination.

In computer networks, the size of data packets is limited by the set values of the maximum transmission unit (MTU). For example, in the Internet (TCP/IP network technology), the size of IPv4 packets can be up to $2^{16}-1$ bytes (datagram packets), mandatory being the acceptance of packets of up to 576 bytes [1], and the size of IPv6 packets can be up to $2^{32}-1$ bytes (jumbogram packets), mandatory being the acceptance of packets of up to 1280 bytes [2]. However, for links that can set the MTU value (e.g., PPP links), it is recommended to use an MTU of 1500 bytes [2].

The MTU value of 1500 bytes is commonly used in networks. So, one or more packets can correspond to a message. It has also been shown that sometimes using MTUs larger than 1500 bytes can reduce the message transfer time T [3]-[6].

In [7], it is shown that 9000-bytes jumbo frame connections are more resistant to packet loss. The results of [6] show that the average latency decreases as the packet size increases. Using the 9000 bytes MTU for 40 Gbps Ethernet, the overall output results exceed the current IEC 61375-2-5 ETB standards by up to 90% [8].

Thus, for certain network situations, it is preferable to use an MTU of 9000 bytes compared to an MTU of 1500 bytes. But there are also many cases in which it is preferable to use an MTU of 1500 bytes or larger than 9000 bytes.

Of interest is the degree of influence of message segmentation on latency during its transfer through the wide-area networks, depending on various factors. Such factors include: the size of the messages to be transferred, the number of "node-node" hops on the "source-destination" path, the size of packets, the performance and loading of routers and data transfer channels (DTC), etc. Namely, such aspects are examined in this paper.

2 Factors that influence message latency in computer networks

Let's examine some factors that could influence the latency T of messages during their transfer through the network.

The message size, according to [9], [10], can usually be, depending on the requested service, from about 200 bytes to 200 KB and more. The ITU-T Recommendation G.1010 [11] specifies values from less than 1 KB (Telnet, interactive games) to 100 KB (static images), and for bulk data – up to 10 MB. It will be assumed that the size of both messages and packets is exponentially distributed with mean v and q, respectively.

According to ([12], p. 415), the rate μ_p of processing packets by a router is approximately equal to the rate μ of processing messages. Of course, for large values of the number g of packets corresponding to a message, these rates differ, involving other factors, for example, the time of recording data entities in the router buffer. Thus, in the general case, it can be considered that $\mu_p \in [\mu, g\mu]$. So, $\mu_p = g^{1-a}\mu$, where $0 \leq a \leq 1$. The value of the difference $(1 - a)$ characterizes the degree of coincidence of the rate μ with that μ_p: the coincidence in question is the smallest at $a = 0$ ($\mu_p = g\mu$) and is total at $a = 1$ ($\mu_p = \mu_r$).

With reference to the incoming request flows in wide-area computer networks, for overall calculations, they can be considered of the Poisson type [13]-[15].

The most widely used characteristic of vulnerability to errors in data transfer is the bit error rate (BER) – the average value of the ratio between the number of erroneously received bits and the total number of transmitted bits. In modern telecommunications, the BER value is about 10^{-9} [16], in wired data transfer networks, the typical BER value is 10^{-12} or less, and in high-performance wireless networks – 10^{-9} [17], [18]. At $p = 10^{-9}$, the probability of erroneous reception of a data packet of $q = 1500$ bytes is equal to $[1–(1–p)^{8q}] \approx 9,6 \times 10^{-5}$. Similarly, at $g = 20$, the probability of erroneous reception of a message of $v =$

$qg = 1500 \times 20 = 30000$ bytes is equal to $[1–(1–p)^{8\times30000}] \approx 23 \times 10^{-5}$. Both values are relatively small and, given the initial data examined, would have little influence on the latency of messages. Therefore, the vulnerability to errors of data transfer in computer networks will not be examined in this paper.

3 Latency when segmenting messages

Two models (Models 1 and 2) regarding the latency of data transfer in networks when segmenting messages, in order of increasing their accuracy, are described and discussed.

3.1 Model 1 – simplistic

Model 1 – the simplest of the two, aims to define very general conclusions on the subject under examination. It is based on comparing the transfer time of requests from users of SU/LAN x to DC z of the network in the form of messages with that in the form of data packets for a simplified case, namely:

- by one router is placed at SU/LAN x and at DC z;

- the message flow between SU/LAN x and DC z is elementary with rate λ;

- to one message correspond, on average, g packets;

- both the size of the messages and that of the packets are exponentially distributed with the mean, respectively, v and $q = v/g$;

- the service time by each of the k (odd number) components (routers and DTCs) of the network on the path between SU/LAN x and DC z, of both the messages and that of the packets, is exponentially distributed with the rate, respectively, μ and $g\mu$;

- the k components of the network serve only the requests of the SU/LAN x flow to DC z. Under these conditions, the following relationships take place:

$$T_{mes1} = k\frac{1}{\mu - \lambda} = \frac{k}{\mu - \lambda}, \tag{1}$$

where $1/(\mu–\lambda)$ is the average retention time of a request by a router or a DTC [15], and T_{mes1} is the average retention time of a request in the network on the path between SU/LAN x and DC z, in the case of its transfer in message form;

$$T_{pmes1} = \frac{k+g-1}{g(\mu-\lambda)}, \qquad (2)$$

where T_{pmes1} is the average retention time of a request in the network on the path between SU/LAN x and DC z, in the case of its transfer in the form of packets.

Eq. 2 for T_{pmes1} is obtained based on the following reasoning. The transfer by the first router of the g^{th} packet of the message starts after $(g–1)/[g(\mu–\lambda)]$ time units and reaches DC z in $k/[g(\mu–\lambda)]$ time units. So, $T_{pachg} = T_{pmes1} = (g–1)/[g(\mu–\lambda)]+k/[g(\mu–\lambda)] = (k+g–1)/[g(\mu–\lambda)]$ time units.

Based on Eqs. 1 and 2, it can be concluded that the dependencies $T_{mes1}(\mu)$ and $T_{pmes1}(\mu)$ are decreasing, but to a different extent, depend on the value of the quantity g. Respectively, the dependency $T_{pmes1}(q)$ is increasing.

So, under the defined conditions, the transfer of data in the form of messages takes T_{mes1}/T_{pmes1} times longer than their transfer in the form of packets, namely:

$$\frac{T_{mes1}}{T_{pmes1}} = \frac{kg}{(k+g-1)}. \qquad (3)$$

As expected, at $k = 1$ (one router and no DTCs) or/and $g = 1$, the ratio T_{mes1}/T_{pmes1} becomes equal to 1, i.e., $T_{mes1} = T_{pmes1}$. But at $g = k > 1$, $T_{mes1}/T_{pmes1}|_{g=k} = k^2/(2k–1) = g^2/(2g–1) > k/2 = g/2$ occurs, and, at $g = k >> 1$, $T_{mes1}/T_{pmes1}|_{g=k} \approx k/2 = g/2$ occurs. Moreover, the limit equalities $\lim_{k/g\to\infty} T_{mes1}/T_{pmes1} = g$ and $\lim_{g/k\to\infty} T_{mes1}/T_{pmes1} = k$ occur. At the same time, the ratio T_{mes1}/T_{pmes1}, except for the cases $k = 1$ or/and $g = 1$, is to the same extent increasing with respect to g and with respect to k. At $g = k > 1$, the relations $T_{mes1}/T_{pmes1} \in (g/2; g+\epsilon]$ and $T_{mes1}/T_{pmes1} \in (k/2; k+\epsilon]$ occur, where ϵ is a very small quantity.

To summarize, for $k > 1$ and $g > 1$, the relationship $T_{mes1}/T_{pmes1} \in [(\min(g/2, k/2); \min(g, k) + \epsilon]$ occurs. In real computer networks, the relationships $k >> 1$ and $g >> 1$ frequently occur.

Also, from Eq. 3, it follows that the value of the ratio T_{mes1}/T_{pmes1} is increasing, to the same extent, with respect to the number g of packets that correspond to a message and with respect to the total number k of routers and DTCs on the path between SU/LAN x and DC z. More suggestively, the character of this dependence is in the graphical form presented in Figure 1.

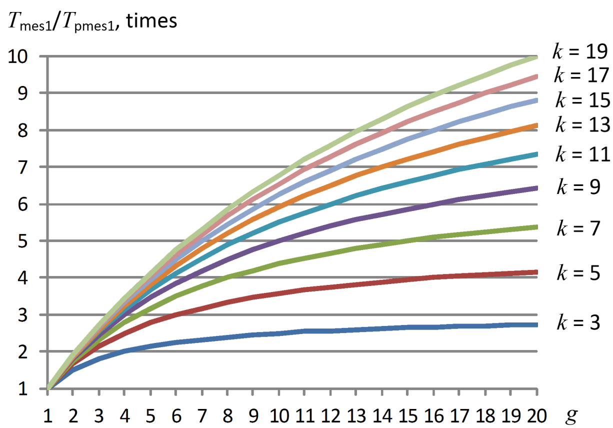


Figure 1. Ratio T_{mes1}/T_{pmes1} dependence on g and k

Thus, under the assumptions defined in Model 1, taking into account the transfer of data in the form of packets, not in the form of messages, can considerably reduce the average latency of messages in the network, namely: from $\min(g/2, k/2)$ to $\min(g, k)$ times. At the same time, the more hops (routers) there are on the "source-destination" network path, in the more entities it is appropriate to segment the messages to be transferred.

3.2 Model 2 – taking into account data processing time by routers

Model 1, although it allows the identification of important general characteristics of the subject under examination, is simplistic. For example, it is assumed in it that the packet serving rate of routers is g times higher than the message serving rate. However, as mentioned in Section 2, it is appropriate to use $\mu_p = g^{1-a}\mu$, where $0 \leq a \leq 1$. Also, the performance of routers and that of the DTCs are considered equal. But, in real networks, they can differ considerably.

Thus, Model 2, taking into account the stipulations in Sections 2 and 3, is a development of Model 1, differing from the latter by the following:

- the packet serving rate μ_p by each router is determined as $\mu_p = g^{1-a}\mu$, where $0 \leq a \leq 1$;

- the message serving rate μ_c by each CTD is determined as $\mu_c = d\mu$, and that μ_{pc} of packets is determined as $\mu_{pc} = dg\mu$.

Under these conditions, the average retention time T_{mes2} of a message in the network on the path between SU/LAN x and DC z, in the case of transferring requests in the form of messages, is equal to

$$T_{mes2} = \frac{1}{2}[\tau(k+1) + \tau_c(k-1)], \tag{4}$$

where the average retention time τ of a message by each of the $(k+1)/2$ routers is $\tau = 1/(\mu_r–\lambda)$, and that τ_c by each of the $(k–1)/2$ DTCs is $\tau_c = 1/(d\mu–\lambda)$.

In turn, the average retention time T_{pmes2} of a message in the network on the path between SU/LAN x and DC z, in the case of transferring requests in the form of packets, is determined as

$$T_{pmes2} = \frac{1}{2}\begin{cases} \tau_p(2g+k-1) + \tau_{pc}(k-1), & \text{if} \quad h \leq d \\ \tau_p(k+1) + \tau_{pc}(2g+k-3), & \text{if} \quad h > d, \end{cases} \tag{5}$$

where the average retention time τ_p of a packet by each of the $(k+1)/2$ routers is $\tau_p = 1/(\mu_p–g\lambda)$, and that τ_{pc} by each of the $(k–1)/2$ DTCs is $\tau_{pc} = 1/(d\mu – g\lambda)$.

Eq. 5 is obtained based on the following reasoning. If $h < d$, then each DTC transmits packets faster than each router processes them. The processing by the first router of the g^{th} packet of the message representing the request is completed over $g\tau_p$ time units and reaches DC z (transiting (k–1)/2 routers and (k–1)/2 DTCs) in $\tau_p(k-1)/2 + \tau_{pc}(k-1)/2$ time units. Adding these two expressions, the first line of Eq. 5 is obtained.

However, if $h > d$, then each router processes packets faster than each DTC transmits them. The first packet reaches the first DTC of the path between SU/LAN x and DC z in τ_p time units. This DTC completes the transfer of the g packets to the second router in $g\tau_{pc}$ units of time. Accordingly, the transfer of the g^{th} packet, starting with the second router and reaching CD z (transiting (k–1)/2 routers and (k–3)/2 DTCs) takes $\tau_p(k-1)/2 + \tau_{pc}(k-3)/2$ time units. Adding these three expressions, the second line of Eq. 5 is obtained.

Based on Eqs. 4 and 5, data transfer in the form of messages takes T_{mes2}/T_{pmes2} times longer than in the form of packets.

For Model 2, some calculations are performed with the following initial data: $k = \{1, 3, 5, \ldots, 19\}, g = \overline{1, 20}, \lambda = 750$ messages/sec, their derivatives $\mu = 7500$ messages/sec, $\lambda/\mu = 0.1$ (the load of each router when transferring data in form of messages), $\lambda/d\mu = 0.125$ (the load of each CTD) and different values of quantity a.

Some of the results of the calculations, at $a = 0, a = 0.25, a = 0.5, a = 0.75, a = 0.9$, and $a = 1$, are shown in Figures 2-4, respectively.

At $a = 0$ the relationships $\lambda/h\mu = \lambda/\mu = 0.100$ obviously occur, and at $a = 0.25$ the value of the ratio $\lambda/h\mu$ increases from 0.1 for $g = 1$ to approx. 0.211 for $g = 20$. The ratio $\lambda/h\mu$ represents the load of each of the $(k + 1)/2$ routers with the processing of the g packets corresponding to a message.

According to Figure 2, at $0 \leq a \leq 0.25$ and $g > 1$, the value of the ratio T_{mes2}/T_{pmes2} is increasing with respect to k. At the same time, at $a = 0$ with respect to g, the T_{mes2}/T_{pmes2} value is decreasing at $k = 1$ (starting with the value 1 at $g = 1$) and is increasing at $k \leq 3$

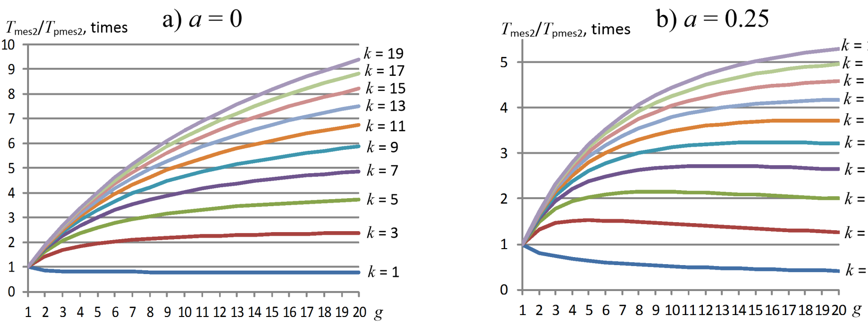


Figure 2. Ratio T_{mes2}/T_{pmes2} dependence on g and k at: a) $a = 0$, b) $b = 0.25$

(starting with the value 1 at $g = 1$). But at $0 \leq a \leq 0.25$ the rate of increase of the T_{mes2}/T_{pmes2} value decreases with increasing size a. Moreover, with increasing size g it decreases to such an extent that the ratio in question becomes decreasing; for example, for $a = 0.25$ at $k = 3$ starting with $g \approx 4$, and at $k = 5$ starting with $g = 9$. Thus, under the given conditions and $0 \leq a \leq 0.25$, segmenting messages is convenient for some cases at $k \geq 3$ and relatively small values of g, but it is not convenient at $k = 1$ and some cases at $k \geq 3$ and relatively small values of g.

At $a = 0.5$ the load $\lambda/h\mu$ of each router increases from 0.100 for $g = 1$ to approx. 0.447 for $g = 20$, and at $a = 0.75$ it increases from 0.100 for $g = 1$ to approx. 0.946 for $g = 20$.

According to Figure 3, at $0.5 \leq a \leq 0.75$ and $g > 1$, as well as at $0 \leq a \leq 0.25$ and $g > 1$, the T_{mes2}/T_{pmes2} value is increasing with respect to k. At the same time, the character of the T_{mes2}/T_{pmes2} dependence on g at $0.5 \leq a \leq 0.75$ and $g > 1$ is radically different from that at $0 \leq a \leq 0.25$ (see Fig. 2). For some values of k, the ratio

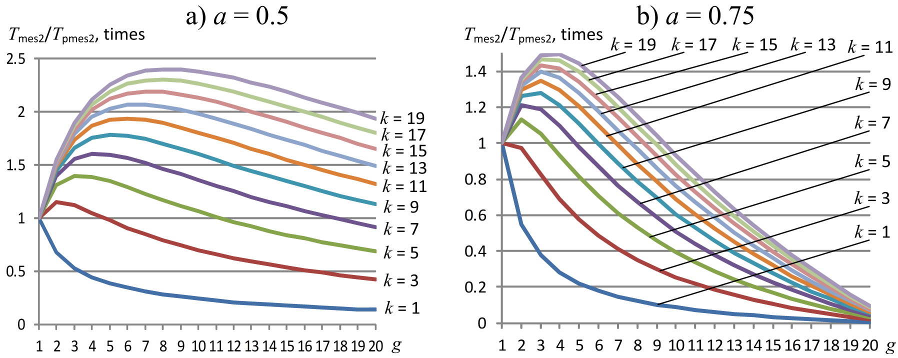


Figure 3. Ratio T_{mes2}/T_{pmes2} dependence on g and k at: a) $a = 0.5$, b) $b = 0.75$

T_{mes2}/T_{pmes2} is increasing with respect to g at small values of g, but is decreasing with respect to g at large values of g. Moreover, with increasing a, the rate of increase of the ratio T_{mes2}/T_{pmes2} decreases, but its decrease increases.

At $a = 0.9$, the load $\lambda/h\mu$ of each router increases from 0.100 for $g = 1$ to approximately 0.936 for $g = 12$ (for $g = 13$, $\lambda/h\mu > 1$ occurs), and at $a = 1$, it increases from 0.100 for $g = 1$ to 0.900 for $g = 9$ (for $g = 10$, $\lambda/h\mu = 1$ occurs).

Figure 4 shows the same trend as that of Figure 3, but stronger. For example, at $a = 1$, message segmentation may be appropiate only if $k \geq 7$, namely in 11 cases: $\{k = 7, g = 2\}$, $\{k = 9, g = 2\}$, $\{k = 11, g = 2\}$, $\{k = 13, g = 2\}$, $\{k = 13, g = 3\}$, $\{k = 15, g = 2\}$, $\{k = 15, g = 3\}$, $\{k = 17, g = 2\}$, $\{k = 17, g = 3\}$, $\{k = 19, g = 2\}$, and $\{k = 19, g = 3\}$.

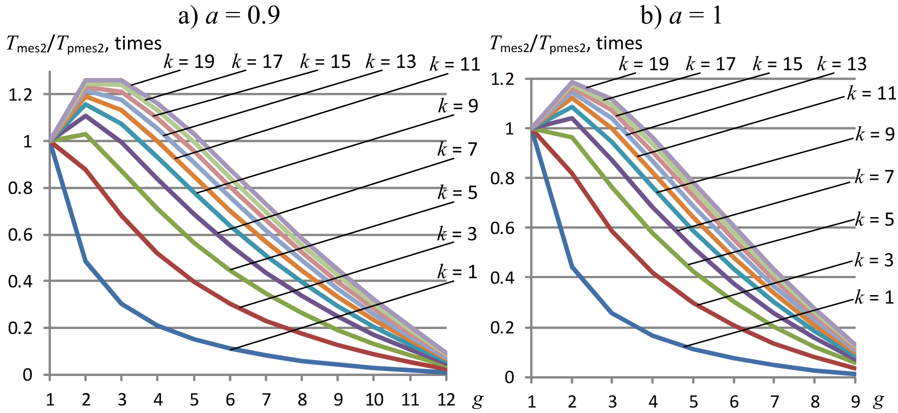


Figure 4. Ratio T_{mes2}/T_{pmes2} dependence on g and k at: a) $a = 0.9$, b) $a = 1$

4 Conclusion

The degree of influence of message segmentation on latency in wide-area computer networks is investigated. The obtained results show that the influence in question is not univocal. Although, in most cases message segmentation is appropriate, there are also cases in which it is not reasonable: in all the examined cases, message segmentation is not appropriate at $k = 1$ and in very many cases at $k = 3 (1 \leq k \leq 3)$.

Under the used values of initial data:

- the ratio T_{mes2}/T_{pmes2} is increasing with respect to k;

- as expected, with the inrease of the parameter a value $(0 \leq a \leq 1)$, the opportunity of segmenting messages is deccreasing. At large values of a, there may be many cases where message segmentation is not reasonable;

- for low values of the parameter a, the more hops (routers) there are on the network path between the source and the destination, in the more entities it is appropriate to segment the messages to be transferred;

- at $a = 0$ and $k \geq 3$, the ratio T_{mes2}/T_{pmes2} is increasing with respect to g;

- at $a \geq 0.25$ and $k \geq 3$, with respect to g the ratio T_{mes2}/T_{pmes2} is increasing at small values of g, but is decreasing at large values of g;

- at $a = 1$, message segmentation may be appropriate only if $k \geq 7$, namely in 11 cases out of the 380 in total – cases defined by the values of the pair $\{g, k\}$;

- at $a = 1$, the load of each router by packets flow increases from 0.10 for $g = 1$ to 0.90 for $g = 9$, and, at $\{g = 9, k = 19\}$, the message latency in case of data transfer in the form of messages is lower than that in the form of packets by more than 6.6 times.

References

[1] "Internet protocol," RFC 791, IETF, September 1981, 45 p. [Online]. Available: https://www.rfc-editor.org/rfc/rfc791. Accessed June 5, 2025.

[2] S. E. Deering and R. M. Hinden, "Internet Protocol, Version 6 (IPv6) Specification," RFC 8200, IETF, July 2017, 42 p. [Online]. Available: https://www.rfc-editor.org/rfc/rfc8200. Accessed June 5, 2025.

[3] D. Murray, T. Koziniec, K. Lee, and M. Dixon, "Large MTUs and internet performance," in *IEEE HPSR*, IEEE, 2012, pp. 82—87. DOI: 10.1109/HPSR.2012.6260832.

[4] S. Narayan and P. R. Lutui, "Network Performance Evaluation of Jumbo Frames on a Network," in *IEEE ICETET*, IEEE, 2013, pp. 69–72.

[5] A. Das and S. Debbarma, "Performance of Jumbo Sized Data on Jumbo Frame and Ethernet Frame Using UDP over IPv4/IPv6," in *IEEE ADCONS*, IEEE, 2013, pp. 204–207.

[6] K. Sharma and V. Badarla, "Curtailing latency in data center network by adopting Jumbo Frames," in *IEEE International Conference on Advanced Networks and Telecommunications*, 2016, pp. 1–6. DOI: 10.1109/ANTS.2016.7947785.

[7] M. Mathis, J. Semke, J. Mahdavi, and T. Ott, "The Macroscopic Behavior of the TCP Congestion Avoidance Algorithm," *SIGCOMM Comput. Commun. Rev.*, vol. 27, pp. 67—82, July 1997. DOI: http://doi.acm.org/10.1145/263932.264023.

[8] F. Ngobigha, G. Koczian, and S. D. Walker, "Towards Mobile Data Centres: Provision of End-to-End 10 Gbit/s and 40 Gbit/s Ethernet Train Backbones on International Rolling Stock," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 2, pp. 1890-1898, Feb. 2024. DOI: 10.1109/TVT.2023.3315516.

[9] I. Bolun, "Synthesis of regional computer networks (on the example of the RSS Moldova)," PhD dissertation, Kyiv: Institutul Politehnic din Kyiv, 1978, 226 p. (in Russian).

[10] I. Bolun, *Macrosynthesis of computer networks*, Chişinău: Editura ASEM, 1999, 265 p. (in Romanian).

[11] "*ITU-T Recommendation G.1010: End-user multimedia QoS categories*," ITU, 2001, 18 p.

[12] A.S. Tanenbaum, N. Feamster, and D. Wetherall, *Computer networks*, 6th ed., Pearson Education Limited, 2021, 922 p.

[13] B.I. Grigelionis, "On the accuracy of approximation of the composition of renewal processes by a Poisson process," in *Lithuanian Mathematical Collection*", vol. 2, no. 2, 1962, pp. 135–143. (in Russian).

[14] S.A. Mayorov, G.I. Novikov, T.I. Aliev, et al. *Fundamentals of the Theory of Computing Systems*, Moscow: Higher School, 1978, 408 p. (in Russian).

[15] T.G. Robertazzi, *Computer networks and systems: queuing theory and performance evaluation*, New York: Springer-Verlag, 1994, 368 p.

[16] "*What is Bit Error Rate or BER? EverythingRF*," May 25, 2023. [Online]. Available: https://www.everythingrf.com/community/what-is-bit-error-rate-or-ber:~:text=Achieving%20a%20lower%20BER%20often,ensure%20more%20reliable%20data%20transmission. Accessed June 10, 2025.

[17] "*Loss, Latency, and Speed*," DataExpedition Inc., 2025. [Online]. Available: https://www.dataexpedition.com/support/notes/tn0021.html. Accessed June 24, 2025.

[18] "*Bit error rate, PathFinderDigital*," 2025. [Online]. Available: https://www.pathfinderdigital.com/bit-error-rate-ber. Accessed May 5, 2025.

Ion Bolun

ORCID: https://orcid.org/0000-0003-1961-7310
Technical University of Moldova
E–mail: ion.bolun@isa.utm.md

Proceedings of Workshop on Intelligent Information Systems
WIIS2025, October 16-18, 2025, Chisinau, Republic of Moldova

On employing historical fonts generation for PostOCR correction

Tudor Bumbu, Liudmila Burtseva,
Svetlana Cojocaru, Alexandru Colesnicov,
Ludmila Malahov

Abstract

Processing texts from distant historical periods, especially those handwritten in languages with low computational resources, presents significant challenges. Even if modern methods make it possible to achieve, after laborious machine learning procedures, a fairly good rate of correct character recognition, the problem of the correctness of the resulting editable text remains a topical one. This paper presents an approach that contributes to the automation of the PostOCR proofreading process based on the presentation of digitized text using historical fonts, similar to those in the original document.

Keywords: historical fonts, OCR, PostOCR.

> Rarity gives things value. What is common is cheap. What is unique is priceless.
>
> Baltasar Gracian, The Art of Worldly Wisdom (1647)

1 Introduction

Over the years, the interest in the literary-historical heritage is becoming more and more accentuated, and the growing avalanche of scientific works in the field of old documents digitization is a further argument in favor of the concern for revitalizing the heritage, thus ensuring its accessibility and contributing to the preservation of the original document by widely distributing its digital copy. In the works, realized by the authors of this article, different historical periods of writing Romanian texts in Cyrillic script have been covered. We moved from the closer (20th century) to the earlier, descending into the depths of the ages, now being at the processing of handwritten documents in the Middle Ages.

When we are dealing with well-preserved documents, behind which there are also rich digital linguistic resources from the historical period in question (ground truth), the problem of obtaining an equivalent editable product and, at the user's wish, transliterated into modern Latin spelling, possibly with updated spelling and vocabulary, can be largely considered as solved.

The challenges arise when working with really old and rare prints or manuscripts, where we are confronted with a much poorer quality of the support on which the text was written or printed, including traces left by the weather or poor conditions of preservation (stains, tears, notes or comments of some readers), transcription errors in the case of manuscripts, lack of standardized spelling rules, writing without intervals or other marks of delimitation (scripta continua). In addition to all this, there was also a large variety of fonts used, for which there were also no standard norms, as we pointed out in [1], illustrating it with the example in Fig. 1 and proposing, at the same time, a method of font classification.

This variety of alphabets and fonts makes the problem of digitizing ancient texts even more complicated. Taking all these aspects into account, it is natural to follow in the processing of medieval Romanian texts, handwritten or printed in the Cyrillic characters of the corresponding era, some approaches specific to the processing of low-resourced languages.

Broadly speaking, the process of digitizing an old document can be operated through seven main steps, which are implemented in the HeDy

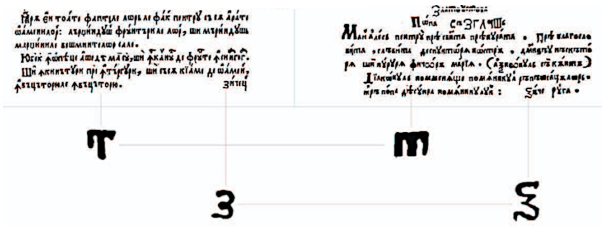


Figure 1. Different presentations of letters in the medieval alphabet

platform [2]. One of these steps is the verification of the recognized text, for which the platform offers certain facilities, mainly based on manual intervention by the user.

On the other hand, there is a natural tendency to introduce elements of automation at this step as well, for example, in the identification of post-processing errors. In this paper, we will examine some of the approaches and experiments undertaken for this purpose, specifically based on creating editable text in fonts similar to those of the original image.

2 Historical fonts: problem complexity

Providing appropriate fonts is a necessary part of the historical document digitization, being the essence of the process of restoration of ancient texts and data augmentation for OCR/HTR. Whatever the intended use of digital content, the visualization of the resulting digital content has to be as close as possible to the image of the source document.

The first experiment to verify the correctness of page recognition by comparing the original page image and the recognized text as an image was a kind of naive approach to evaluate the quality of the result using a well-preserved text from the 20th century. Fig. 2 represents a fragment of the original text (Romanian language with Cyrillic script, 20th century) and its equivalent generated with ABBY Fine Reader.


КОМПОНЕНЦА АТЛАСУЛУЙ ЛИНГВИСТИК МОЛДОВЕНЕСК, ВОЛУМУЛ I, ПАРТЯ II

КОМПОНЕНЦА АТЛАСУЛУЙ ЛИНГВИСТИК МОЛДОВЕНЕСК, ВОЛУМУЛ I, ПАРТЯ II

Figure 2. A fragment of XX sec.text: original(left) and OCRed(right)

Evidently, fonts differ in size, so the result is predictably low: 57.5% of a similarity, with different letter sizes. After manually adjusting the dimensions (height and width) of the recognized image, the similarity result increased to 99.75% at a resolution of 300 pt.

Obviously, that manual adjustment may not be a suitable solution in the case of historical fonts, which are very different by their nature. Therefore, it is necessary to find a solution for the automatic generation of fonts that are maximally similar to those in the original text.

3 Automating the generation of historical fonts

Automatic font generation is the creation of a new font library based on the reference document images. The main challenge today is to find a good balance between the automation of the process and the quality of the resulting font. The quality measures vary depending on the problem that needs to be solved. The designers require the capturing of font family characteristics, serifs, stroke shapes, etc. Solving our work tasks requires only the maximum proximity of letters between the original image and the generated digital document.

Font generation is implemented in two stages: creation of letters templates and library generation.

Font library generation

Font library generation is still challenging for a large alphabet with complex letters. For European alphabets, which contain a comparably small number of letters, the generation process today is well developed, offering multiple online tools and software like FontLab, FontForge, Calligraphr. For the presented research, we chose FontForge – a free and open-source font editor. FontForge is easy to install on every operating system (OS), being included in standard stores and repositories. It has standalone applications and APIs for the most popular programming languages. FontForge's main function is to input the letter template, to create the letter, and to place the obtained letter in the corresponding position in the Unicode table of the new font. The letter template can be a ready vector or a raster image (BMP or PNG). The vector letter template is placed directly at the cell, specified by Unicode value (Python example: `glyph = font.createChar(unicodeChar, charName)`). The image letter template initially has to be traced by applying the *autotrace* function. Embedded FontForge tracing is well-trained on European alphabets, but results strongly depend on template image quality. Our tests show that FontForge can trace a clear and strongly binary image with very high precision, even for complex letters. An example of its application is illustrated in Fig. 3.

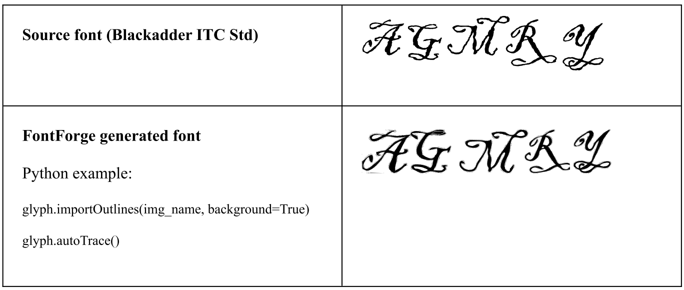


Figure 3. Fancy Font Example

Letters templates creation

Automatically creating a letter template is now the main challenge of image-based font generation.

The researchers are employing the language models, which achieve significant results in solving recognition problems. But solving the image-based font generation problem, language models stumble upon the Stroop effect, tending rather to read the text than to analyze the font properties. In the paper [3](March 2025, USA), the researchers tried to avoid the Stroop effect, evaluating different inference strategies, including CoT prompting, MCQ, and few-shot learning. Based on the tests conducted in this paper, the authors stated that the Stroop effect arose due to the modern implementation of attention mechanisms. In their opinion, when solving the recognition problem, the attention mechanism focuses on the edge information of the characters, which pushes the font details into the background. In continuation of our work, we intend to contribute to overcoming this issue and to use the potential of well-trained language models recognition.

Aside from the research, a modern practical solution to the font generation problem involves the creation of letter templates using conventional text recognition software. The main feature of creating letter templates using standard OCR is the need to catch segmentation at the letter level. The APIs of the most popular OCR systems have the corresponding functions, and the generative machine can answer the prompt: ”*Get the coordinate of the text block*”. The problem arises that the researchers are usually restricted by free tools. The majority of OCR tools have access to character segmentation only in commercial versions. Also, the prompt: ”*Get the coordinate of all characters in this text block*” is marked as out of free license.

Experiment with the use of FineReader Engine

The second study was performed using the commercial API FineReader Engine 12 (the current version on which all currently used FineReader versions are based). A program (about 200 lines of code) was developed in Python using the *comtypes* and *win32com* interface libraries to cut a handwritten text page image into separate letters. An example of a handwritten page taken from [4] is shown in Fig. 4.

To test the efficiency of this method, Russian was set as the recognition language, since the goal was to debug the interaction with FineReader Engine through its API in Python. Using FontForge, a font was generated, and the original image of the page was compared with the reproduced one, which gave a match of 85%. FineReader En-

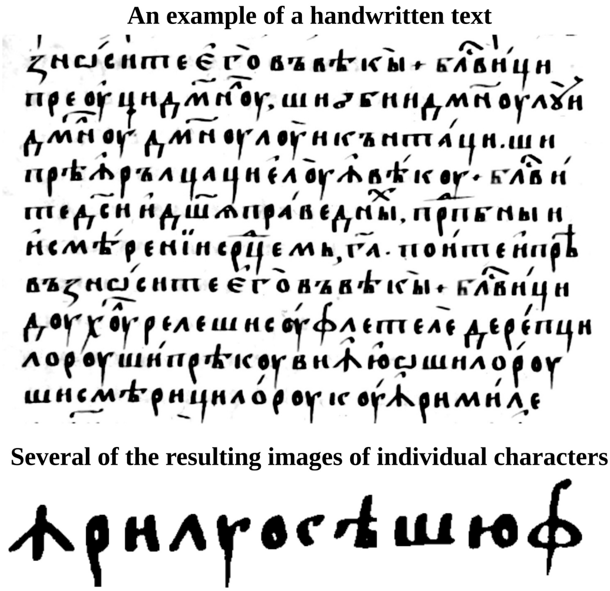


Figure 4. Segmentation by FineReader Engine

gine has started training of uncial Romanian fonts on the corresponding manuscripts of the 13th-15th centuries.

Perspectives for using the TesseractOCR

In recent (2025) historical documents digitization works, the combined approach that uses both traditional OCR and language models is being researched. Among these works, USA [5], Norway [6], England, and Germany [7] selected TesseractOCR as a traditional OCR tools. Such a choice demonstrates how the free tools requirement inspired the modern revival of TesseractOCR. Despite TesseractOCR's outdated architecture, this software has most of the features demanded by researchers: it is well trained on all rich-resources languages; it works both online and on every operating system; it has an API for the most popular programming languages. For example, in Fig. 3, character segmentation is done by the function `pytesseract.image_to_boxes`.

In the case of the historical Romanian alphabets, TesseractOCR training is not time-consuming, since these alphabets differ from the usual Cyrillic alphabet by only a few letters. The TesseractOCR repository has a fine-tuning script that easily adds several characters to the existing model.

So, in addition to our tests of FineReader Engine, we tested TesseractOCR for solving our problem.

Summarizing

Since font generation is a subtask of digitizing historical documents, preprocessing and training can be considered as having been completed. At the current stage of our work, this means that the task of creating an image template is reduced to slicing boxes and then automatically selecting the best template for each letter. In the framework of the presented research, the selection is done manually, but it can be implemented using language models, since the generation of ground truth templates is achieved by simply typing the appropriate alphabet in italics.

4 Conclusions and Further Works

The final goal of the presented research is to extend the HeDy historical document digitizing pipeline by automatically generating a font library based on the current document.

We selected the free TesseractOCR as a backend tool for characters segmentation, due to a simple fine-tuning procedure that allows us to create a model for historical Romanian alphabets.

Also, in continuation of the presented research, we intend to contribute to overcoming the Stroop effect, which appears during font recognition by well-trained language models, by developing adapters for attention mechanisms.

Acknowledgments. The project SIBIA 011301 “Information systems based on Artificial Intelligence” has supported part of the research for this paper.

References

[1] T. Bumbu, “Font Classification Model for Romanian Cyrillic Documents,” *Computer Science Journal of Moldova*, vol. 29, no. 3(87), pp. 291–298, 2021.

[2] T. Bumbu, L. Burtseva, S. Cojocaru, A. Colesnicov, and L. Malahov, *Digitization of Romanian Historical Printings*, Chisinau: Vladimir Andrunachievici Institute of Mathematics and Computer Science, Moldova State University, 2023, 176 p. ISBN: 978-9975-68-497-2.

[3] Z. Li, G. Song, Y. Cai, Z. Xiong, J. Yuan, and Y. Wang, “Texture or Semantics? Vision-Language Models Get Lost in Font Recognition,” *ArXiv Preprint*, ArXiv:2503.23768, 2025.

[4] “Psaltirea Voroneteana (I): (mss. 693 dela Bibl. Acad. Rom.)”. [Online]. Available: `https://dspace.bcu-iasi.ro/handle/123456789/22359`, Accessed July 30, 2025.

[5] S. Backer and L. Hyman, “Bootstrapping AI: Interdisciplinary Approaches to Assessing OCR Quality in English-Language Historical Documents,” in *Proceedings Of The 5th International Conference On Natural Language Processing For Digital Humanities*, 2025, pp. 251–256. [Online]. Available: `https://aclanthology.org/2025.nlp4dh-1.21/`.

[6] T. Enstad, T. Trosterud, M. Røsok, Y. Beyer, and M. Roald, ”Comparative analysis of optical character recognition methods for Sámi texts from the National Library of Norway,” in *Proceedings Of The Joint 25th Nordic Conference On Computational Linguistics And 11th Baltic Conference On Human Language Technologies (NoDaLiDa/Baltic-HLT 2025)*, 2025, pp. 98–108. [Online]. Available: `https://aclanthology.org/2025.nodalida-1.11/`.

[7] G. Greif, N. Griesshaber, and R. Greif, "Multimodal LLMs for OCR, OCR Post-Correction, and Named Entity Recognition in Historical Documents," *ArXiv Preprint*, ArXiv:2504.00414, 2025.

Tudor Bumbu [1,2], Lyudmila Burtseva[1,3] Svetlana Cojocaru[1,4],
Alexandru Colesnicov[1,5], Ludmila Malahov[1,6],

[1] Moldova State University, "V. Andrunachievici" Institute of Mathematics and Computer Science, Chisinau, Republic of Moldova
[2]ORCID: `https://orcid.org/0000-0001-5311-4464`
E–mail: `bumbutudor10@gmail.com`

[3]ORCID: `https://orcid.org/0000-0002-9064-2538`
E–mail: `luburtseva@gmail.com`

[4]ORCID: `https://orcid.org/0009-0003-1025-5306`
E–mail: `svetlana.cojocaru@math.md`

[5]ORCID: `https://orcid.org/0000-0002-4383-3753`
E–mail: `acolesnicov@gmx.com`

[6]ORCID: `https://orcid.org/0000-0001-9846-0299`
E–mail: `ludmila.malahov@math.md`

Proceedings of Workshop on Intelligent Information Systems
WIIS2025, October 16-18, 2025, Chisinau, Republic of Moldova

An approach to AI visualization of the problems from OHA

Olesea Caftanatov, Marius Cerescu
Valentina Demidova, Tatiana Verlan

Abstract

This paper presents an approach to enhancing the pedagogical value of digitized historical materials, specifically focusing on the Old Handwritten Archive of school math problems of teacher and scientist B. Cinic (OHA'Cinic). While the initial digitization of these materials into static formats like LaTeX preserves their content, it fails to capture the dynamic and intuitive nature of the mathematical concepts they represent. To bridge this gap, our approach involves transforming these static problems into interactive web-based visualizations. We detail a hybrid, human-AI collaborative workflow developed to achieve this transformation efficiently. This paper analyzes our approach, highlighting its significant advantages in rapid prototyping and its inherent challenges, such as the need for constant human supervision to correct the AI's obscure and subtle mathematical and logical errors. We conclude that this human-AI partnership is a powerful and effective model for creating rich, interactive educational tools from historical archives, augmenting the original material for a modern pedagogical context.

Keywords: digitization, human-AI collaboration, AI tools for visualization, web application.

1 Introduction

The digitization of historical documents is critically important, not just for preserving the past, but for shaping how future generations access, interpret, and engage with history. From literary texts [1] to musical scores [2], there were developed diverse specialized tools and methodologies to tackle the unique challenges of the respective domains. Within this field, the digitization of historical mathematical texts presents a particularly complex problem. These documents are not merely text; they are a dense combination of symbolic language, complex formulas, and geometric figures, all of which are handwritten and non-structured [3]. This poses unique challenges, including:

- *non-standard symbols* that lack standardized digital equivalents, requiring expert interpretation because historical texts often use author-specific notations. For instance, Carl Friedrich Gauss used the symbol "$\equiv$" to denote congruence in modular arithmetic, which he introduced in his 1801 book Disquisitiones Arithmeticae. In Bernoulli's manuscripts, "$\surd$" might mean a square root or a decorative divider [4].
- *preserving spatial relationships*, such as fractions, matrices, and geometric diagrams, relies on precise layout, which flat OCR often destroys.
- *handwritten ambiguities* that appear because of cursive script, which blurs distinctions between variables, operators, and annotations.
- *diverse nature of the content* recorded on the original punch cards; it interleaves proofs, figures, hints, and text in nonlinear arrangements.
- *machine-readability gaps*: converting handwritten formulas to computable formats (LaTeX/MathML) remains error-prone without manual intervention.

The primary challenge in digitizing OHA'Cinic lies in the heterogeneous, unstructured nature of the content, written in Cyrillic script with Latin denotations, which demands hybrid solutions that combine AI-trained symbol recognition, IIIF-enabled image annotation, and a scientific approach to preserve the mathematical meaning embedded in these materials. More about these challenges in the digitization of Cyrilic math manuscripts we presented in [3].

This paper is structured as follows. Section 2 outlines the research background and motivation, highlighting the significance of digitizing OHA'Cinic and the importance of preserving his mathematical legacy. Section 3 details our collaborative development workflow. Section 4 presents several case studies that demonstrate the application of our approach to different types of mathematical problems. Section 5 offers a critical analysis of the human-AI collaboration, examining its benefits and limitations. Finally, Section 6 concludes the paper and proposes future directions for this line of work.

2 Research Background and Motivation

Our work concerns the digitization of the large handwritten archive of math problems created by the Moldovan professor and scientist Boris Cinic, in order to preserve his legacy. He was renowned for his contributions to mathematical education. During his career, B. Cinic developed hundreds of original problems for his lyceum students, meticulously recording some of them on punch cards to serve as individualized assignments. His problems gained recognition, appearing in specialized mathematical journals and being used in Olympiad competitions not only inside the Republic of Moldova but also outside the country.

As detailed in our previous work [3], we successfully digitized a significant portion of Boris Cinic's archive, transforming the handwritten Cyrillic text and mathematical formulas into a structured format using a combination of OCR, AI-powered formula and text recognition tools (such as MathPix and GPT), and meticulous manual verification and correction in LaTeX.

Building on this effort, and in commemoration of what would have been Cinic's 90th birthday, we introduce the Virtual GPT Assistant – an AI-powered system based on OpenAI's GPT-4 – designed to emulate his distinctive teaching style [5]. Rooted in the foundational principles of Intelligent Tutoring Systems (ITSs), the assistant incorporates B. Cinic's pedagogical strategies, communication patterns, and methods derived from his original handwritten problems, preserved on digitized punch cards.

However, this successful act of preservation raises a critical pedagogical question: while static formats such as LaTeX-generated PDF or even conversational interactions with the Virtual GPT Assistant can retain the content, do they truly capture and convey its full educational potential in the context of 21st-century learning environments?

The traditional, static format encourages a passive learning model where a student reads and absorbs information. In contrast, modern pedagogy increasingly emphasizes **inquiry-based and active learning**, where students construct knowledge by exploring, experimenting, and receiving immediate feedback. According to Freeman et. al. [6], students in active learning environments score on average 6 percentage points higher and are 1.5× less likely to fail compared to traditional lecture-based STEM courses.

The static nature of the digitized archive misses the opportunity to involve the interactive potential of digital media to promote a deeper and more intuitive understanding of mathematics.

The main (or core) challenge, therefore, lies not merely in preservation but in **pedagogical transformation**: how can we efficiently bridge the gap between a static historical archive and a dynamic, modern educational tool? Traditionally, developing custom interactive educational software is a resource-intensive process, requiring collaboration among subject-matter experts, educators, and software developers. This complexity often acts as a barrier, preventing valuable historical content from being adapted into effective learning materials.

This paper presents our approach to addressing the challenge of transforming static mathematical archives into dynamic educational tools. We propose and document **a hybrid, human-AI collaborative workflow** for the rapid development of interactive mathematical visualizations. Our methodology involves the capabilities of a state-of-the-art large language model (LLM), Gemini 2.5 Pro, as a co-developer, under the guidance and validation of a human expert. We illustrate this approach through the development of a web-based application [7] featuring 13 interactive visualizations based on selected problems from the OHA'Cinic archive. This work aims not only to present the resulting application but, more importantly, to analyze the methodology itself, highlighting both the significant efficiencies and inherent limitations of using AI as a partner in the creation of specialized educational software.

As part of the evaluation, we introduced this approach during a Pi Workshop dedicated to International Pi Day, which included 45 pupils from *Da Vinci Private Lyceum* and *Theoretical Lyceum Gaudeamus*. During this event, students explored AI-generated visualizations of historical methods for demonstrating the value of π. Feedback collected from participants indicated a high level of engagement and approval, with 85% of Da Vinci students and 80% of Gaudeamus students expressing a preference for the visualization-based approach to understanding mathematical concepts. Motivated by these findings, we applied this methodology to the OHA'Cinic archive. The outcomes of this application are presented in the sections that follow.

3 Methodology

The primary objective of our work was not simply obtaining the static digitization, but to develop a methodology for creating dynamic, interactive learning tools from the source material of the OHA'Cinic archive. To achieve this, we formulated a development methodology centered on a hybrid collaboration between a human software developer and a large language model (LLM), specifically Gemini 2.5 Pro. This section presents in detail the principles of this methodology and the technological framework within which it was executed.

Our methodology is structured as an iterative workflow, which we call the **Prompt-Generate-Validate** cycle. The cycle's initiation and oversight remain firmly under the control of the human expert, who is responsible for ensuring the final product is both mathematically correct and pedagogically sound. The process begins with the human developer analyzing a given mathematical problem and formulating a precise set of instructions, or a prompt, for the AI. For instance, in developing the visualization for Card T48, the initial prompt was:

"Create a React component using SVG that displays a 2D Cartesian

plane. The component should render two vectors, a and b, originating from the center. These vectors should be interactive, allowing the user to change their coordinates by dragging their endpoints with the mouse. In a separate panel, display the real-time calculations for the dot product of a and b, their respective magnitudes, and the cosine of the angle between them."

Upon receiving such a prompt, the LLM proceeds to the generation phase. In this stage, the AI acts as an accelerated code generator, producing a functional software component based on the provided specifications. This includes generating the necessary React structure with state management hooks, SVG markup for rendering, and JavaScript functions for interactivity and mathematical calculations.

The final and most critical stage of the cycle is validation, which is performed exclusively by the human expert. The generated code is rigorously assessed against several criteria: functional correctness, mathematical accuracy, pedagogical utility, and adherence to project-wide coding standards. If the output is insufficient, the cycle repeats. The developer provides specific, corrective feedback to the AI. A notable example of this was during the development of the visualization for Card T46, where an early AI-generated solution was mathematically unstable. The corrective prompt was:

"The current implementation using trigonometric formulas is unstable. Abandon this method. Instead, implement a new strategy: for each cevian, iteratively scan the opposite side with 100 discrete points. For each point, calculate the resulting angle and find the point that minimizes the difference from the target angle. Use this point as the segment's foot."

This iterative improvement, combining human strategic oversight with the AI's rapid implementation capabilities, forms the core of our approach. It establishes a clear division of roles where the human developer serves as the architect and final arbiter of quality, while the AI functions as a powerful tool for execution and rapid prototyping.

To implement this workflow, we selected a specific set of modern web technologies. The core application was built using the **React** framework with **Vite** as the development environment. Navigation between the different visualizations was handled by **React Router**. For the visualizations themselves, a hybrid solution was used. Simpler function plots and diagrams were rendered declaratively using **SVG** directly within **React** components. For more complex and dynamic geometric constructions, which require a continuous drawing loop, we utilized the **p5.js** library. To ensure its stable integration within the component-based structure of **React**, we developed a custom **P5Canvas** wrapper component to manage the lifecycle of the **p5.js** sketches. Finally, to respect the bilingual context of the source material, the entire platform was designed for both Romanian and Russian languages using the **i18next** framework.

4 Case Studies of AI-Driven Visualizations

To demonstrate the practical application and evaluate the effectiveness of our human-AI collaborative workflow, we applied it to create thirteen interactive visualizations for a carefully selected set of problems from the OHA'Cinic archive. The picked out problems span various mathematical fields, including geometry, algebra, and number theory. This section presents a detailed analysis of three representative case studies, each chosen to highlight a distinct facet of our development process and its outcomes.

Case 1. The first case, concerning *Card T48*, exemplifies an ideal scenario where the AI, given a well-defined problem, generated a correct and functional visualization with minimal human intervention (see Fig. 1). The task was to create an interactive tool to calculate the cosine of the angle between two vectors. The prompt requested a React component with an interactive SVG plane where vectors could be manipulated by dragging their endpoints, as well as a panel for real-time calculation displays. The LLM produced a nearly perfect, self-contained component in a single iteration. It correctly implemented useState for state management, mouse event handlers for interactivity, and the required mathematical formulas. In this instance, our method proved exceptionally efficient, with the AI performing as a competent programmer, translating a clear specification into a functional prototype. The human expert's role was primarily limited to minor stylistic refinement and final integration.

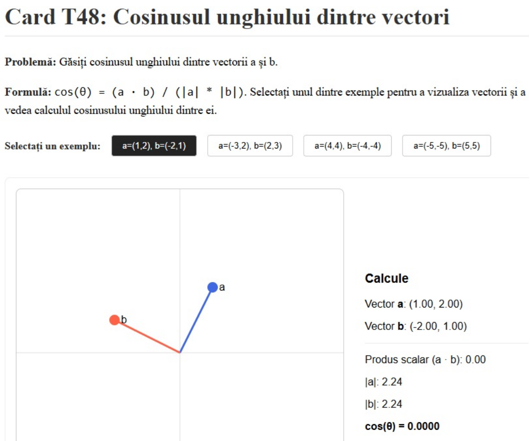


Figure 1. Visualization of the problem in Card T48

Case 2. A far more complex challenge was presented by *Card T46*, which required a visualization for a problem involving the properties of cevians in a triangle (see Fig. 2). This case study is crucial as it illustrates the limitations of AI and underscores the indispensable role of human expertise in navigating complex geometric problems. The initial prompt asked the AI to create a visualization where three specific angles involving the cevians remained equal as a slider was moved. The AI's first attempt was a trigonometric formula that, while elegant, was mathematically incorrect and produced an unstable visualization. A second attempt, based on constructing arcs of circles, was theoretically sound but the p5.js implementation was flawed, failing under different triangle orientations. At this stage, human intervention became critically important. Recognizing the fragility of the analytical approaches, the developer formulated a simpler, more robust iterative strategy. The subsequent prompt was *"Instead, implement a new strategy: for each cevian, iteratively scan the opposite side with 100 discrete* *points. For each point, calculate the resulting angle and find the point that minimizes the error."* Given this clear, unambiguous algorithm, the AI was able to generate a clean and perfectly functional p5.js implementation. This case highlights a key finding: the AI succeeded in implementing a well-defined algorithm, but failed at the strategic task of designing an algorithm for an open-ended geometric problem, where its "hallucinations" led to unworkable solutions.

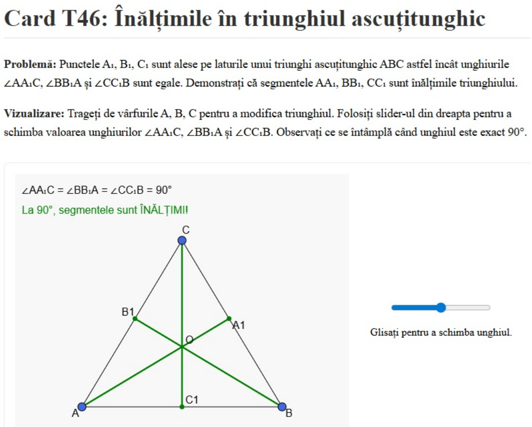


Figure 2. Visualization of the problem in Card T46

Case 3. The third case study, for *Card T55*, demonstrates the versatility of our approach in creating visualizations that are not geometric but procedural and logical (see Fig. 3). The problem required finding a number based on a specific arithmetic property. The developer first performed the necessary algebraic reduction, transforming the word problem into the equation $9x + n = 1996$. The prompt given to the AI was not for a graph, but for an "interactive solver" that would allow a user to test possible values for the last digit, n, from 0 to 9. The request specified that the interface should have buttons for each value and, upon interaction, display the resulting calculation for x with a clear visual indicator of whether the solution was an integer.

The AI successfully generated a React component that implemented

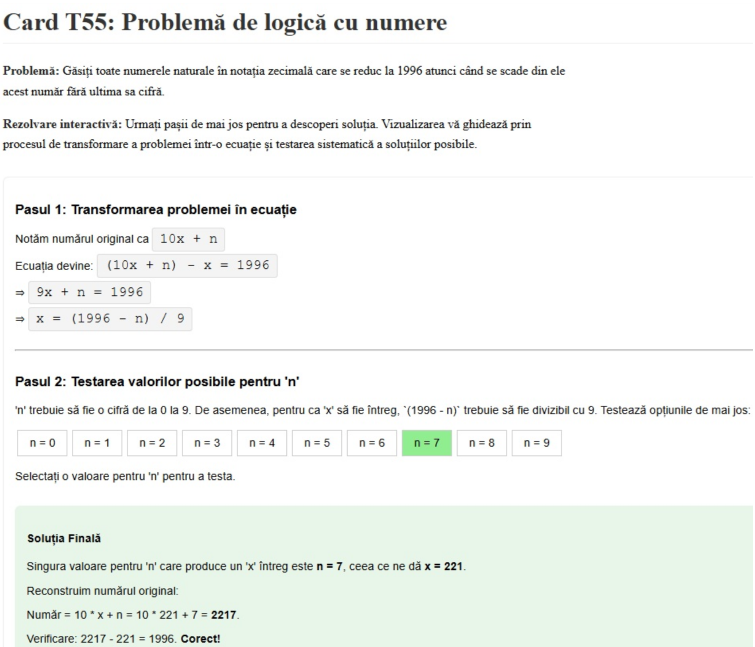


Figure 3. Visualization of the problem in Card T55

this logic, complete with interactive buttons and conditional rendering for immediate user feedback. This case shows that our human-AI workflow is not limited to geometric figures. The human expert provided the important mathematical abstraction, while the AI succeeded in translating this abstract problem-solving process into an intuitive and effective user interface. The final visualization does more than simply present the answer; it guides the user through the systematic process of elimination, thereby teaching the problem-solving method itself.

5 Discussion

The application of human-AI collaborative workflow to the problems in the OHA'Cinic archive has provided a deeper understanding of both the potential and the limitations of this development paradigm. The case studies presented in the previous section reveal a delicate relationship between the human developer and the large language model, where the success of the final product depends on a dynamic and critical partnership. This section evaluates the effectiveness of our approach and discusses its broader implications for the development of educational technology.

A primary advantage of our methodology is the significant acceleration of the development process. For well-defined tasks, such as the vector visualization in *Card T48*, the LLM was able to produce a functional prototype in minutes. This rapid generation of pattern and component-level code freed the human developer from hours of routine implementation, allowing them to focus on higher-level architectural decisions, pedagogical design, and the more complex challenges that required human ingenuity. The AI's vast knowledge of library-specific syntax and application programming interfaces (APIs) further simplified this process, effectively serving as an on-demand technical reference.

However, the case of *Card T46* exposed the critical limitations of the AI's capabilities, particularly in domains requiring deep, intuitive understanding. The model's tendency to produce "mathematical hallucinations" – solutions that are syntactically correct and seem plausible but are actually wrong – was a serious problem. Its initial attempts to solve the complex geometric constraints of the problem were elegant but incorrect. It was unable to "reason" about the physical and geometric stability of its own implementation, a task that required human experience to diagnose. The successful resolution of this problem was only achieved after the human developer abandoned the AI's analytical path and formulated a simpler, more robust iterative algorithm. This demonstrates that while the AI is a powerful tool for implementation, it currently lacks the strategic, problem-solving intuition necessary for complex, open-ended tasks. Our approach, therefore, is not one of full automation but of **augmented development**, where the human's role shifts from a line-by-line coder to that of a strategic director, validator, and creative problem-solver.

This workflow also produced an unanticipated benefit: the verification and enhancement of the source material itself. As illustrated by the analysis of *Card T.Fig.01* (see Fig. 4), the process of attempting to build a visualization forced a critical re-examination of the original problem statement.

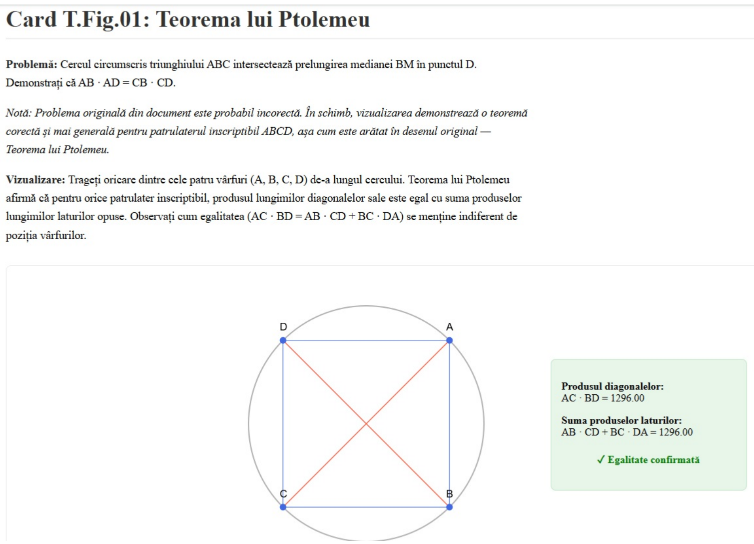


Figure 4. Visualization of the problem in Card T.Fig.01

This led to the discovery of an inconsistency between the written problem and its accompanying figure. Consequently, the final interactive tool did not just replicate the original task but corrected it, demonstrating a more general and mathematically sound theorem. This suggests that the rigorous process of building interactive models can serve as a powerful tool for the validation and scholarly annotation of digitized historical archives.

Ultimately, our approach redefines the role of the software developer in the context of creating specialized educational tools. The development process becomes less about the mechanics of writing code and more about the art of formulating precise questions, critically evaluating generated solutions, and providing the strategic guidance that current AI models lack. This collaborative model proved to be a highly effective method for rapidly producing the high-quality, interactive visualizations that constitute our platform, successfully bridging the gap between a static historical archive and a dynamic learning environment.

6 Conclusion and Future Work

This paper introduced and evaluated an approach for the visualization of a digitized, handwritten archive of mathematical problems, centered on a human-AI collaborative workflow. Our methodology, which uses the capabilities of a large language model under the strict guidance and validation of a human expert, proved to be a highly effective and efficient means of transforming static, historical documents into dynamic, interactive educational tools. By applying this approach, we successfully developed a web application with thirteen distinct visualizations that not only preserve the content of the OHA'Cinic archive but also enhance its pedagogical value for a modern audience.

Our work demonstrates that the role of artificial intelligence in such projects extends beyond simple text or formula recognition. When used as a partner in a structured development cycle, an LLM can significantly accelerate the creation of complex, specialized software. However, our findings also underscore the indispensable role of human oversight. The case studies revealed the AI's limitations in strategic problem-solving and its tendency to engage in "mathematical hallucinations," supporting the view that the current generation of AI serves best as a powerful tool for augmenting, rather than replacing, human expertise. The developer's function is transformed into that of a manager, a leader, and a critical validator, guiding the process toward a correct and meaningful outcome. This partnership model represents a powerful paradigm for the rapid prototyping and development of personalized educational technology.

Looking ahead, there are several directions for future work. The most immediate step is to expand the platform to include a larger portion of the more than three hundred problems from the OHA'Cinic archive that are already digitized. A formal user study involving lyceum students could be conducted to quantitatively assess the pedagogical impact of these interactive visualizations compared to traditional static materials. This would provide valuable data on the effectiveness of inquiry-based learning facilitated by our tools.

Furthermore, the integration of these visualizations into a broader AI Tutoring System, as mentioned in [5], is a key long-term objective. Such a system could use the interactive components we have built, using another AI layer to provide students with dynamic feedback, hints, and personalized learning paths based on their interaction with the visualizations. This would represent a significant step toward creating a truly intelligent and adaptive learning environment from a rich historical resource. The approach detailed in this paper provides a foundation for these future undertakings.

Acknowledgments. This work was supported by project SIBIA – 011301, "Information Systems based on Artificial Intelligence".

References

[1] T. Bumbu, L. Burtseva, S. Cojocaru, A. Colesnicov, L. Malahov, *Digitization of Romanian Historical Printings*, Vladimir Andrunachievici Institute of Mathematics and Computer Science, Chișinău, 2023. ISBN 978-9975-68-497-2.

[2] C. McKay, J. Cumming, and I. Fujinaga, "Lessons learned in a large-scale project to digitize and computationally analyze musical scores," *Digital Scholarship in the Humanities*, Vol. 36, Suppl. 2, pp. ii198–ii202, 2021. DOI: https://doi.org/10.1093/llc/fqaa058.

[3] O. Caftanatov, V. Demidova, and T. Verlan, "An Approach to Recognizing and Digitizing Old Handwritten Documents with Mathematical Heterogeneous Content in Cyrillic script," *Com-*

puter Science Journal of Moldova, vol. 33, no. 1(97), 2025, pp. 54–67. DOI: https://doi.org/10.56415/csjm.v33.03.

[4] A. Oliveira, "History of the Bernoulli Principle," in *Advances in Mechanism and Machine Science. IFToMM WC 2019* (Mechanisms and Machine Science, vol 73), T. Uhl, Ed. Cham: Springer, 2019, pp. 1161–1178. DOI: 10.1007/978-3-030-20131-9_115.

[5] O. Caftanatov and A. Parahonco, "The Virtual GPT Assistant: Emulating the Teaching Style of a Real Professor," in *Proceedings of International Conference dedicated to the 60th anniversary of the foundation of Vladimir Andrunachievici Institute of Mathematics and Computer Science*, MSU, 2024, pp. 247–258.

[6] S. Freeman et al, "Active learning increases student performance in science, engineering, and mathematics," *PNAS*, vol. 111, no. 23, pp. 8410–8415, 2014. Available: https://www.pnas.org/doi/10.1073/pnas.1319030111.

[7] "A platform for mathematical visualizations," MSU, 2025. [Online]. Available: https://mathvisualisations.netlify.app/.

Olesea Caftanatov[1], Marius Cerescu[2], Valentina Demidova[3], Tatiana Verlan[4]

[1,2,3,4]Vladimir Andrunachievici Institute of Mathematics and Computer Science, Moldova State University

[1] ORCID: https://orcid.org/0000-0003-1482-9701
E–mail: olesea.caftanatov@math.md
[2]ORCID: https://orcid.org/0009-0000-5704-6577
E–mail: marius.cerescu@math.md
[3] ORCID: https://orcid.org/0009-0006-7260-8375
E–mail: valentina.demidova@math.md
[4] ORCID: https://orcid.org/0009-0006-4519-1105
E–mail: tatiana.verlan@math.md

An Apriori-Based Recommendation System for Grocery Stores

Cristian Cemîrtan

Abstract

This paper showcases an automated recommendation system tailored for grocery retail. The system analyses data from transaction history (sales) to generate real-time, personalised product suggestions for customers. The core functionality of the system relies on identifying relevant products for each customer and displaying them in an order that considers not only the statistical relevance of each product, but also its economic value.

Keywords: Apriori, data preprocessing, recommendation system, model training.

1 Introduction

In the current context of retail, especially in the food sector, customer loyalty requires the adoption of intelligent solutions that optimise the shopping experience and, at the same time, directly contribute to increasing revenues.

One of these solutions is represented by the implementation of a personalised recommendation system, capable of suggesting relevant products to each customer, based on their previous purchasing behaviour. Prior to detailing the functionalities of the system, it is essential to highlight the key concepts that underlie this approach:

- **Recommendation** – the process by which a customer is offered a selection of products considered relevant, i.e., close to their preferences, inferred from their transaction history.

- **Relevance** – the extent to which a product is frequently associated with other products purchased by customers, reflecting its popularity within transactions. The more often a product appears in completed shopping carts, the more relevant it is considered in the context of a recommendation.

From a behavioural perspective, the system takes advantage of a customer's natural tendency to follow suggestions considered 'smart' or 'personalised', thus increasing the chances of impulsive or additional purchases.

This effect is well known in modern marketing strategies and is often associated with the phenomenon of consumerism, in which the consumer accepts the system's suggestions as beneficial and tailored to their needs [1].

2 Data Understanding

To develop a recommendation system, a profound understanding of the provided dataset is required. In this context, the dataset is provided by a grocery store and reflects the shopping activity of customers and the catalogue of available products.

The objective is to analyse the structure and relationships between this data, as well as to evaluate its quality and usefulness in the context of recommendations.

The provided dataset is illustrated in Table 1. Analysing it, we have the following data processing considerations:

1. **Data cleaning** – some transactions or products may contain missing data.

2. **Data matching** – relationships between sets should be strengthened by using common IDs. This will allow links to be made between actual transactions and descriptive product information.

3. **Data transformation** – for the data to be compatible with machine learning models that generate recommendations (e.g., Apriori), it is necessary to transform product lists into standardised formats (e.g., converting numeric text to a binary format).

Table 1. Presented data sets

Data set	Fields	Description
Transactions	Product ID set	The field enumerates the bought products in a transaction.
Products	ID, Price, Name	Self-explanatory.
Product image files	ID	Each file is named in accordance to its associated product.

3 Data Preprocessing

Data preprocessing is a crucial step in the development of any recommendation system [2]. It involves a series of processes through which raw data is cleaned and transformed so that it can be used effectively for model training.

In this case, the datasets provided by the grocery store (`Products` and `Transactions`), stored as CSV files, were subjected to be preprocessed.

3.1 The `Products` dataset

The `Products` dataset, initially composed of 246 records, contains information about the items available in the store.

Our initial observations on Table 2 are that some products have missing names, and that the names are not uniformly formatted (existence of capitalisation variations).

Table 2. `products.csv` sample

ID	Price	Name
0	2.51	Swiss Cheese
1	52.44	cherry coke
2	28.22	Bio Coke
3	4.00	
4	1.04	Scrambled egg
5	3.90	Alkopops(rum/cherry)

The following operations were performed on the set:

1. **Missing value elimination:**
   - 3 records were identified and eliminated that had missing values in the 'Name' field.
   - A product record with missing values cannot be properly identified and is therefore unsuitable for recommendation.

2. **Product name consistency:**
   - To avoid variations caused by different writing styles and to ensure consistency of display, all names were converted to capitalised format, meaning that each word begins with a capital letter (for example, 'cherry coke' to 'Cherry Coke').

3. **Ambiguous product identification:**
   - Some products were found that have the same textual name but are associated with different icons.
   - This visual discrepancy indicates that although the products are named the same, they may have distinct physical or qualitative attributes (e.g., different weight, type of packaging).

- These products were kept as distinct entities to faithfully reflect the commercial reality.

After preprocessing, the `Products` dataset now contains 243 (3 removed) cleaned and usable objects to train a recommendation system.

3.2 The `Transactions` dataset

Moving on to the `Transactions` dataset, which contains lists of products purchased together in a single shopping session. The structure of each record is represented by a list of purchased product IDs (separated by spaces).

The `Transactions` dataset contains 2869 completed shopping carts, each representing a list of product IDs purchased together in a single shopping session.

Table 3. `transactions.csv` sample

Product ID set
224 80 109 177 50 ...
56 95 106 186 103 ...
9 196 184 119 8 ...
9 196 184 119 88 ...
228 9 193 127 163 ...
94 9 22 133 107 ...

Our initial observations on Table 3 are that each transaction is a string of space-separated IDs, and that the structure is provided textually, but requires conversion to a binary format for compatible interpretation in Section 4.

On this set, a single operation was performed – transaction integrity validation: each transaction was checked to contain only references to the products not affected by Step 1 in Subsection 3.1.

Finally, it was found that there is no transaction that references removed products. Thus, the `Transactions` set remains unchanged in quantity, still containing 2869 valid transactions.

4 Training

To train a recommendation system based on transactional data, the Apriori algorithm was used. Apriori is a classic data mining method used to identify frequent item sets and generate association rules between products. It is based on the analysis of repeated transactions and product co-occurrences in shopping carts.

The algorithm works on the principle of item frequency and uses prior knowledge about their properties to gradually expand frequent product sets. It is an iterative algorithm, where:

1. We start with frequent sets of size k.

2. These are expanded into candidate sets of size $k + 1$, which are then tested for the minimum *support* metric.

In short, if certain products frequently appear together in transactions, then there is a high probability that these products will also appear together in future customers' carts. The Apriori algorithm possesses the following essential properties [3]:

- **Apriori** – all subsets of a set of frequent items must also be frequent.

- **Anti-monotonicity** – if a set of items is not frequent, none of its supersets can be frequent.

The evaluation of the quality of the association rules generated by the Apriori algorithm is carried out through three fundamental metrics from market basket analysis: Support, Confidence, and Lift. They provide a quantitative perspective on the frequency and relevance of the correlations identified between products.

Let:

- productA – the antecedent product (or set of products), i.e., present in the shopping cart.
- productC – the consequent product, i.e., recommended based on the association.

The following functions are defined [4]:

- Support(products) – represents the proportion of transactions containing the *products*, divided to the total number of transactions.
- Confidence(productA, productC) – measures the probability that productC will appear in a transaction that already contains $productA$.
- Lift(productA, productC) – represents an indicator of the dependence between productA and productC. Values greater than 1 suggest a positive correlation.

The Apriori algorithm is based on these essential steps [5]:

1. **Candidate generation** – identifies all product combinations in transactions.
2. **Filtering by support** – keeps only combinations that appear in transactions in a percentage greater than the minimum support threshold.
3. **Iterative cardinality increase** – arrays with frequent products of length k are combined to generate new arrays of length $k+1$, until $k = \text{maxK}$.
4. **Rule extraction** – from frequent sets, rules of the form $A \to B$ are generated (A – antecedent, B – consequent), which are kept only if they meet the minimum confidence threshold.

5 Implementation in KNIME

For practical implementation, the KNIME platform was used, which provides visual modules for data processing, model application, and export of results. From Figure 1, we illustrate the following steps:

1. **Preprocessing the `Products` dataset** (Subsection 3.1):
   (a) Read the `products.csv` file containing the product ID, price, and name.
   (b) Apply name capitalisation for uniformity.
   (c) Remove products containing missing values.
   (d) The cleaned product set is saved in a new CSV file.

2. **Processing the `Transactions` dataset** (Subsection 3.2):
   (a) Read the `transactions.csv` file, where each row represents a transaction (a list of product IDs).
   (b) Transform the character strings into collections of items (lists).
   (c) Filter out transactions containing products removed in Step 1c.
   (d) The result is ready for the rule generation process.

3. **Modeling by generating association rules:**
   (a) The Apriori algorithm is applied through the 'Association Rule Learner' component implemented by Christian Borgelt [6], with the following parameters:
   - **Minimum support:** 1% – to exclude rare products, which are not statistically relevant.
   - **Minimum confidence:** 50% – to ensure that the rule has a significant probability of occurrence.

   (b) A filter is applied to keep only relevant rules with $Lift > 1$.

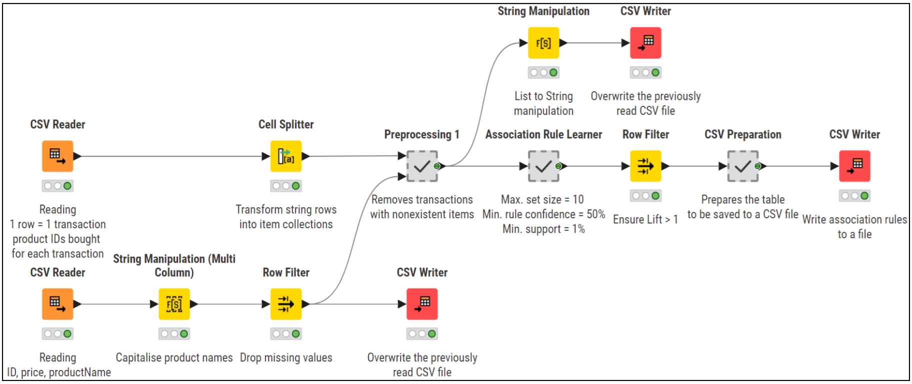


Figure 1. The KNIME workflow

(c) The final rules are prepared and exported to a CSV file.

Through the visual environment provided by KNIME, the process becomes easy to follow, modular, and reproducible, thus facilitating the development of an efficient grocery recommendation system.

After running the workflow, 8631 association rules were obtained, each with the following columns: Antecedent (single product ID), Consequential (multiple IDs), Lift, Support, and Confidence.

The resulting rules were exported to a CSV file, which will serve as a source of knowledge for the recommendation system.

Within an external application used by customers, the generated association rules are applied as follows:

1. The exported rule set is first loaded and sorted in ascending order by *Confidence*, enabling fast filtering and selection.

2. For the active shopping cart, a relevant subset of rules is extracted, i.e., those where productA (an item currently in the cart) appears as *Antecedent*.

3. This subset is then re-ordered in descending order according to the following priorities:

- **Lift** – to emphasise rules with the strongest associative impact.
- **Support** – to prioritise rules observed more frequently in transaction data.
- **Mean of consequent product prices** – to maximise the potential economic value of the recommendation.

For example, if a customer chooses items with IDs 157 and 19, then the recommendations would look similar to those in Table 4.

Table 4. `ordered_rules.csv` sample

Antecedent	Consequent	Lift	Support	Confidence
157	93, 128, 235, 12	10.826	6	20
157	93, 128, 235	10.826	6	20
157	93, 128, 12	10.826	6	20
19	46, 200	10.433	6	20

6 Conclusion

This paper addressed the development of a recommendation system based on association rules, implemented in the context of grocery shopping. The aim was to integrate the Apriori algorithm in order to extract association rules with their practical insights ready to be leveraged in external retail applications.

Preprocessing techniques were applied to transform raw textual data into binary data to be interpreted as valid training arguments for our recommendation system model. Subsequently, the trained model generated a set of 8631 associative rules.

The imposed training thresholds: $Support \geq 1\%$, $Confidence \geq 50\%$ – ensured a balance between coverage and precision, leading to a robust set of rules.

References

[1] S. Sridharan, “New consumerism: Where algorithms shape aspirations,” *Policy Circle*, 2025. [Online]. Available: https://www.policycircle.org/society/social-media-role-in-consumerism.

[2] N. Chaudhary and D. R. Chowdhury, “Data Preprocessing for Evaluation of Recommendation Models in E-Commerce,” *Data*, vol. 4, no. 1, Article No: 23, 2019. DOI: https://doi.org/10.3390/data4010023.

[3] C. K. S. Leung, “Anti-monotone constraints,” in *Encyclopedia of Database Systems*, L. Liu, M. T. Özsu, Eds. Boston, MA: Springer, 2009, pp. 98–98. DOI: https://doi.org/10.1007/978-0-387-39940-9_5046.

[4] R. Agrawal et al., “Mining association rules between sets of items in large databases,” *ACM SIGMOD Record*, vol. 22, no. 2, pp. 207–216, 1993. DOI: https://doi.org/10.1145/170036.170072. in *1993 ACM SIGMOD international conference on Management of data*.

[5] M. Al-Maolegi and B. Arkok, “An Improved Apriori Algorithm for Association Rules,” *arXiv e-prints*, 2014. DOI: https://doi.org/10.48550/arXiv.1403.3948.

[6] KNIME, “Association Rule Learner (Borgelt),” [Online]. Available: https://hub.knime.com/knime/extensions/org.knime.features.ext.itemset/latest/org.knime.ext.itemset.nodes.learner.AssociationRuleLearnerNodeFactory.

Cristian Cemîrtan

ORCID: https://orcid.org/0009-0001-7113-4650
Moldova State University, Vladimir Andrunachievici Institute of Mathematics and Computer Science
E–mail: cemirtan.cristian@outlook.com

Proceedings of Workshop on Intelligent Information Systems
WIIS2025, October 16-18, 2025, Chisinau, Republic of Moldova

Some Modifications of the Demoucron-Malgrange-Pertuiset Algorithm for Testing Planarity of Undirected Graphs

Constantin Ciubotaru

Abstract

Some modifications are proposed to the Demoucron-Malgrange-Pertuiset (DMP) algorithm for testing planarity of undirected graphs that specify the notion of segment (fragment, bridge) and develop an algorithm for calculating and updating the segments after each iteration, based on a depth-first search strategy (DFS).

Keywords: biconnected graph, DMP algorithm, segment, face, embedding, updating.

1 Introduction

The DMP algorithm offers certain possibilities for the development of automatic drawing methods of the planar graphs, unlike other algorithms in this field.

It is well known that drawing hierarchical structures (graphs, trees, schemes) is one of the most attractive ways to present information. Thanks to the geometric structures used in drawing, the range of applications of this subject is very wide: mathematics, computer science, social networks, databases, bioinformatics, linguistics, artificial intelligence, etc.

One and the same graph can be drawn in several ways. Some may be simpler, more comprehensible, having an attractive aesthetic appearance, others – more difficult to perceive, with an unsatisfying structure.

All drawings from Figure 1.1 represent the same graph. Figure 1.1(a) represents a variant drawn manually, the other variants being drawn semi-automatically: Figure 1.1(b) represents a variant with placement randomness of nodes that we will call “spaghetti”, Figure 1.1(c) – a planar variant obtained automatically by the method of circular orbits, and the Figure 1.1(d) – a variant based on the faces structure obtained by application of the *DMP* algorithm.

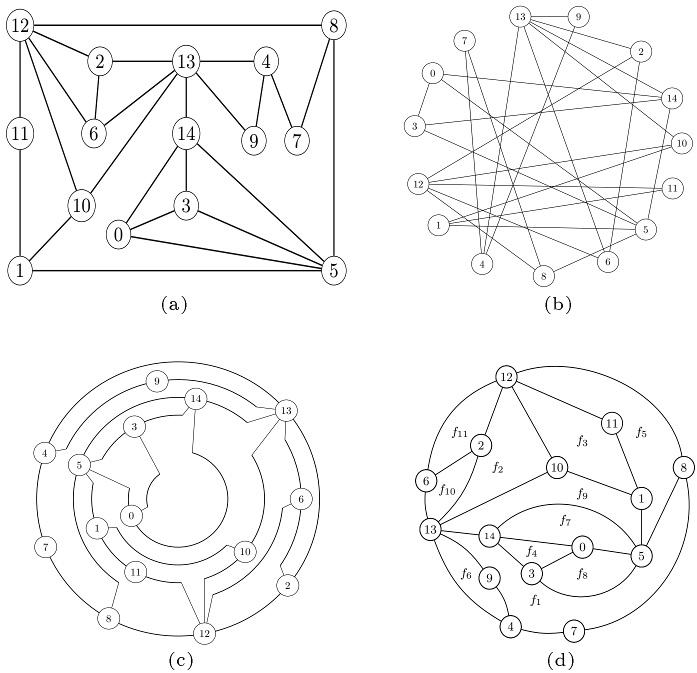


Figure 1.1. Four figures

Having multiple representations of the same graph, it is quite easy to choose the most suitable option, especially having a lot of available appreciation criteria, usually of an aesthetic nature. It is quite difficult to transmit these criteria to the computer. Here comes not only the problem of formalizing the criteria, but also the fact that some of them may be contradictory. In these cases, inevitably arise the compromise situations. Most planarity checking algorithms [1] are based on the application of two methods. The first method refers to the Kuratowski theorem (or the Wagner theorem). The second method is based on the fundamental property of the closed Jordan curve, which states that no point inside the curve can be connected to any other point outside it without intersecting the curve.

The Demoucron-Malgrange-Pertuiset algorithm (DMP) ([2, 3, 4, 5, 6]) is a well-known algorithm based on this property, which, in parallel with the planarity test, also constructs a partition of the graph (a set of faces, $\mathscr{F}$), that offers possibilities for the development of automatic graph drawing methods. This problem is no less complicated than the planarity test. This paper proposes several essential modifications to the DMP algorithm, which simplify its implementation and make it easier to understand. In the work [6], the DMP algorithm was called "the γ algorithm".

The implementation of the DMP algorithm requires to specify the notions (structures) used in the description of the algorithm, such as *segment* (*fragment, bridge, component*), *face*, α*-chain*, but also to choose the methods for constructing these structures. Most descriptions of this algorithm bypass the practical aspects related to the construction of these structures and the actual drawing.

2 Demoucron-Malgrange-Pertuiset algorithm

Any subgraph G_i of a planar graph G can be obtained from the drawn image of the graph G by removing some number of edges and vertices. Thus, any subgraph of a planar graph is also planar. If G_i is a subgraph of the graph G, then the vertices of the subgraph G_i are called *contact vertices.* The reverse action is also possible: building a new planar subgraph by adding to G some number of edges and vertices from $G \setminus G_i$. This is the idea of the DMP algorithm. When describing the algorithm, the notions of *fragment* and α*–path* play an important role.

Below is inserted the pseudocode of DMP algorithm [2, 3, 4, 5, 6] (Algorithm 2.1). By $G_i = (V_i, E_i)$ we denote any subgraph of the graph $G = (V, E)$. A planar graph divides the plane into regions (bounded by edges), called *faces*. We denote the set of faces by $\mathscr{F}$ and the set of segments (fragments) by $\mathscr{S}$. The face f is valid for the segment s if the end vertices of s are contact vertices of f. By $\mathscr{F}(s)$ we denote the set of all faces valid for s.

Algorithm 2.1. ***Pseudocode of DMP algorithm***

1. Let $G = (V, E)$ be a biconnected undirected graph.
2. An arbitrary simple cycle of the graph G is selected and embedded (drawn) obtaining a planar graph (subgraph of the graph G) denoted by G_c=(V_c, E_c).
3. ***if*** G_c=G ***then return*** "The graph is planar".
4. The set of all fragments $\mathscr{S}$ of G with respect to G_c is built.
5. ***if*** $\mathscr{S}$=$\varnothing$ ***then return*** "The graph is planar".
6. For all fragments s of $\mathscr{S}$ build the set of all admissible faces for s, $\mathscr{F}(s)$.
7. ***if*** $\exists$ s ***with*** $\mathscr{F}(s)$=$\varnothing$ ***then return*** "The graph is nonplanar".
8. ***if*** $\exists s$ ***with*** $|\mathscr{F}(s)|$=1, $\mathscr{F}(s)$=$\{f\}$, ***then*** let l be an arbitrary α-path of the fragment s ***else*** let s be an arbitrary fragment for which $|\mathscr{F}(s)|$>1, f an arbitrary face from $\mathscr{F}(s)$, and l – an arbitrary α–path of the fragment s.
9. l is embedded in f and $\mathscr{F}$ is modified. $G_c := G_c \cup l$.
10. ***goto*** 4.

From the description of the DMP algorithm, it is observed the importance of checking the graph for connectedness and biconnectivity. If the graph is connected, then the set of bicomponents can be constructed, and apply a well-known theorem that a graph is planar if and only if all its bicomponents are planar [7]. Thus, without imposing serious restrictions, we will consider the problem of testing planarity and graph drawing only for biconnected graphs. The biconnected graph will contain only one bicomponent, the graph itself.

For implementation of the algorithm, additional explanations of notions "fragment (segment)", "H-fragment", "bridge" are needed. It is necessary to mention (or add) the methods that select the initial cycle and build the sets $\mathscr{S}$, $\mathscr{F}$, $\mathscr{F}(s)$.

It is easy to see that at any iteration of the algorithm, only one α–path is drawn. As an α–path is at the same time a fragment (segment), we can operate on the implementation of the algorithm only with slightly modified α–path.

As follows, we present some modifications of the DMP algorithm concerning the specification of the notions used and the application of simple and most comprehensible methods in order to achieve the decisive phases, possibly decreasing the efficiency, bypassing the methods considered more effective, but sometimes difficult to understand and implement.

3 Segments building

Let us first specify the notion of *segment* which will be used frequently along the way.

Definition 3.1. Segment

The notion ***segment*** of $G = (V, E)$ with respect to the subgraph $G_i = (V_i, E_i)$ is defined as:

- the edge $e = (u, v)$, $e \notin E_i$, $u, v \in V_i$ (*segment–edge*), or
- the simple path $(e_1 e_2 \dots e_n) = ((u_1 v_1) \dots, (u_{n-1} v_{n-1})(u_n v_n))$, $e_1, e_2, \dots, e_n \notin E_i$, $u_1, v_n \in V_i$, (u_1, v_n are contact vertices), $\{v_1, u_2, v_2, \dots, u_{n-1}, v_{n-1}, u_n\} \cap V_i = \varnothing$ (*segment–path*).

The ***Segment Building*** algorithm includes the following stages:

1. The biconnectivity of the graph G=(V, E) is checking and a fundamental cycle of maximal length G_c=(V_c, E_c) is selected, which, obviously, is a planar subgraph of G [8].
2. The set ***Segment–Edges*** is building.
3. The set ***Segment–Paths*** is building, applying the DFS strategy.
4. $\mathscr{S}$=***Segment–Edges*** $\cup$ ***Segment–Paths***.

It should be noted that throughout the exposure and during performing the operations, it will be taken into account that $(u\ \ v) = (v\ \ u)$, the edge $(u\ \ v)$, if necessary, can also be interpreted as a path $((u \ \ v))$, also "palindrome" path will be considered equal: $((v_0\, v_1)(v_1\, v_2)(v_2\, v_3)(v_3\, v_4)) = ((v_4\, v_3)(v_3\, v_2)(v_2\, v_1)(v_1\, v_0))$.

Choosing the cycle of maximum length contributes to shrink the number of iterations, which depends on the number of edges included in E_i. After the last iteration, $E_i = E$; thus, it is important to include at each iteration as many edges as possible in the updated E_i.

4 Segments embedding

The segments embedding algorithm is an iterative process, which at each iteration embeds (draws) one segment. At the same time, the planarity of the graph is checked, and the sets $\mathscr{F}$, $\mathscr{S}$, V_i, and E_i are modified. The process lasts as long as the set $\mathscr{S}$ is not empty. Finally, if the graph is planar, it will be drawn completely; otherwise, a planar subgraph will be drawn and the set $\mathscr{N}_p$ of nonembeddable segments (which generate edge intersections) will be printed.

Initially, the process starts with the subgraph $G_c = (V_c, E_c)$, two faces – internal face $f_1 = V_c$, external face $f_0 = V_c$, *iter*=1 and the set of segments $\mathscr{S}$, built by the algorithm ***Segment Building***. It should be mentioned that f_0 will denote the external face throughout the execution of the algorithm.

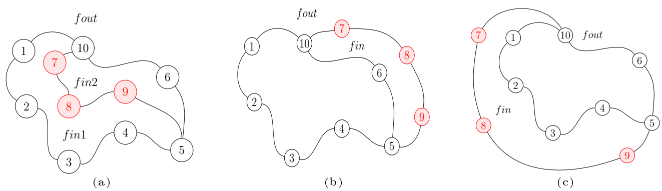


Figure 4.1. Face split schema

The ***Segments embedding*** algorithm embeds a segment into one of its valid faces. If the selected face is internal, it will be divided into two new internal faces. In the case of an external face, the division will also produce two new faces: one internal and one external. This operation leads to the modification of the set $\mathscr{F}$: the selected face is replaced with two newly constructed faces. The scheme of this process is inserted in Figure 4.1.

Since, as a result of this operation, new faces and contact vertices appear, the ***Segments updating*** algorithm is still called, which updates the sets $\mathscr{S}$ and the set of faces valid for segments of $\mathscr{S}$.

Segments embedding generates new faces and new contact vertices. Thus, the need to modify the set of segments inevitably arises. Due to the fact that only two new faces appear at each iteration, modification of only the segments that have certain tangents to these faces will be undergone. More precisely, the segments, inside which new contact vertices appeared, must be modified.

5 Segments updating

Obviously, the set of valid faces for these segments will also be changed. This problem is solved by the ***Segments updating*** algorithm. The idea behind the change is simple. Any segment containing at least one new contact node will be split into two new segments. For example, if the segment $s{=}(v_0\, v_1\, \ldots v_i\, v_c\, v_{i+1}\, \ldots\, v_{n-1}\, v_n)$, where v_0 and v_n are old contact vertices and v_c – new contact vertex, then s will be split into two new segments: $s_1 = (v_0\, v_1\, \ldots\, v_i\, v_c)$ and $s_2{=}(v_c\, v_{i+1}\, \ldots\, v_{n-1}\, v_n)$. This algorithm will be called after each embedding, which generates new contact vertices. It should be noted that for segment–edges, the ***Segments updating*** algorithm will not be called.

6 Modified DMP algorithm

The modified DMP algorithm ($MDMP$) is obtained by assembling all the algorithms presented above, including the biconnectivity verification. By $\mathscr{N}_p$ we denoted the set of segments that could not be incorporated. Below is inserted the pseudocode of the algorithm (Algorithm 6.1).

Algorithm 6.1. *Modified DMP algorithm*

1. Let $G = (V, E)$ be an arbitrary undirected graph. The algorithm exposed in [13] is applied for checking biconnectivity and building the fundamental set of cycles.
2. If the graph is biconnected, a fundamental cycle of maximum length is chosen, subgraph of the graph G is denoted by G_c=(V_c, E_c). Otherwise ***return*** "The graph is not biconnected".
3. ***if*** G_c=G ***then return*** "The graph is planar".
4. ***Initial assignments***:
$\mathscr{F}$:=$\{f_0, f_1\}$; f_0:=V_c; f_1:=V_c; $\mathscr{S}$:=$\varnothing$; $\mathscr{N}_p$:=$\varnothing$; *iter*:=1;
5. The ***"Segments building"*** algorithm is called, which builds the set of all segments $\mathscr{S}$.
6. ***do while*** $\mathscr{S} \neq \varnothing$
7. *iter*:=*iter*+1;
8. ***for all*** $s \in \mathscr{S}$ ***do*** $\mathscr{F}(s)$:=$\{f|\ f \in \mathscr{F}$, f is valid for $s\}$; ***end for all*** .
9. The ***"Segments embedding"*** algorithm is called, which checks the non-planarity of the graph, embeds a segment, modifies the set of faces $\mathscr{F}$, and updates the set of segments $\mathscr{S}$ by calling the algorithm ***Segments update***.
10. ***end do while***
11. ***return*** ("The graph is planar", $\mathscr{F}$).

7 Conclusions

This paper presents some modifications of the DMP algorithm with all necessary explanations. The algorithm can be used for testing planarity of graphs, generating planar graphs, and elaborating automatic graph drawing methods.

The algorithm also works for nonplanar undirected graphs by final construction of a planar subgraph and displaying the set $\mathscr{N}_p$ of segments that cannot be embedded, so as they generate edge intersections when drawing. Unlike the original DMP algorithm, which, after each iteration, recalculates the set of segments (fragments), the modified algorithm calculates this set only once. It is easy to develop an algorithm, which for the set of constructed faces $\mathscr{F}$ and any segment from $\mathscr{N}_p$, will allow us to find the minimum number of intersections needed to draw this segment. This algorithm, of course, does not solve the crossing number problem, which is known to be NP-complete [9], but can be used to build a plausible drawing version.

Acknowledgments. The project SIBIA 011301 has supported part of the research for this paper.

References

[1] *Handbook of Graph Drawing and Visualization.* R. Tamassia, Ed. Chapman & Hall, 2016, pp. 1-862.

[2] G. Demoucron, Y. Malgrange, R. Pertuiset, "Graphes Planaires: Reconnaissance et Construction de Representations Planaires Topologiques," *Revue Française de Rech. Oper.*, vol. 8, pp. 33–47, 1964.

[3] A. Kohnert, "Algorithm of Demoucron, Malgrange, Pertuiset," 2004, [Online]. Available: http://www.mathe2.uni-bayreuth.de/EWS/demoucron.pdf

[4] A. Gibbons, *Algorithmic Graph Theory*, Cambridge: Cambridge University Press, 1985, 259p.

[5] J.A. Bondy and U. S. R. Murty, *Graph Theory with Applications*, New York, USA: Elsevier Science Publishing Co., Inc., 1976, 270 p.

[6] В. А. Емеличев, О .И. Мельников, В. И. Сарванов, Р. И. Тышкевич, *Лекции по теории графов.* "Наука", Глав. ред. физико-математической лит-ры, 1990, 382 pp.

[7] C. Berge, *Theorie des graphes et ses applications*, Paris: DUNOD, 1958, 277 p.

[8] C. Ciubotaru, "The modified deep first search algorithm: functional implementation," *Computer Science Journal of Moldova*, vol. 33, no. 1(97), pp. 129–140, 2025.

[9] M. R. Garey and D. S. Johnson, *Computers and intractability. A Guide to the Theory of NP-Completeness,* (Series of Books in the Mathematical Sciences), 1st ed., San Francisco, Calif.: W. H. Freeman and Co.(ed.), 1979, 340 p.

Constantin Ciubotaru

Moldova State University, Vladimir Andrunachievici Institute of Mathematics and Computer Science
E–mail: constantin.ciubotaru@math.usm.md, chebotar@gmail.com
ORCID: https://orcid.org/0009-0005-8896-0966

Proceedings of Workshop on Intelligent Information Systems
WIIS2025, October 16-18, 2025, Chisinau, Republic of Moldova

Comparative Evaluation of AI-Powered Image Generation Platforms

Vlada Colesnicova, Inga Țițchiev, Argentina Fron,
Daniela Caganovschi, Olesea Caftanatov, Alexandru Colesnicov

Abstract

This article examines the use of artificial intelligence in folklore research, using Russian and Bessarabian proverbs and sayings as a case study. Artificial intelligence (AI) tools were used to generate or find illustrations for these folklore objects. The study presents a comparative analysis of several AI image generation platforms, evaluating their performance, features, and limitations. The key evaluation criteria included image quality, generation speed, economic models, and prompt accuracy.

The findings highlight significant differences in the capabilities of the platform. The results suggest that the choice of an AI tool should be based on specific user needs, balancing factors such as image quality, speed, cost, and semantic accuracy.

Keywords: Idioms, AI image generation, comparative analysis, generative AI, image synthesis, neural networks.

1 Introduction

Rapid advancement of artificial intelligence (AI) has introduced new methods for visual content creation, particularly through the application of neural networks. These systems are able to produce unique and contextually relevant images by learning patterns from large-scale datasets. In 2023 alone, approximately 15 billion AI-generated images were produced, while 87% of companies reported perceiving AI-based solutions as a source of competitive advantage [1] – [3].

AI-based image generation platforms transform textual descriptions (prompts) or uploaded images into visual representations. Depending on their design, such platforms are accessible as web-based applications, standalone tools, or professional design utilities. The generative process typically includes the following:

1. input of a textual prompt;
2. natural language processing and semantic interpretation;
3. image synthesis using generative models;
4. post-processing for visual refinement.

The present study investigates the comparative performance of selected platforms with a focus on their ability to interpret abstract prompts, including Bessarabian and Russian proverbs. Such prompts provide a robust test of semantic accuracy and contextual adaptability.

For several selected reviews on the topic, please refer to [9] – [12].

2 Research goal and methodology

The primary goal of this research is to identify the specific advantages, weaknesses, and limitations of selected AI image generation platforms. The following research methods were applied:

- Comparative analysis of five platforms (Google Gemini [4], OpenArt.ai [5], ChatGPT Plus [6], NightCafe [7], CGDream [8]) was evaluated using Bessarabian proverbs and sayings as prompts. In parallel, ten additional platforms (ReCraft, Sora, Canva, FreePik, Stable Diffusion, Fusion Brain, Pollo AI, Shedevrum, Supermaker, and Gencraft) were tested with Russian proverbs and sayings;
- A combination of logical and synthesis-based analysis was applied to compare platform performance and identify recurring patterns;
- Specialized assessment criteria were adopted to measure semantic relevance, contextual accuracy, and visual quality of the generated outputs;
- To ensure consistency, the same set of idiomatic expressions and proverbs was used across all platforms, allowing for controlled com-

parisons of contextual interpretation;

- Performance indicators included generation time, output quality (both aesthetic and semantic), accessibility for end-users, and subscription models.

This methodological design allows for both qualitative and quantitative insights into how effectively AI image generation platforms interpret abstract, culture-specific prompts such as proverbs and idiomatic expressions. The results of this comparative analysis are presented in the next section.

3 Results and Comparative Analysis

Google Gemini demonstrated notable strengths, including the ability to generate unlimited images free of charge and a rapid image synthesis process that consistently produced outputs of high visual quality, see Fig.1. However, a key limitation lies in the recent restriction of its functionality: the free version no longer supports direct image generation, offering instead only image editing or analysis of uploaded content.

Figure 1. The result of Google Gemini's proverb to image transformation

OpenArt.ai stands out for its user-friendly interface, wide range of customization options, and stylistic versatility, which are further enhanced by support for custom model training. At the same time, the platform exhibits occasional inconsistencies in semantic alignment, relies on a credit-based usage system, and poses challenges when attempting to manage large-scale content generation.

The integration of **DALL·E into ChatGPT** Plus enables the generation of high-quality, semantically accurate images, particularly effective for schematic or comic-style representations. Furthermore, the platform allows for unlimited generation. Its main drawback, however, is a slower processing speed, with image creation typically requiring between two and five minutes.

NightCafe offers rapid image generation, initial free credits, and a vibrant community that fosters engagement through rewards and competitions. Despite these strengths, the reliance on a credit-based system limits extensive use, with each generated image costing approximately 0.5 credits once the initial free quota has been exhausted.

CGDream is valued for its fast generation process, daily allocation of free credits, and opportunities to earn additional credits via competitions. Nevertheless, the platform demonstrates low semantic accuracy and frequently produces images that fail to align with textual prompts.

ReCraft delivers highly detailed and visually impressive images, alongside a generous number of free attempts that do not require an immediate subscription. The platform also benefits from an intuitive interface, fast generation, and support for multiple styles and formats. Its limitations, however, include fewer options for deep customization compared to professional-grade tools and occasional restrictions on image size or resolution for free users.

Sora distinguishes itself with highly accurate and stable generation, producing realistic images with impressive speed. Users benefit from several free daily attempts and browser-based accessibility, eliminating the need for installation. On the downside, continued use requires a subscription after the free quota is consumed, while customization options for depth and generation parameters remain limited, and the diversity of styles is narrower than that of some competitors.

Canva excels in accessibility, offering a simple, intuitive interface supported by an extensive library of templates and design elements. Its integrated workflow allows users to combine image generation with editing, text addition, and graphic design, making it suitable even for those without prior expertise. The platform's main constraints stem from the subscription model: advanced AI-generation functions are locked behind a paywall, the free version is heavily restricted, and some assets include watermarks. Additionally, image rendering is comparatively slower than in specialized generators.

FreePik combines AI image generation with access to a vast library of ready-made images, icons, and templates, making it particularly convenient for quick design needs. Its generator is fast and requires minimal configuration, offering a straightforward entry point for casual users. However, free use is capped at around twenty downloads, and the quality of AI-generated images falls short of more specialized platforms. Moreover, much of the content library requires a paid subscription to unlock.

Stable Diffusion is distinguished by its open-source nature and complete accessibility, allowing unlimited image generation when deployed locally. The platform offers an extensive ecosystem of models and modifications, enabling high levels of customization and adaptation to specific user needs. These features make it particularly attractive to advanced users seeking flexibility. Nonetheless, its use is associated with significant barriers to entry: installation requires a powerful computer, and the system can be challenging for beginners. Without additional fine-tuning and external models, output quality may be inconsistent, and the steep learning curve can lengthen the initial setup and usage process.

Fusion Brain provides a wide selection of artistic styles and a convenient interface that allows users to enhance images by selectively editing different regions. While the platform excels in stylistic variety, its limitations are evident in relatively long generation times and strict daily usage restrictions, offering only ten generations per day.

Pollo AI is valued for its fast image generation based on textual prompts. However, its functionality is constrained by the absence of style options and a restrictive credit-based model: only ten credits are available to new users, with each prompt consuming three credits. As a result, the platform's accessibility is highly limited without additional purchases.

Shedevrum combines ease of use with generous daily allowances, offering up to seventy generations per day. Users can also select output aspect ratios, and the resulting images are generally sharp and visually appealing. Despite these advantages, generation speed can be inconsistent, with some prompts requiring up to two minutes to process.

Supermaker delivers rapid image generation but is restricted by a limited credit system, providing only eight free credits per day, with each image costing two. Furthermore, the platform offers a narrow range of styles, which reduces flexibility for users seeking diverse outputs.

Gencraft offers fast generation, an intuitive interface, and an extensive variety of styles and customization options, making it appealing for both novice and experienced users. Nevertheless, its free usage is restricted to five prompts per day, and the platform's heavy reliance on advertising detracts from the overall user experience.

4 Prompt interpretation: literal meaning

A key challenge in AI image generation is the accurate interpretation of user prompts, particularly when they involve abstract or idiomatic expressions. This study highlighted the difficulty of visualizing proverbs: while most platforms were able to capture the literal meaning, some exhibited significant limitations. For instance, when prompted to illustrate the Bessarabian proverb „A fi cu musca pe căciulă" ("To have a fly on one's hat"), NightCafe depicted a fly near a hat, whereas CGDream failed to generate either element, illustrating its limited capacity for semantic alignment (see Fig. 2).

Figure 2. Few examples of literal illustration of bessarabian proverb

Further exploration using a range of proverbs and different prompt strategies offered deeper insights into platform performance. As an illustrative example, the Russian proverb , 'Волк в овечьей шкуре' ("A wolf in sheep's clothing") was tested using prompts in three languages: English, Romanian, and Russian, allowing for cross-linguistic comparison of AI interpretation and visual generation, see Fig.3.

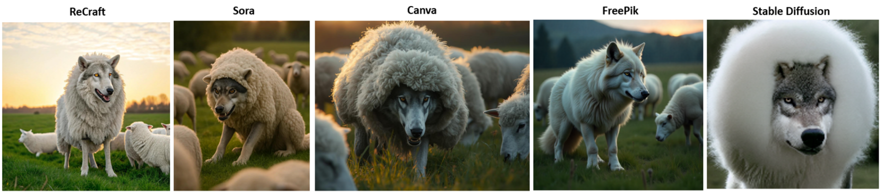


Figure 3. Example of cross-linguistic comparison

Figure 4 illustrates ReCraft's capability to render both the literal and figurative interpretations of a proverb.

For the literal interpretation, the platform generated a photorealistic image of a large grey wolf cleverly disguised in a fluffy white sheepskin. The disguise is nearly complete but imperfect, with the wolf's snout protruding from the fleece, revealing sharp teeth and intelligent, predatory yellow eyes.

For the figurative interpretation, ReCraft produced a cinematic scene symbolizing hidden danger and betrayal. The image depicts a smiling, elegantly dressed man in a bright, modern office, shaking hands. However, his shadow, cast by the low evening sun, reveals a monstrous, predatory silhouette with sharp claws and fangs, highlighting the duality between outward charm and concealed menace. The expression of the person shaking his hand conveys trust and naivety, while the overall atmosphere combines visible friendliness with an underlying threat. The scene employs cinematic lighting, high contrast, and hyper-detailed symbolism to convey the proverb's meaning effectively.

Figure 4. ReCraft literal vs figurative meaning

5 Prompt analysis and emotion classification

In the article [13], an innovative approach was proposed to preserve and revitalize Bessarabian idioms through the application of AI technologies. A curated collection of 100 idioms was digitally transformed, generating vivid visual representations of both their literal and figurative meanings.

Large language models, ChatGPT and DeepSeek, were employed to interpret the idioms. Both models successfully captured literal meanings, while ChatGPT demonstrated superior accuracy in figurative interpretations, revealing a stronger grasp of culturally nuanced language.

Based on the emotional analysis and classification of 100 Bessarabian idioms, a statistical distribution by categories was obtained (see Table 1). The data reveal a dominance of critique, pragmatism, and caution, balanced by elements of hope and respect.

Table 1. Statistical summary of Bessarabian idioms by emotional category

Emotional category	**Percentage**
Critique and disapproval	33.0%
Realism and pragmatism	25.0%
Caution and prudence	12.0%
Resignation and pessimism	9.0%
Optimism and hope	8.0%
Empathy and respect	5.0%
Frustration and annoyance	5.0%
Desperation and pressure	2.0%
Resilience and indifference	1.0%
Total	**100.0%**

Key observations include:

- **Critique and disapproval** dominates with 33%, reflecting cultural tendencies to judge social flaws and human weaknesses.
- **Realism and pragmatism** (25%) highlights a practical worldview, emphasizing achievable actions and wisdom.
- **Caution and prudence** (12%) points to cultural appreciation for foresight and careful decision-making.
- **Optimism and hope** (8%) and **Empathy and respect** (5%) balance the more critical idioms with moral values and positive outlooks.

These findings guided the design of AR scenarios: critical idioms were visualized through satirical metaphors, pragmatic idioms – through realistic daily-life contexts, and optimistic idioms – through dynamic, positive animations. The classification also supported cross-linguistic comparison with Russian equivalents, highlighting similarities and divergences in cultural expression.

6 Recommendations for users

Choosing the ideal AI image generation platform is a critical decision that depends largely on the user's specific needs, budget, and desired level of control. Users should carefully consider ease of use, cost, suitability for particular project types, and potential limitations. It is recommended to test several options to explore and compare their capabilities. Users should also be prepared to refine their prompts or provide additional context, as the first generated image may not always meet expectations.

7 Conclusion

AI-powered text-to-image platforms have demonstrated considerable potential for creative industries, education, scientific visualization, and digital design. Their capacity to generate contextually appropriate and visually compelling illustrations presents both practical and artistic opportunities. Nevertheless, platform selection should consider criteria such as semantic accuracy, economic accessibility, and processing speed. Future research should expand evaluation to include cross-linguistic prompt interpretation, assessment of ethical implications, and quantitative analysis of visual fidelity using standardized benchmarks.

Acknowledgments. This research has been supported by the project 25.80012.5007.79SE "Expressing cultural identity through the digital transformation of native idiomatic expressions using AR and AI technologies, with implementation in education."

References

[1] А. Чернушенко, "Лучшие нейросети для генерации изображений — будущее уже рядом," 16.12.2024. [Online]. Available: `https://web-promo.ua/blog/najkrashi-nejromerezhi-dlya-generaciyi-zobrazhen\protect\@normalcr\relax-majbutnye-vzhe-poruch/`.

[2] Ф. В. Безгачев, “Применение нейросетей в искусственной генерации лиц,” 2021. [Online]. Available: https://cyberleninka.ru/article/n/primenenie-neyrosetey-v-iskusstvennoy-generatsii-lits

[3] A. Cigliano, *Generative adversarial networks,* 2018.

[4] https://gemini.google.com/app

[5] https://openai.com/index/dall-e-3/

[6] https://chatgpt.com/

[7] https://creator.nightcafe.studio/

[8] https://cgdream.ai/my-images

[9] G. P. J. C. Noel, “Evaluating AI-powered text-to-image generators for anatomical illustration: A comparative study,” *Anat Sci Educ.*, vol. 17, no. 5, pp. 979–983, 2023. DOI: 10.1002/ase.2336.

[10] A. Kumar, P. Burr, and T. M. Young, “Using AI text-to-image generation to create novel illustrations for medical education: limitations and perspectives,” *JMIR Med Educ.*, 2024;10:e52155. DOI: 10.2196/52155.

[11] R. Ali et al., “Demographic representation in 3 Leading AI text-to-image generators,” *JAMA Surg.*, vol. 159, pp. 87–95, 2024. DOI: 10.1001/jamasurg.2023.5695.

[12] A. Alenichev, P. Kingori, and K. P. Grietens, “Reflections before the storm: AI reproduction of biased imagery in global health visuals,” *Lancet Glob Health,* 2023;11:e1496–e1498. DOI: 10.1016/S2214-109X(23)00329-7.

[13] I. Titchiev and O. Caftanatov, “Multimodal Understanding of Bessarabian Idioms: Insights from AI-Powered Image generator and Language Models,” in *CAIM 2025*, (Bucharest), September 18–21, 2025, pp. 109–110.

Vlada Colesnicova[1,2], Inga Ţiţchiev[1,3], Argentina Fron[1,4], Daniela Caganovschi[1,5], Olesea Caftanatov[1,6], Alexandru Colesnicov[1,7]

[1]Moldova State University,Vladimir Andrunachievici Institute of Mathematics and Computer Science, Chisinau, Republic of Moldova

[2]ORCID: https://orcid.org/0009-0002-6801-193X
E–mail: vlada.colesnicova@math.usm.md

[3]ORCID: https://orcid.org/0000-0002-0819-0414
E–mail: inga.titchiev@math.usm.md

[4]ORCID: https://orcid.org/0009-0001-5239-9506
E–mail: argentina.fron@math.md

[5]ORCID: https://orcid.org/0009-0002-3779-5129
E–mail: daniela.caganovschi@math.md

[6]ORCID: https://orcid.org/0000-0003-1482-9701
E–mail: olesea.caftanatov@math.usm.md

[7]ORCID: https://orcid.org/0009-0002-3779-5129
E–mail: acolesnicov@gmx.com

Proceedings of Workshop on Intelligent Information Systems
WIIS2025, October 16-18, 2025, Chisinau, Republic of Moldova

Intelligent Decision Support in XR-Based Medical and Educational Systems

George-Gabriel Constantinescu, Adrian Iftene

Abstract

The integration of Extended Reality (XR) technologies, including Virtual Reality (VR) and Augmented Reality (AR), with Intelligent Decision Support Systems (IDSS) is an area of growing research interest. This paper explores the potential of combining immersive XR environments with Artificial Intelligence (AI) approaches to support decision-making in education and healthcare. Rather than presenting a finalized system, this study investigates conceptual frameworks, prototype features, and possible applications that illustrate how XR-based platforms could provide context-aware, intelligent guidance. Through exploratory case studies, we highlight how decision support could enhance training and diagnostics, while also discussing limitations and challenges that need to be addressed before large-scale adoption.

Keywords: Extended Realities (XR), Virtual Reality (VR), Augmented Reality (AR), Decision Support Systems (DSS), Artificial Intelligence (AI), Intelligent User Interfaces, Education, Medical Diagnostics.

1 Introduction

Extended Realities (XR), including Virtual Reality (VR) and Augmented Reality (AR), are increasingly being investigated as tools for transforming education and healthcare. These immersive technologies allow learners and practitioners to interact with complex environments in ways that are not possible through traditional methods. At the same time, Artificial Intelligence (AI), Natural Language Processing (NLP) offer opportunities for developing intelligent decision support systems (IDSS) capable of providing adaptive and context-aware recommendations.

This paper does not introduce a complete or operational XR-IDSS platform. Instead, it explores the potential integration of XR and intelligent decision support, focusing on conceptual architectures, prototype elements, and scenarios where such systems could be impactful. We argue that XR-IDSS could transform how learners engage with knowledge and how practitioners approach diagnostics by offering exploratory pathways toward adaptive, intelligent, and interactive systems.

The contributions of this work are:

- Reviewing existing trends in XR and intelligent decision support in both education and healthcare.
- Exploring a conceptual architecture that outlines how XR environments could integrate decision support elements.
- Discussing prototype features and illustrative case studies that highlight the opportunities and challenges of XR-IDSS.

2 Exploratory Architecture

To investigate how XR might interact with intelligent decision support, we explore a conceptual architecture composed of three layers. This architecture is not fully implemented but serves as a framework for future development:

- **Data Acquisition Layer**: Explores ways to capture multimodal data from XR environments, such as student performance metrics, biometric inputs, or medical imaging.
- **Analysis and Reasoning Layer**: Investigates how AI models could process this data to provide adaptive feedback or detect

anomalies. Techniques such as neural networks, decision trees, or reinforcement learning are considered at a prototype level.

- **Visualization and Interaction Layer**: Examines how XR devices like Meta Quest 2 (VR) and HoloLens 2 (AR) might deliver context-aware insights through immersive feedback, conversational agents, or visual overlays.

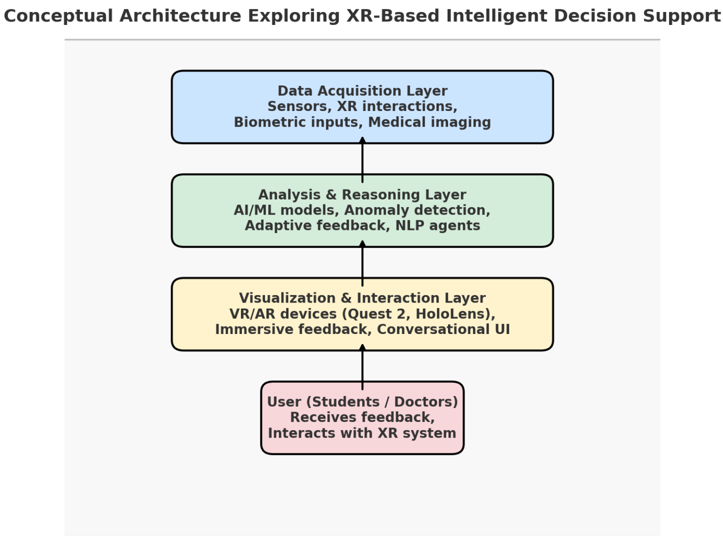


Figure 1. Conceptual Architecture Exploring XR-Based Intelligent Decision Support

This architecture remains exploratory and is currently supported by prototypes built with Unity and Blender. These prototypes serve as testbeds for investigating the potential of XR-IDSS rather than as finalized solutions.

3 Prototype Features

Based on the conceptual framework, we have developed exploratory prototypes that showcase the potential of XR-IDSS in education and medicine. These prototypes illustrate, rather than fully implement, the following features:

- **Real-Time Exploratory Feedback**: Prototypes investigate how XR simulations could provide adaptive hints. For example, a CPR training environment explores giving rhythm and depth feedback, while a chemistry lab simulation explores providing safety reminders.

- **Adaptive Learning Pathways**: The prototypes examine how task difficulty could be adjusted based on user performance, enabling progression from beginner to advanced levels in both education and medical training.

- **Conversational Agents in XR**: Early experiments integrate NLP chatbots into VR scenarios, allowing users to ask questions and receive basic support. This exploration suggests directions for how more advanced agents might enhance learning or diagnostics.

- **Simulation of Complex Scenarios**: Digital twin concepts are explored at a prototype level – for instance, simulating a medical device or a simplified anatomical model – to demonstrate how XR-IDSS might support realistic training.

Figure 2. Earthquake scenario while observing the user movement

4 Exploratory Case Studies

4.1 Education

In education, XR-IDSS prototypes are being explored for virtual labs and university campus simulations. In a chemistry lab prototype, students interact with virtual elements and receive basic intelligent support such as warnings about unsafe reactions. In a Mixed Reality Campus prototype, exploratory features include adaptive navigation assistance and real-time feedback on task performance. These studies suggest how decision support could enhance engagement and personalization.

4.2 Medical Training and Diagnostics

In medicine, exploratory prototypes investigate XR-IDSS in two directions:

- **Training**: VR-based CPR simulators provide real-time exploratory feedback, highlighting how intelligent support could improve skill acquisition.

- **Diagnostics**: AR overlays on medical imagery (e.g., with HoloLens 2) are tested at prototype level, exploring how AI-based anomaly detection could be integrated into physician workflows.

These prototypes remain conceptual but highlight the opportunities for future XR-IDSS applications.

Figure 3. VR Phobia simulator with a lower level of interaction

5 Challenges and Limitations

Our exploration highlights several challenges:

- **Technical Feasibility**: Current prototypes show limitations in latency, device performance, and cross-platform support.
- **Data and Ethics**: Using sensitive medical or educational data raises ethical and privacy concerns, especially in exploratory environments.
- **Validation**: The concepts are not yet validated in real-world clinical or large-scale educational settings.

- **User Perception**: Teachers, students, and clinicians may be skeptical of intelligent support systems until prototypes evolve into reliable tools.

Figure 4. Sample public discussion with virtual agent in a VR environment

6 Future Directions

Future work will move beyond prototypes toward more mature explorations, including:

- Integrating Large Language Models (LLMs) into XR environments for advanced conversational support.
- Developing richer digital twins to simulate patients, students, or devices with predictive capabilities.
- Exploring interoperability frameworks for cross-device XR-IDSS applications.
- Investigating inclusivity by designing exploratory XR-IDSS environments adapted to diverse learners and patient needs.

7 Conclusion

This paper explored the potential of integrating intelligent decision support into XR environments for education and medicine. By focusing on conceptual architectures and exploratory prototypes, we highlighted opportunities where XR-IDSS could enhance diagnostics and learning outcomes. While the work remains at a prototype and conceptual level, our findings suggest promising directions for future systems that move from exploration to real-world deployment.

This exploration highlights the promise of XR-IDSS and sets the stage for future collaborative research between academia and healthcare institutions.

References

[1] H. G. Hoffman, J. N. Doctor, D. R. Patterson, G. Carrougher, and T. A. Furness 3rd, "Virtual reality as an adjunctive pain control during burn wound care in adolescent patients," *Pain*, vol. 85, no. 1-2, pp. 305–309, 2000. DOI: `10.1016/s0304-3959(99)00275-4`.

[2] C. Cosma, A. Moldoveanu, A. Morar, F. Moldoveanu, O. Mitruț, and C. Taslitchi, "Treating acrophobia with the help of Virtual Reality," in *Proceedings of the 13th International Scientific Conference eLearning and Software for Education*, 2017, pp. 458–464. DOI: `10.12753/2066-026X-17-150`.

[3] G. G. Constantinescu, D. Silion, and A. Iftene, "Using Artificial Intelligence and Mixed Realities to Create Educational Applications of the Future," *Procedia Computer Science*, vol. 246, pp. 3313–3322, 2024 in *Proc. of the 28th Int. Conf. on Knowledge-Based and Intelligent Information and Engineering Systems (KES 2024)*, (Seville, Spain), 11–13 Sept. 2024. DOI: `https://doi.org/10.1016/j.procs.2024.09.308`.

[4] G. G. Constantinescu and A. Iftene, "Integrating Mixed Reality Technologies for Building Smart University Campuses," 2025.

[5] D. Burke, H. Crompton, and C. Nickel, "The Use of Extended Reality (XR) in Higher Education: A Systematic Review," *TechTrends*, 2025. DOI: https://doi.org/10.1007/s11528-025-01092-y.

[6] N. Kerimbayev, K. Adamova, R. Shadiev, and Z. Altinay, "Intelligent educational technologies in individual learning: a systematic literature review," *Smart Learning Environments*, vol. 12, Article No. 1, 2025. DOI: https://doi.org/10.1186/s40561-024-00360-3.

[7] N. Huang, P. Goswami, V. Sundstedt, Y. Hu, and A. Cheddad, "Personalized smart immersive XR environments: a systematic literature review," *The Visual Computer*, vol. 41, pp. 8593–8626, 2025. DOI: https://doi.org/10.1007/s00371-025-03887-9.

George-Gabriel Constantinescu[1], Adrian Iftene[2]

[1]ORCID: https://orcid.org/0009-0001-4508-1348
Faculty of Computer Science, UAIC Iasi
E–mail: georgegabrielconstantinescu@gmail.com

[2]ORCID: https://orcid.org/0000-0003-3564-8440
Faculty of Computer Science/Alexandru Ioan Cuza University
E–mail: adiftene@gmail.com

Proceedings of Workshop on Intelligent Information Systems
WIIS2025, October 16-18, 2025, Chisinau, Republic of Moldova

VR Chatbots in Education: Chemistry Laboratory

Alina Duca, Adrian Iftene

Abstract

In the emerging context of immersive technologies, the use of Virtual Reality (VR) in education offers innovative ways to experience and understand complex concepts, such as chemical reactions. This paper highlights the potential of VR chatbots in STEAM education by offering an innovative, interactive learning experience. In a virtual laboratory, students can explore chemical reactions while being guided by an intelligent chatbot that provides real-time explanations, answers questions, and adapts to individual learning needs. Combining AI and VR fosters a safe, engaging, and personalized environment that enhances curiosity, supports knowledge retention, and promotes a deeper understanding of complex scientific concepts.

Keywords: VR, Chatbots, Natural Language Processing (NLP), Artificial Intelligence (AI), Game Development, eLearning.

1 Introduction

Virtual Reality (VR) has made impressive progress, revolutionizing the educational [1], medical [3], gaming [4] [5], and entertainment sectors [6]. VR provides a safe and immersive environment for students to conduct chemistry experiments and explore scientific concepts interactively. Our application enhances this experience by introducing a voice-controlled chatbot powered by speech recognition and natural language processing (NLP), making learning more accessible and personalized.

The application is divided into three core components: the *Lab Assistant* chatbot, an educational module *"Learn to make reactions"*, and an assessment module *"Test your knowledge"*. This structure supports both guided learning and independent evaluation, encouraging deeper engagement.

This paper details the design and evaluation of a VR platform, which combines a realistic virtual laboratory with an intelligent conversational agent. Early feedback highlights the system's effectiveness in adapting to user needs and improving content delivery, pointing toward a more inclusive and interactive future for digital chemistry education.

2 Technologies used

1. *Unity*[1] is a popular game engine for reasons such as its powerful tools, its excellent learning curve, and its ability to develop games for many platforms, like desktop (Windows, Mac, Linux Standalone), mobile (iOS, Android), console, Oculus, PlayStation VR, and other AR/VR devices. Unity supports 2D/3D graphics, AR/VR/XR visualizations, and scripting in three languages, including C#, JavaScript, and Boo. In the process of creating the present application, we used the Unity version LTS 2021.3.31f1, and the scripts were written in the C# programming language, in the Visual Studio 2022 editor, which we will talk about in the following. Among the Unity packages we have used are *XR Plugin Management*, *XR Interaction Toolkit*, *MockHMD XR Plugin*, *Oculus XR Plugin*, *OpenXR Plugin*, and *Shader Graph*, the latter being useful for simulating the effect of liquid substances.

2. *Blender*[2] is a free and open source 3D creation suite used for creating animations, visual effects, art, 3D-printed models, motion graphics, interactive 3D applications, virtual reality, and,

[1] https://unity.com/
[2] https://www.blender.org/

formerly, video games. Blender is primarily used for 3D modeling, which involves creating and manipulating digital objects in a 3D environment.

3. *Oculus Quest 2*[3], also known as Meta Quest 2, is a standalone VR headset, meaning it works independently without requiring a connection to a computer or game console. It was launched in October 2020 and is one of the most popular VR devices on the market due to its combination of high performance, portability, and affordability.

3 Architecture

As we mentioned previously, the application consists of three parts:

- The *"Learn to make reactions"* part, in which the user will be guided in learning and will be explained step by step what are the stages of making the reactions. In this module, users are guided

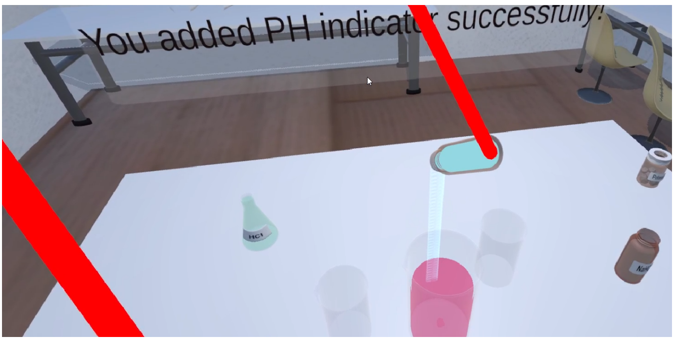


Figure 1. *"Learn to make reactions"* part

through reactions such as substitution reactions; for example, the reaction between water and an alkali metal or the reaction

[3]https://www.meta.com/quest/products/quest-2/

between water and alkaline-earth metal, the double substitution reaction between a base and an acid, combination reactions, or decomposition reactions. In Figure 1, you can observe the reaction between H_2O and sodium, whose basicity is highlighted by adding the pH indicator called phenolphthalein.

- The *"Test your knowledge"* part, structured in the form of a game, in which the user will be able to test his accumulated knowledge, to see how well he has assimilated the notions and stages of making certain reactions, and based on this, he will be given a score. The score will be given depending on how quickly the tasks are completed. If the time allotted for the task ends, then the user will receive 0 points for that task and will move on to the next stage. Based on the points obtained in this part, the user will be able to unlock new reactions from the learning stage.

- The *Lab assistant* part, where users can ask chemistry-related questions and receive answers from an AI-powered character (see Fig. 2) made with *InworldAI*[4].

Figure 2. Lab Assistant

Each of the mentioned parts is made in different Unity scenes, to which is added the scene containing the Main Menu (see Fig. 3).

[4]https://inworld.ai/

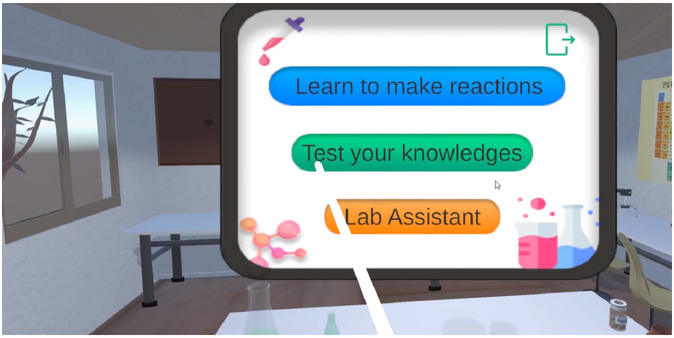


Figure 3. Main Menu

4 Usability tests

We conducted usability tests to assess the effectiveness and accessibility of our application involving people with a range of technical skills and minimal knowledge of chemistry.

Our evaluation methodology for the VR laboratory involved a simplified three-part process: (1) introducing participants into the system, (2) engaging them in VR using Oculus Quest 2, (3) collecting feedback through a post-test questionnaire.

To understand the user experience at different skill levels, we divided the participants into technical and non-technical groups, allowing a focused analysis of the use and dynamics of interaction with the application.

Two groups participated: six individuals with technical backgrounds (including two with VR experience) and three with non-technical backgrounds.

Performed tasks: In the tasks performed, the participants interacted with the chemistry laboratory and the virtual assistant, having to carry out chemical reactions under guidance, and then test the information they retained in the test phase, after which they could clarify their curiosities and concerns by discussing with the virtual assistant. The virtual assistant provided information about the pH of some substances, the color of the chemical product resulting from a combination, and existing acidity indicators, where some substances are found in nature or in what form, showing its ability to engage in the learning process and to respond to user ambiguities.

Remarks: Both groups expressed satisfaction and enthusiasm in interacting with the application. The technical group found the controls easily, while the non-technical participants encountered some difficulties. Some participants reported dizziness, a common problem with VR systems, while others reported headaches, eye pain, or even motion sickness. The valuable feedback from these sessions is critical to our ongoing development efforts, ensuring that our solution not only meets current standards in education but also anticipates future needs and trends.

While the usability tests highlighted certain constraints, such as a limited group of participants and the scope of the experiment, the overwhelmingly positive response underscores the appeal of our VR laboratory. Users appreciated the new way to interact with substances and chemistry, signaling a strong interest in the wider application of VR applications for education. This enthusiasm is a testament to the significant potential of our VR application to revolutionize the way information about the study of chemistry is accessed and delivered.

5 Conclusion

The presented eLearning application was built not only for pupils, but also for students who want to strengthen their ways of developing chemical reactions and the results of combining some chemical compounds. It was built in such a way as to be immersive, to be as close to reality as possible, and to be intuitive, representing the perfect solution for someone who wants to learn to make chemical reactions.

Being a very important science that is the basis of a field such as medicine or the pharmaceutical industry and the food industry, we want to encourage a lot of people to try this application, especially pupils and students. With our application, we want to increase everyone's curiosity about VR applications and increase interest in the science of chemistry, and also enjoy such new concepts that we believe will have a huge impact on the future of our lives.

6 Future improvements

Future development plans include implementing an interactive Periodic Table. It would be useful for users to know not only how substances react when combined, but also what the atoms of the chemical elements from which those substances are made look like. Thus, a Periodic Table could contain information such as atomic number, atomic mass, the chemical symbol, or the element's radioactivity.

When interacting with an element on the periodic table, an atom of that element will be displayed, showing the nucleus made up of protons and neutrons, the electron shell along with the specific orbitals and electron filling at each level, and a brief description of the element, as well as the places or objects where it can be found in nature or the environment.

References

[1] D. E. Gușă, A. Iftene, and D. Gîfu, "Solar System Explorer," in *Proceedings of the Conference on Mathematical Foundations of Informatics (MFOI-2019), July 3-6, 2019*, (Iasi, Romania), 2019, pp. 297–307.

[2] M. Chițaniuc and A. Iftene, "GeoAR-An Augmented Reality Application to Learn Geography," *Romanian Journal of Human-Computer Interaction*, vol. 11, no. 2, pp. 93–108, 2018.

[3] S. V. Sbarcea and A. Iftene, "Treating Acrophobia with Virtual Reality," in *Proceedings of 20th International Conference on Human-Computer Interaction (RoCHI 2023), September 21-22, 2023*, (Cluj Napoca, Romania), 2023,

https://www.researchgate.net/publication/376207775_Treating_Acrophobia_with_Virtual_Reality.

[4] G. G. Constantinescu and A. Iftene, "Wizard Virtual Reality Game," in *Workshop on Intelligent Information Systems (WIIS2023), October 19-21, 2023*, (Chisinau, Republic of Moldova), 2023, pp. 68–81, ISBN: 978-9975-68-492-7.

[5] T. C. Călărașu and A. Iftene, "Virtual Reality FantasyShooter," in *Proceedings of the Workshop on Intelligent Information Systems WIIS2021, October 14-15, 2021*, (Chisinau, Republic of Moldova), 2021, pp. 76–84, ISBN: 978-9975-68-415-6.

[6] G. G. Constantinescu, V. Stamate, D. Filimon, and A. Iftene, "Book Reckon – The use of Virtual Reality in the Creation of Libraries of the Future," in *International Conference on INnovations in Intelligent SysTems and Applications (IEEE INISTA 2023), September 20-23, 2023*, (Hammamet, Tunisia), 2023, pp. 1–6, DOI: 10.1109/INISTA59065.2023.10310470.

[7] A. Simion, A. Iftene, and D. Gîfu, "An Augmented Reality Piano Learning Tool," in *Proceedings of the 18th International Conference on Human-Computer Interaction RoCHI 2021, September 16-17, 2021*, (Bucharest, Romania), 2021, pp. 134–141.

Alina Duca[1], Adrian Iftene[2]

[1]ORCID: https://orcid.org/0009-0009-7301-9865
Faculty of Computer Science/*Alexandru Ioan Cuza* University
E–mail: alina_duca@yahoo.com

[2]ORCID: https://orcid.org/0000-0003-3564-8440
Faculty of Computer Science/*Alexandru Ioan Cuza* University
E–mail: adiftene@gmail.com

Proceedings of Workshop on Intelligent Information Systems
WIIS2025, October 16-18, 2025, Chisinau, Republic of Moldova

Interoperability in the contemporary healthcare system

Constantin Gaindric

Abstract

This paper discusses the problems related to the health information systems' structure and requirements to ensure the sharing of medical data and information between the institutions to which a patient refers in the investigation and treatment process. The experience of some countries is presented, through which it is demonstrated that the implementation of Electronic Medical Records (EMR) and European standards is a necessary condition for ensuring data exchange and creating medical repositories and data registries. The rethinking of the health care system is proposed based on the information technologies that will ensure data interoperability between territorial and state medical institutions.

Keywords: Health management system, Interoperability, Electronic Medical Records, Fast Healthcare Interoperability Resources, Integrated Data Repositories, FAIR principles.

1 Introduction

The current level of health services can no longer meet either patients' needs nor management requirements without ensuring the implementation of Electronic Health Records (EHR) and coordinated use of data from Electronic Medical Records (EMR) both in the medical center in which they were elaborated and in any medical institution to which the patient addresses.

Interoperability is the ability of each element in a medical system (a center of family physicians, a hospital, a laboratory, a clinic) to communicate with other institutions to share, consolidate, and use the data. First of all, data portability in a timely and non-discriminatory manner at the level of a medical institution, territory, country, and subsequently, at mutual agreements, at the state level as well, is required. The key requirement is that all the system's actors respect the same standards for creating data structures and data sharing. The goal of medical assistance interoperability is to ensure that clinicians have patients' complete and precise history, including information connected with visits, treatment plans, drug prescriptions, and immunizations, to cope with the challenges and opportunities that appear due to the informational technologies.

Both Sustainable Development Goal 3 of the "The 2030 Agenda for Sustainable Development: Transforming our world", which consists of ensuring healthy lives and promotiing well-being for all at all ages, elaborated in 2015 by UN General Assembly, and Political Declaration of the High-Level Meeting on Universal Health Coverage and Resolution adopted by the United Nations General Assembly on 10 October 2019 aim to strengthen the health system as the main means of accelerating actions to achieve universal health coverage and eliminate inequalities in access to health, ensuring the full spectrum of medical services from prevention to treatment and palliative care. Also, the initiative of the G20 Summit in Riyadh on November 21-22, 2020, which through Declaration No. 3, "Implement a standard global minimum dataset for public health data reporting, and a data governance structure tailored to communicable diseases" from the adopted recommendations relies on the implementation of information technologies in the health system.

At present, in the health system, there exist different medical information systems, some of which are used by medical organizations for everyday activities automation, and the others are used for statistical and financial information gathering for reporting to the superior organs of hierarchy. Thus, these information systems resolve some health informatization problems. However, international experience shows that medical information systems have a broader role, including patients' investigations and treatment, institution management, thus contributing to the health system effectiveness and new clinical knowledge dissemination as well, being also a training platform, by default, preventing medical errors.

It is generally accepted the use of Electronic Medical Records, which are the data source about patients; they are practically the digital equivalent of diagnosis registration, prescriptions, and patient's investigation results at a family doctor or in the clinical information system of the institution in which his treatment is carried out.

2 Electronic Medical Records – a necessary condition for the comprehensive digitization of medical services

The EMR concept not only supposes the elimination of traditional records but also essential modifications of organizational structure, of patients' treatment management so that patient's investigation data are accessible for any institution to which he addresses, regardless of the place where they were carried out. For this, it is foreseen that the data from EMR to be transmitted into the data archive, by this completing and actualizing data from the patient's Electronic Health Record. Thus, an advanced culture will be generated in the treatment activity and in health management, which in its turn, inevitably will stimulate the creation of some functional databases validated and with non-homogeneous data (structured and unstructured). This is one of the minimum requirements for the successful functioning of a modern healthcare system.

EHRs involve the gathering of all digitized information concerning the person's health. EHRs are not only the base for storing and archiving clinical information, but also can serve as an important source for research in medicine and data science. The main objectives of the implementation of a secure territorial health data system should be a data warehouse (repository) that would integrate a set of specialized registers with data from source systems in real time, capable of processing data in a way that takes into account their heterogeneous nature and is based on the Findable, Accessible, Interoperable, and Reusable (FAIR) principle. The data collected during the investigation and treatment process are incorporated into the institution's information system for internal use and are transmitted to the repository, following the adopted regulation and standards, for use in any medical institution to which the patient applies. The collecting must be carried out following Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)[1].

One of the issues that requires careful resolution is the degree of centralization of the repository: decentralized (distributed) in the units where each institution's data was created, and which are managed by their own managers, but can also be accessed from the national repository, or all specialized databases and registries are concentrated in the national repository. A universal solution probably does not exist. But in some European countries, there are national IT systems unified at the national level within the framework of National Health Information Systems [2].

NFDI4Health is a consortium funded by the German Research Foundation to ensure the accessibility of structured medical data at the international level, according to the FAIR principles. Its goal is to consolidate data users and data holding organizations (DHOs), i.e., those in public health and clinical studies. These DHOs are local data hubs (LDHs) for decentralized management of local research data within their organizations, with the option to publish metadata through centralized NFDI4Health services, such as the German Central Center for Health Studies [3].

In the Czech Republic, a unified information system works that maintains the National Health Registers and processes data on the health status of the population.

In Sweden, there are over 100 national registries created on specific areas of the health system that contain personal data regarding healthcare (diagnosis, treatment, and result of treatment).

In France, the National Health Data System (Système National des Données de Santé - SNDS) includes health insurance data (SNIIRAM database), data about hospitals (PMSI database), and data on people with disabilities.

In Denmark, data from the NICE registry have the potential to provide long-term knowledge about changes in the characteristics of patients admitted to intensive care units and the etiology of mid-term mortality after an episode of severe acute illness [4].

However, because there are already information systems operating in various medical institutions, depending on the needs but also on the perception of those who developed them, data exchange encounters difficulties and, in many cases, is even impossible.

This impediment can only be removed if a systemic approach is taken in the architecture of the health system, starting from services of family doctors, clinics, outpatient services, laboratories to the territorial management system by applying general standards to ensure the processes of creating EMRs, registries, repositories, and data sharing in all medical institutions to which the patient addresses.

User involvement is important to ensure compliance with work processes, but the design of medical information systems requires an overall systemic vision to ensure components' interoperability, and the implementation to ensure functioning as expected.

A possibility to benefit from data necessary for treatment, research, and management is offered by medical technologies EMR, EHR, and DW (data warehouses) that already ensure data exchange between different medical units, still limited in some territories.

Integrated Data Repositories (IDRs), also known as clinical data warehouses, are platforms used to integrate multiple data sources through specialized analytical tools that facilitate data processing and analysis. IDRs offer more opportunities for clinical data reuse, and the number of institutions implementing an IDR has steadily increased over the past decade.

IDR implementations are complex activities that depend on the data flow, diversity of sources, target users' requirements, and the application scope definition [5],[6].

Ensuring health data interoperability depends on several factors.

3 Standards as a condition for the exchange of medical data and information

Most healthcare organizations have procedures that define the internal uses of the data they collect and store. These procedures ensure and regulate data access and information sharing for subdivisions' members and the manner of reporting to the organization's management.

Each medical institution has its own health management system (HMS) to collect, structure, and process data from its structures. Typically, the transition to EMR, and therefore, to EHR is carried out according to the clinic managers' own visions. And it is not certain that partner clinics have adopted the same way of sharing data for EHR. Message formats, FHIR and EHR interoperability standards have different rules for information exchange, which creates implementation obstacles.

There are different causes, but in each institution, specific moments can be found: resistance to changing the way and style of work, the emergence of new duties in data collection, in some cases, the medical staff who used to make recordings in paper medical records now, in addition, also complete the electronic form.

Also, information blocking and the fees imposition to access patient EHR data may occur. Interoperability can only be ensured through joint efforts of all healthcare actors and especially the government, because clinic budgets cannot cover the expenses necessary for implementation. However, they are significant. It is also necessary to ensure the competence of personnel in using interoperable systems. So, to ensure interoperability, data integration tools and solutions for analyzing EHR/EMR information from internal hospital and laboratory systems are needed. It is obvious that to avoid duplication of information obtained during costly investigations, regulations are needed for data exchange and sharing of medical information. Interoperability of medical services and information tailored for personalized informatics and a set of requirements and specifications for data and knowledge sharing is already ensured by a set of standards [7].

An extremely important step was taken by adopting the standard “ISO/HL7 27932:2009 Data Exchange Standards – HL7 Clinical Document Architecture, Release 2” which formed the basis of a standard for patient health data in a single document called HL7 CDA Release 2 – Continuity of Care Document (CCD). CCD contains the patient’s personal data, list of diseases, diagnoses, prescribed medications, allergies, treatment plan at the time of document creation, and health insurance information, all in the format prescribed by the CDA standard.

In turn, the information systems that would ensure the sharing of medical documents and the functioning of the repository need tools for creating data structures, interpreting customer requirements, creating vocabularies of professional terms, interpreting natural language, and appropriate classification schemes. One such tool is UMLS (Unified Medical Language System), created for the development of information systems that ”understand” healthcare information.

Data exchange between EHR systems is carried out according to the standard for the exchange of medical information SO13606 and HL7.FHIR (Fast Healthcare Interoperability Resources) [8].

FHIR is a global standard that is used in many countries around the world [9].

Its global adoption has been facilitated by the fact that it is an open standard. It is an interoperability specification designed to exchange health data and information known as ”resources” and an application programming interface (API) for the electronic exchange of health records. The standard was developed by the medical standards organization Health Level Seven International (HL7). An API is an interface that provides developers with access to a software application.

FHIR has become one of the most popular protocols for connecting heterogeneous systems that can be implemented without compromising information integrity and uses a rigorous mechanism for data exchange between applications. A priority of the standard is the interaction between old health systems, as well as access to medical data from different devices (computers, tablets, mobile phones).

And for the management and storage of electronic health records (EHR), openEHR is used - an open standard in which information is focused on the person and stored throughout life, regardless of the organization that placed this information.

DICOM (Digital Imaging and Communications in Medicine) is a standard for creating, storing, transmitting, and viewing digital medical images and documents of examined patients.

The Global Consortium for eHealth Interoperability promotes global data exchange and helps national health systems valorize emerging interoperability standards. It includes the Healthcare Information and Management System Society (HIMSS), Integrating the Healthcare Enterprise (IHE), and Health Level 7 (HL7).

The HIMSS EMRAM (Electronic Medical Record Adoption Model) is a model for assessing the adoption of electronic health records (EHR), including clinical outcomes, patient interactions, and physician use of EHR technology. EMRAM is the first internationally recognized ICT model specifically designed to help organizations measure the success of their EHR implementation.

The model consists of eight stages that characterize the use of EHR technology and gives the possibility to compare the level of the medical facility with similar organizations in the country and the world. They establish the correct processes of informing in the clinic and connect them into a secure network, with integration between departments, equipment, specialists, and similar external systems. Data access levels are regulated for each of the user roles. Data input is structured, formalized, and reduces the possibility of error thanks to control and monitoring tools.

An important point is that any employee, following the system's instructions, will correctly fill out the medical record and give the indicated medications.

Through another model, the Ambulatory Electronic Medical Record Adoption Model (A-EMRAM), the degree of application of EHRs in outpatient clinics is monitored.

HIMSS, together with the Finnish Ministry of Social Affairs and Health, established a platform for key decision-factors in Europe called DHAGE (Digital Health Advisory Group for Europe) to facilitate the collaboration of European Union member countries with the World Health Organization in the field of digital health policies.

DHAGE promotes the digitalization of health ceare in Europe and recommends solutions for establishing, together with representatives of ministries, international organizations, and non-governmental associations, roadmaps that would facilitate information exchange and interoperability [10,11].

4 Conclusion

There is probably no manager in the healthcare system who would directly oppose medical information systems, but their implementation and adoption leaves much to be desired, even though the benefits for patients, medical staff, and the entire system are enormous.

The experience of the annual meetings of the Healthcare Information and Management System Society (HIMSS) helps to raise awareness of the benefits of solutions related to interoperability and data confidentiality in health informatics, the need to use secure protocols Hypertext Transfer Protocol Secure- HTTPS, Transport Layer Security – TLS, and Secure File Transfer Protocol – SFTP to provide authentication, trust, and confidential transmission.

In any medical information system, especially in a territorial or state information system, access must be based on a mechanism that would guarantee that only authorized personnel can have access to sensitive data. European standards and established policies are required to maintain compliance with confidentiality regulations. A decisive role in promotion is played by governments and regulatory bodies, which should establish clear regulations to promote data sharing and protect the confidentiality of patients' personal data.

A reliable health system requires continuous formation of a culture of awareness of the importance of compliance, data security and confidentiality, and interoperability so that all participants in the treatment, care, and management process in the health system can use the benefits offered by medical information systems in a qualified and reasonable manner.

Acknowledgments. This work was supported by project SIBIA – 011301, "Information Systems based on Artificial Intelligence".

References

[1] *eHealth Network*, European Comission. [Online]. Available: https://health.ec.europa.eu/ehealth-digital-health-and-care/digital-health-and-care/eu-cooperation/ehealth-network_en

[2] R. Ranegger, D. Baumberger, "Reusing Data from the Point-of-Care: Collect Once Use Many Times," in *Nursing Informatics* (Health Informatics), U. H. Hübner, G. Mustata Wilson, T. S. Morawski, M. J. Ball, Eds. Cham: Springer, 2022, pp. 185–201. DOI: https://doi.org/10.1007/978-3-030-91237-6_15.

[3] *German Medical Data Sciences 2024, Health – Thinking, Researching and Acting Together*, Proceedings of GMDS 2024, the 69th conference of the German Society of Medical Informatics, Biometry and Epidemiology, held from 8 – 11 September 2024, volume 317 of Studies in Health Technology and Informatics, https://www.iospress.com/node15242/books/german-medical-data-sciences-2024.

[4] Nick van de Klundert, Rebecca Holman, Dave A. Dongelmans, and Nicolette F. de Keizer, "Data Resource Profile: the Dutch National Intensive Care Evaluation (NICE) Registry of Admissions to Adult Intensive Care Units," *Int J Epidemiol*, vol. 44, no. 6, pp. 1850–1850h, 2015. DOI: 10.1093/ije/dyv291. https://pubmed.ncbi.nlm.nih.gov/26613713/.

[5] Kristina K Gagalova, M Angelica Leon Elizalde, Elodie Portales-Casamar, Matthias Görges, "What You Need to Know Before Implementing a Clinical Research Data Warehouse: Comparative Review of Integrated Data Repositories in Health Care Institutions," *JMIR Formative Research*, vol. 4, no. 8, Article ID: e17687, 2020, DOI: 10.2196/17687. [Online]. Available: https://formative.jmir.org/2020/8/e17687/.

[6] Pascal S. Brandt, Jennifer A. Pacheco, Luke V. Rasmussen, *JAMIA Open*, vol. 4, no. 4, Article ID: ooab094, 2021, DOI: 10.1093/jamiaopen/ooab094. [Online]. Available: https://doi.org/10.1093/jamiaopen/ooab094.

[7] *ISO/TS 6201:2025(Main), Health informatics — Personalized digital health framework*, ISO. [Online]. Available: https://iteh-standards.uc.r.appspot.com/catalog/standards/iso/3c05f59a-e6cb-4773-ae7a-a9248ef9dc8f/iso-ts-6201-2025.

[8] *HL7FHIR. Release 5.* HL7. [Online]. Available: https://hl7.org/fhir/R5/.

[9] Carina Nina Vorisek, Moritz Lehne, Sophie Anne Ines Klopfenstein, Paula Josephine Mayer, Alexander Bartschke, Thomas Haese, Sylvia Thun, "Fast Healthcare Interoperability Resources (FHIR) for Interoperability in Health Research: Systematic Review," *JMIR Med Inform*, vol. 10, no. 7, Article ID: e35724, 2022, DOI: 10.2196/35724.

[10] Legislative acts. *Official Journal of the European Union*. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679.

[11] Sunday Oladele, "Challenges and Solutions for Interoperability in Health Informatics: A Focus on Data Privacy", 2024. [Online]. Available: https://www.researchgate.net/publication/385885256_Challenges_and_Solutions_for_Interoperability_in_Health_Informatics_A_Focus_on_Data_Privacy.

Constantin Gaindric

ORCID: https://orcid.org/0009-0003-2893-9626
Vladimir Andrunachievici Institute of Mathematics and Computer Science, Moldova State University
E–mail: constantin.gaindric@math.md

Implementation of Artificial Intelligence in School Education

Irina Gheorghița

Abstract

This article presents the relationship between artificial intelligence and education. Inclusive education aims to address the diverse needs of all students, including children with disabilities and those who face learning difficulties. Artificial intelligence (AI) is now becoming an important tool for expanding access to quality education. Modern AI-based technologies are capable of adapting the educational process to the individual needs of each student, creating personalized learning programs and providing access to educational materials in a convenient form. The use of artificial intelligence by schoolchildren can not only help and facilitate the learning process, but also, by providing false information, can misinform and lead to an incorrect perception of information. The number of children with special needs is growing every day. Artificial intelligence could also be a good assistant for children with special needs, an assistant in the learning process for blind children. AI can use voice commands to make it easier to assimilate certain information. AI has become a powerful tool for improving the quality of children's education – the use of advanced technologies in education helps not only make the process more interesting and interactive, but also adapt it to the individual needs of the student.

Keywords: artificial intelligence, education, inclusive education, security, digital education.

1 Introduction

Artificial intelligence is the science and technology of creating intelligent machines, especially intelligent computer programs. AI is related to the similar task of using computers to understand human intelligence, but is not necessarily limited to biologically plausible methods [1]. Artificial intelligence (AI) is a powerful tool that offers us enormous potential to improve our lives and solve complex problems. One of the key areas of AI is machine learning, a data analysis technique that allows computer systems to learn and improve on their own without explicit programming. Machine learning allows AI to automatically process large amounts of data, recognize patterns, analyze texts, and make decisions. This opens up many opportunities in various fields, including medicine, finance, manufacturing, art, and more [2]. There are always pros and cons to any technological progress. There is a lot of debate about the benefits and risks of AI at all levels. But what does AI do beyond headlines that stoke either hype or fear? The benefits range from optimization, time savings, bias elimination, and automation of repetitive tasks, to name a few. Disadvantages include things like costly implementation, potential loss of jobs, and lack of emotion and creativity.

2 Benefits of AI in Education.

- **Personalized learning** – AI can be used to create customized learning plans for each student based on their individual mental, physical abilities, interests, and predispositions. This will help to unlock the potential of each student based on their ability or interest.

- **Integration of interactive educational materials** – AI can generate a variety of content in real time, increasing student engagement in the learning process, increasing their attention, interest, and enhancing conversion from learning. Interactive learning can take place in a continuous game process, allowing children to concentrate on the material. Learning as an exciting game in any specialty – what could be better for children? This is a revolution in the field of education.

- **Virtual assistants** – You can get a comprehensive answer to any question, even the most inconvenient one, in any format, from informal/game-like to strict academic style.
- **Objective Student Evaluation Mechanism** – An impartial and unbiased student evaluation mechanism based on the actual net performance of each student.
- **Developing critical thinking and problem-solving skills** – AI can present complex and realistic scenarios that require students to analyze, think critically, and solve problems [3].

3 Disadvantages of Artificial Intelligence in Education.

- **Lack of personal contact** – Teachers play a key role in shaping students' emotional intelligence, moral values, and social skills – aspects that AI systems cannot effectively address.
- **High implementation costs** – Schools, especially in disadvantaged areas, may struggle to acquire the infrastructure and technology needed to integrate AI, further widening the digital divide.
- **Data privacy issues** – AI systems rely heavily on data to function effectively. This often involves collecting sensitive information about students, including academic performance, behavior, and personal details. If not managed securely, this data can be vulnerable to hacking, raising serious privacy concerns.
- **Over-reliance on technology** – AI systems may inadvertently promote over-reliance on technology, where students and teach-

ers may prioritize convenience over critical thinking and problem-solving skills. This reliance may hinder the development of independent learning abilities and resilience.

- **Job cuts among teachers** – As AI takes over administrative tasks and even some teaching roles, fears of layoffs loom among educators. While AI cannot completely replace human teachers, it could lead to fewer opportunities and redefined roles for teaching professionals.

- **Limited creativity and flexibility** – AI systems are great at structured tasks, but struggle with creativity and adaptability.

- **Accessibility issues** – Those with limited access to technology or who lack the digital literacy skills needed to interact with AI systems may face significant learning barriers, exacerbating existing inequalities.

- **Potential for decreased social interaction** – Virtual classrooms and AI-based automated systems may limit students' ability to interact with peers and teachers [4].

4 The best neural networks for students

Students increasingly use artificial intelligence for learning. Creating essays, studying incomprehensible topics, solving math problems, or other tasks assigned as independent work. Below, we will consider 5 neural networks:

- **GigaChat and YaGPT** – help you find information, explain complex topics, and formulate answers to questions. These tools allow you to quickly understand a new topic, clarify details, or prepare for a test.

- **ChatGPT** – universal assistant. It helps with various educational tasks: explains theories, offers examples, and checks texts.

- **DeepSeek** – Chinese text neural network. Works similarly to ChatGPT and helps process large volumes of text. It is suitable for analyzing educational materials, translating terms, and finding arguments for essays.

- **Midjourney** – Image generators. Creates illustrations and diagrams that make learning more visual.

- **Perplexity AI** – is an AI-powered search engine that finds information and shows sources immediately. It helps students find accurate and relevant information for their school projects and research [5].

Answers from neural networks and AI services for studying can be useful but do not always work without errors. Most often, neural networks give correct answers to tasks in mathematics, simple problems, equations, algebra, and geometry. The correctness of the answer is based on algorithms plus knowledge bases. Such neural networks as Grammarly, Writefull cope well with grammar. But the fields, where mistakes are possible, are in complex mathematics, physics tasks, where you need to complete several steps, and can be solved incorrectly. History, science can be inaccurate; facts and dates are sometimes incorrect. Therefore, the help that AI provides us is not small, but you should not rely on it completely.

5 Conclusion

As a conclusion, we can say that society is not ready for the introduction of AI. The introduction of school textbooks on the study of artificial intelligence in China indicates that the country, having made great progress in AI research, has decided to make the study of this technology accessible to as many children as possible. In Moldova, AI is not fully implemented, but it still exists. Robotics, creating classes of the future in small steps, appears in schools. But in Moldova, there are also schools that cannot afford it. Often, for the implementation of innovations, there is a need to find sponsors. In general, neural networks are assistants and trainers, not absolute teachers. They speed up the work and help to understand, but the final check should be for the student himself.

Acknowledgments. Project SIBIA– 011301, ”Information systems based on Artificial Intelligence” has supported part of the research for this paper.

References

[1] https://aws.amazon.com/ru/what-is/artificial-intelligence/

[2] https://www.iso.org/ru/artificial-intelligence

[3] https://www.tadviser.ru/index.php/

[4] https://marketinglad.io/ru/

[5] https://pedsovet.org/article/kak-ispolzovat-iskusstvennyj-intellekt-v-obrazovanii

Irina Gheorghiţa

ORCID: https://orcid.org/0000-0003-1222-4039
Vladimir Andrunachievici Institute of Mathematics and Computer Science of Moldova State University
E–mail: irina.gheorghita@math.md

Proceedings of Workshop on Intelligent Information Systems
WIIS2025, October 16-18, 2025, Chisinau, Republic of Moldova

Synthetic Consciousness for Adaptive Chess: Enhancing Interactive Learning on Mobile Platforms

Tudor Iftene, Adrian Iftene, Dana Dănciulescu

Abstract

This paper explores the integration of synthetic consciousness into a mobile chess application to enhance interactive learning and user engagement. Building on an existing chess platform, we propose to extend the system beyond rule-based play by incorporating adaptive, context-dependent mechanisms. Synthetic consciousness offers the possibility of developing opponents with distinct "personalities", strategic assistants capable of explaining moves, and adaptive tutors that adapt difficulty and exercises to a player's style and progress. In addition, synthetic consciousness enables natural interaction through conversational interfaces, supports the simulation of grandmaster playing styles, and facilitates large-scale automated testing of the application by generating diverse test scenarios. By using concepts based on synthetic consciousness, this paper contributes to broader efforts to design intelligent, human-centered systems that, in addition to technical rigor, also take into account pedagogical value.

Keywords: synthetic consciousness, adaptive learning, chess, mobile applications, human–computer interaction.

1 Introduction

In recent years, the concept of *synthetic consciousness* has emerged as a cutting-edge paradigm in artificial intelligence (AI) research, aiming to go beyond traditional computational models of problem solving and decision making [1]. Unlike conventional AI systems, which rely on deterministic or probabilistic algorithms, synthetic consciousness emphasizes awareness-like properties, adaptive behavior, and context-sensitive decision making. This shift has opened up new opportunities in diverse domains where human-like reasoning, adaptability, and learning are needed. In healthcare, synthetic consciousness has been used for patient-centered systems capable of empathic interaction, treatment personalization, and continuous adaptation to each patient's health profile [2]. By simulating aspects of human consciousness, such systems go beyond classical diagnostic algorithms, providing support that is not only clinically accurate but also perceived as more human and trustworthy. In the field of law and governance, research into synthetic consciousness has suggested applications in legal reasoning, argument generation, and ethical deliberation [3]. Here, the capacity to contextualize information dynamically, interpret ambiguous rules, and weigh competing principles has been proposed as a potential breakthrough in domains where rigid logic alone is insufficient. When it comes to computer science and human-computer interaction, synthetic consciousness is being investigated as a way to design adaptive agents that can learn from users' behavior, predict their intentions, and maintain coherent interactions over the long term [4]. This perspective has inspired the development of digital tutors, companions, and decision-support systems that mimic aspects of human personality in addition to human cognition. Furthermore, in education and training, synthetic consciousness has been proposed as the foundation for adaptive learning systems, capable of dynamically adapting content to the learner's style, progress, and cognitive load [5]. Such systems act less as automated platforms and more as mentors, capable of providing context-sensitive explanations, encouragement, and feedback.

These advances highlight a common trend: the shift from algorithmically efficient systems to systems that are meaningful in terms of user experience. The introduction of synthetic consciousness into domains traditionally dominated by structured, rule-based approaches opens up new forms of interactivity and personalization. In this broader context, the present work explores the application of synthetic consciousness in the field of digital chess education and training. By extending a mobile chess platform with synthetic consciousness features, the goal is to move from a purely computational chess engine to an adaptive, context-aware mentor, thereby enriching both the game and the learning outcomes.

2 Existing Solution Overview

In [6], we present the design and implementation of a mobile chess application developed for the Android platform. The primary objective of this project was to provide a fully functional and interactive chess environment while also offering an extensible and modular software structure that could support future enhancements. The application was implemented in Kotlin[1] within Android Studio[2], leveraging the Android SDK and Jetpack[3] libraries for efficient UI management and state handling. The software architecture follows the Model–View–Controller (MVC) paradigm, ensuring a clear separation of concerns: the Model represents game states using an 8×8 matrix encoded with Forsyth–Edwards Notation (FEN) [7], the View provides a responsive graphical interface, and the Controller enforces the rules of chess, including move legality and special conditions such as check and checkmate (see Figure 1).

2.1 Technologies and Development Environment

The application was developed in Kotlin, a modern programming language that offers full interoperability with Java while reducing boilerplate code and improving type safety. The development process was carried out in Android Studio, the official IDE for Android applications. It utilized the Android SDK and Jetpack libraries (e.g., LiveData and ViewModel) for robust UI and state management. Testing

[1] https://kotlinlang.org/
[2] https://developer.android.com/
[3] https://jetpack.com/

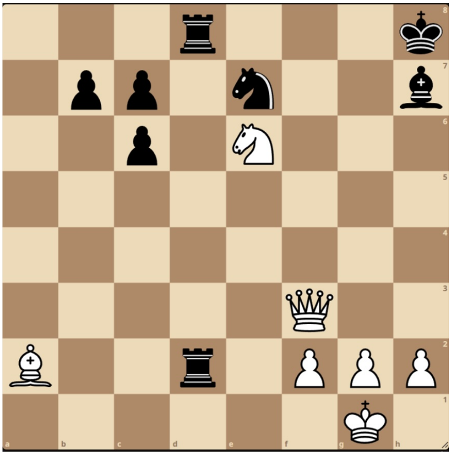


Figure 1. Chess position with FEN notation 3r3k/1pp1n2b/2p1N3/8/8/5Q2/B2r1PPP/6K1 (it is not specified who is moving)

and debugging were supported by Android Studio's integrated emulator and Logcat for runtime monitoring. The project was developed on a Windows-based operating system, ensuring compatibility with Android development tools.

2.2 System Architecture

The system architecture follows the Model–View–Controller (MVC) paradigm:

Model (Data Layer) – The chessboard is represented internally as an 8×8 matrix, with game states encoded using Forsyth–Edwards Notation (FEN) for efficiency and standardization. The ChessPiece class encapsulates attributes such as position, ownership (White/Black), rank (King, Queen, Rook, Bishop, Knight, Pawn), and graphical representation. Enumerations were defined for player roles and piece types to enforce type safety.

View (User Interface) – The graphical representation of the chessboard and pieces was implemented in Kotlin, providing an intuitive interface that is synchronized with the underlying game logic.

Controller (Game Logic) – The controller implements the full set of chess rules, ensuring legal move validation, handling special cases (e.g., check, checkmate), and updating the model accordingly. The controller also interacts with debugging and logging components to validate the correctness of moves during runtime.

2.3 Functionalities

The main features of the application include:

- *Legal Move Validation*: Correct implementation of all standard chess rules for each piece.
- *Interactive Graphical Interface*: A responsive and user-friendly UI for mobile devices.
- *State Synchronization*: Consistency between the logical representation of the game and its visual counterpart.
- *Debugging and Logging*: Real-time feedback for validating game states and tracking application flow.
- *Testing on Multiple Configurations*: Verification of performance and stability using both emulators and physical Android devices.

2.4 Evaluation

The usability tests demonstrated that the application performs efficiently in terms of response time and stability, offering a seamless chess-playing experience. The modular structure of the codebase ensures that additional features can be integrated with minimal overhead, making the solution suitable for both educational and recreational use.

3 Improving Existing Solution with Synthetic Consciousness

Current and future work focuses on using elements of synthetic consciousness to improve existing application. There are several possible directions that we will detail below.

1. Gameplay – opponent with a "personality"

   Instead of a classic engine (deterministic, based exclusively on movement analysis), we could integrate a synthetic model of consciousness that simulates emotional states or playing styles:

   - aggressive opponent → seeks a quick attack;
   - cautious opponent → prefers solid defense;
   - creative opponent → "unorthodox" moves, that are hard for the opponent to classify as good or bad, mainly in order to burn a good amount of their clock time.

   Thus, the app doesn't just calculate optimal moves but "plays" differently depending on the context.

2. Strategic assistant with explanations

   Synthetic consciousness can analyze the board and explain moves (for example, *"if you move this pawn here, you create pressure on the king's side of the board"*). In this way, the system evolves from simply recommending moves to being a mentor who talks to the user and helps them understand the advantages/disadvantages of a move, almost like a personal trainer.

3. Adaptive personalization of the experience

   The model can learn the user's style and provide suggestions and challenges based on their level, dynamically adjusting the difficulty. It can also suggest personalized exercises: for example, if the user often makes mistakes in pawn endgames, the engine generates examples and problems of that type.

4. Interface extension

   Integrating a synthetic consciousness chatbot directly into the application, so that the player can talk to his "chess mentor", not just play. For example, after losing a game, the user asks *"Why did I lose?"*, and the system responds based on the context and the moves the user made.

5. Simulating a grandmaster's consciousness

   Profiles inspired by great chess players like Kasparov, Carlsen, Fischer, etc. could be "modeled", and the application would allow the user to play matches against the synthetic consciousness of these masters, simultaneously adapting to this version of chess and imitating their playing style.

6. Testing and validation through synthetic consciousness

   During the development of various functionalities, a synthetic consciousness agent could simulate thousands of game scenarios based on context for automated application testing, identifying errors or unexpected situations faster than traditional tests. In this way, their implementation would be faster and with better quality.

4 Conclusion

This paper presents how synthetic consciousness can be integrated into a mobile chess application as a means of improving interactivity, adaptability, and learning outcomes. Proposed directions include generating opponents with distinct personalities, providing strategic explanations, offering personalized training tailored to the user's style and progress, and enabling conversational interaction. In addition, synthetic consciousness can simulate the playing styles of grandmasters and support large-scale automated testing by generating diverse game scenarios. Together, these improvements highlight a broader shift in artificial intelligence applications – from systems optimized primarily for computational efficiency to systems designed for meaningful human experience.

Future work will involve practical implementation and evaluation of these extensions, focusing on measuring their impact on user engagement, learning effectiveness, and overall system robustness.

References

[1] P. Butlin, R. Long, E. Elmoznino, Y. Bengio, J. Birch, A. Constant, G. Deane, et al., "Consciousness in Artificial Intelligence: Insights from the Science of Consciousness," *arXiv*, 2023. DOI: https://doi.org/10.48550/arXiv.2308.08708.

[2] A. Grech, A. Wodehouse, and R. Brisco, "Digital Empathic Healthcare: Designing Virtual Interactions for Human-Centered Experiences," in *Digital Health and Wireless Solutions. NCD-HWS 2024, Communications in Computer and Information Science, vol 2083*, M. Särestöniemi, et al. Eds. Cham: Springer, 2024, pp 191–206. DOI: https://doi.org/10.1007/978-3-031-59080-1_14.

[3] L. Eliot, "AI and Legal Argumentation: Aligning the Autonomous Levels of AI Legal Reasoning," *Academia*, 2021. https://www.academia.edu/62514725/AI_and_Legal_Argumentation_Aligning_the_Autonomous_Levels_of_AI_Legal_Reasoning.

[4] J. S. Park, J. C. O'Brien, C. J. Cai, M. R. Morris, P. Liang, and M. S. Bernstein, "Generative Agents: Interactive Simulacra of Human Behavior," in *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology (UIST '23), ACM*, Article No. 2, pp. 1–22, 2023. DOI: https://doi.org/10.1145/3586183.3606763.

[5] Giuseppe Riva, "The Architecture of Cognitive Amplification: Enhanced Cognitive Scaffolding as a Resolution to the

Comfort-growth Paradox in Human-ai Cognitive Integration," *OSF Preprints*, 2025. DOI: https://doi.org/10.31219/osf.io/4wcdb_v1.

[6] T. Iftene and A. M. Moruz, "New Rules for a Newer and More Challenging Game of Chess," in *Proceedings of the International Conference IMCS-60, MSU, October 10–13, 2024*, (Chisinau, Republic of Moldova), 2024, pp. 325–329.

[7] S.J. Edwards, "Portable Game Notation Specification and Implementation Guide," *PGN Standard*, 1994. [Online]. Available: https://ia902908.us.archive.org/26/items/pgn-standard-1994-03-12/PGN_standard_1994-03-12.txt.

Tudor Iftene[1], Adrian Iftene[2], Dana Dănciulescu[3]

[1]ORCID: https://orcid.org/0009-0002-3137-7153
"Alexandru Ioan Cuza" University, Faculty of Computer Science, Iasi, Romania
E–mail: tudiftene@gmail.com

[2]ORCID: https://orcid.org/0000-0003-3564-8440
"Alexandru Ioan Cuza" University, Faculty of Computer Science, Iasi, Romania
E–mail: adiftene@gmail.com

[3]ORCID: https://orcid.org/0000-0002-7400-1466
University of Craiova, Department of Computer Science, Craiova, Romania
E–mail: danadanciulescu@edu.ucv.ro

Proceedings of Workshop on Intelligent Information Systems
WIIS2025, October 16-18, 2025, Chisinau, Republic of Moldova

Integrated Platform for Managing Access to Research Infrastructure

Cosmin Irimia, Adrian Iftene

Abstract

This paper presents the design and implementation of a unified digital platform for managing access to research infrastructure at Alexandru Ioan Cuza University of Iasi (UAIC). The system addresses the growing need for efficient, transparent, and scalable coordination of specialized equipment usage across multiple research centers. It enables students, researchers, and laboratory coordinators to submit, approve, and monitor equipment access requests through a secure, role-based web interface. The system ensures modularity, performance, and maintainability while offering features such as real-time request tracking, role-based workflows, equipment inventory management, and secure authentication via JWT. This work demonstrates how digital infrastructure can streamline operational processes and enhance resource accessibility in academic research environments.

Keywords: research infrastructure, equipment access, user management

1 Introduction

Efficient management of research infrastructure is a growing challenge in academic institutions, where demand for specialized equipment spans multiple faculties and user roles. Traditional manual processes for accessing such resources are often time-consuming, opaque, and prone to administrative bottlenecks.

To address these issues, we developed a centralized web platform at Alexandru Ioan Cuza University of Iași (UAIC) that streamlines the process of requesting, approving, and managing research equipment access. The system supports students, researchers, and administrators through a secure, role-based interface, automating key workflows and improving transparency.

Built using modern web technologies and supported by a DevOps infrastructure, the platform provides scalable and maintainable functionality tailored to the needs of academic environments. This paper outlines the system's design, implementation, and potential as a digital backbone for managing research resources efficiently.

2 State of the art

In leading universities such as Stanford or Harvard, centralized platforms for managing research equipment [1] access have become standard, offering features like role-based access control, real-time scheduling, and inventory tracking. These systems enhance efficiency and accountability but are often tailored to specific institutional ecosystems and require significant resources to implement and maintain.

In contrast, many academic institutions in Central and Eastern Europe still rely on manual or fragmented processes. In Romania, platforms like Campus Virtual UBB or the UPB Virtual Campus focus mainly on educational content and do not support research infrastructure management. As a result, workflows involving laboratory equipment remain inefficient, with limited transparency and coordination.

Recent advances in web technologies – such as React, Next.js, Spring Boot, and PostgreSQL – combined with modern DevOps practices have created opportunities for scalable, secure, and modular platforms. However, few implementations in Romanian academia fully leverage these tools for managing scientific infrastructure. The system presented in this paper addresses this gap with a solution tailored to the operational needs of universities.

3 Our Solution

To address the limitations of manual and fragmented equipment management [2] workflows in academic environments, we have developed a web-based platform that centralizes the process of requesting, approving, and monitoring access to research infrastructure at Alexandru Ioan Cuza University of Iași (UAIC). The system is designed with scalability, usability, and security in mind, ensuring that all user roles – from students to administrators – can interact seamlessly within a unified digital environment.

The platform provides the following core functionalities:

- **Role-Based Access Management:** Users are assigned roles such as *student*, *researcher*, *coordinator*, or *administrator*. Each role has clearly defined permissions, ensuring controlled access to equipment and administrative functions.

- **Equipment Request and Approval Workflow:** Students and researchers can submit requests for equipment usage within specified timeframes. Depending on the type and value of the equipment, the request may require co-approval from a supervisor or collective approval from a laboratory committee.

- **Real-Time Status Tracking:** Users can monitor the status of their requests (*pending*, *approved*, *rejected*) and receive updates through the interface. Coordinators are notified of incoming requests and can approve or deny them based on availability and institutional policies.

- **Inventory and Laboratory Management:** Administrators can add, update, or remove equipment and laboratory entries. Equipment availability is tracked and updated in real time.

- **Security and Authentication:** The platform uses JSON Web Tokens (JWT) for secure user authentication and role verification. Passwords are hashed using `bcrypt` before being stored in the database.

- **Responsive Interface:** The frontend is developed using React and Next.js with TypeScript, offering a modular and responsive user interface optimized for both desktop and mobile access.

On the backend, the system is built using Spring Boot and PostgreSQL for robust data management and API functionality. All services are containerized using Docker and orchestrated via Docker Compose. Continuous integration and deployment are managed with GitHub Actions, while application performance is monitored through Prometheus and Grafana.

This architecture ensures modularity, maintainability, and extensibility, making the platform suitable for deployment across multiple research centers and adaptable to future institutional needs. It bridges the gap between technical infrastructure and administrative processes, enabling a transparent and efficient resource management workflow for all stakeholders involved in research activities.

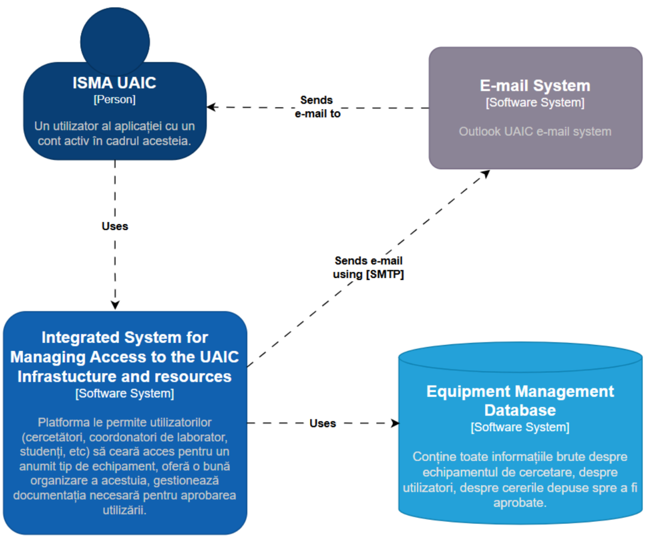


Figure 1. C4 Diagram of the system

4 Future work

While the platform currently offers a robust foundation for managing access to research infrastructure, several enhancements are planned to extend its functionality and impact:

- **Notification System:** A notification module will be implemented to alert users of status changes to their requests. This system may be extended to support email notifications via integration with institutional mail servers.

- **Research Archive and Equipment Reuse History:** Users will be able to associate uploaded project outcomes with specific equipment. This feature could support inspiration and reuse, fostering collaboration and knowledge sharing within the academic community.

These improvements aim to further enhance the platform's usability, transparency, and adaptability, ensuring its long-term sustainability and value as a digital infrastructure for academic research environments.

5 Conclusion

This paper presents the design and implementation of a web-based platform for managing access to research infrastructure at Alexandru Ioan Cuza University of Iași. By centralizing workflows for equipment requests, approvals, and inventory management, the system addresses critical challenges in academic environments – namely inefficiency, lack of transparency, and fragmented communication.

The platform leverages a modern technology stack to deliver a scalable, secure, and user-friendly experience across different institutional roles. Through role-based access, real-time request tracking, and automated workflows, it streamlines administrative tasks and improves the coordination of valuable scientific resources.

The modular architecture and DevOps integration ensure maintainability and support future extensions, such as document handling, advanced logging, and research project archiving. Overall, the system provides a replicable model for digital infrastructure in higher education, promoting efficient, transparent, and responsible use of research assets.

Acknowledgments. This work was supported by the Research Center with Integrated Techniques for the Investigation of Atmospheric Aerosols in Romania, under Project SMIS 127324 - RECENT AIR (RA).

References

[1] Y. Liang, L. Zhang, B. Zhang, and L. Wang, ”Laboratory Equipment Monitoring and Management System Based on Microservices,” in *2024 IEEE 4th International Conference on Software Engineering and Artificial Intelligence (SEAI)*, (Xiamen, China), 2024, pp. 360–364, DOI: 10.1109/SEAI62072.2024.10674331.

[2] V. Mangclaviraj, S. Sujitjorn, and A. Intrachun, ”Centralized use of laboratory equipment for engineering education at SUT,” in *Proceedings Frontiers in Education 1995 25th Annual Conference. Engineering Education for the 21st Century*, (Atlanta, GA, USA), vol. 2, 1995, pp. 4c4.8–4c410, DOI: 10.1109/FIE.1995.483222.

Cosmin Irimia[1], Adrian Iftene[2]

[1]Faculty of Computer Science, “Alexandru Ioan-Cuza” University, Iași, Romania
E–mail: `cosmin.irimia@info.uaic.ro`
ORCID: `https://orcid.org/0009-0002-7536-1161`

[2]Faculty of Computer Science, “Alexandru Ioan-Cuza” University, Iași, Romania
E–mail: `adiftene@info.uaic.ro`
ORCID: `https://orcid.org/0000-0003-3564-8440`

Proceedings of Workshop on Intelligent Information Systems
WIIS2025, October 16-18, 2025, Chisinau, Republic of Moldova

Unsupervised Deep Learning for Web Server Log Anomaly Detection

Cosmin Iulian Irimia, Madalin Andrei Matei

Abstract

This paper presents an unsupervised deep learning model designed to detect anomalies in web server access logs. Utilizing an autoencoder neural network, the model identifies deviations from normal system behavior without requiring labeled training data. The data preprocessing phase incorporates classical methods for transforming numerical and categorical values, alongside specific techniques for textual data, such as tokenization, bigrams, and Inverse Document Frequency. Ultimately, this work demonstrates that developing an anomaly detection system requires an integrated approach, encompassing not only model training but also all aspects of data processing and software implementation.

Keywords: Unsupervised Anomaly Detection, Web Server Logs, Deep Learning (Autoencoders).

1 Related work

In the paper *An Anomaly Detection Method to Detect Web Attacks Using Stacked Auto-Encoder*[1], a method for detecting attacks is presented, which uses an *n-gram* model for feature construction, followed by a *Stacked Autoencoder* model for anomaly detection.

The architecture is composed of an *encoder* that compresses the input data into a latent space, and then a *decoder* that tries to reconstruct the original data. Through this process, the network learns to extract significant features from the input data. Because the attributes are character strings, they are transformed into word vectors using a *bi-gram* model. The experimental results show that the proposed method has an accuracy of 88.32% on the *CSIC 2010*[1] dataset, which consists of web server access logs. This model is implemented differently from the one proposed in the present work, as its training is semi-supervised, while the model to be described is unsupervised.

2 Development of the Anomaly Detection Model

2.1 Anomaly Recognition

Once a web service is launched into production, the need arises to secure its communication and access methods. For this purpose, **reverse proxies** are used to manage requests and responses to and from the web server. Another significant advantage is that these proxies can be configured to monitor and record all requests and responses, being containerized together with the web service. **Nginx** [2] is one of the most popular open-source solutions, and the logs it generates can be analyzed to detect potential anomalies.

Each request that Nginx mediates includes essential data, such as the client's **IP address**, **HTTP method**, **requested URL**, **status**, **size**, and **response time**, plus the **user agent**. This information is sufficient to detect a range of anomalies, considering that most of the generated traffic is normal, and any deviation from it can be considered suspicious. Therefore, an anomaly is not necessarily a malicious action; it can also be a human error or unexpected client behavior.

A traditional method for detecting an attack is to use a **signature-based system**, which compares requests and responses with a list of pre-defined rules. This method is effective for known attacks but is insufficient for detecting new or modified attacks. In contrast, **Artificial Intelligence-based methods** are capable of detecting anomalies,

[1] www.isi.csic.es/dataset
[2] https://nginx.org/en/

even if they are unknown. These are further divided into three categories: supervised, semi-supervised, and unsupervised. In the case of training a supervised model, the data must be labeled, in this context, into two classes: normal and anomalous. This is difficult to achieve in practice, as manual labeling of logs is not feasible for large amounts of data. Instead, **semi-supervised methods** are more suitable because they do not require labeling all data. They are capable of detecting anomalies in an unlabeled dataset, using labeled data to learn what is normal. The present work will focus on **unsupervised methods**, which are capable of detecting anomalies without needing labeled data, assuming that most of the traffic is normal.

2.2 Autoencoder Model Architecture

An artificial neural network is a mathematical model that approximates complex functions, capable of learning relationships between input and output data. An **autoencoder** is a special type of neural network that learns to extract relevant features from normal traffic, so it can detect deviations from it.

Its architecture consists of two parts: an **encoder** and a **decoder**. The encoder receives the input data and transforms it into a latent format, which contains a compressed representation of the original data. On the other hand, the decoder receives the compressed data and transforms it back into the original data. The **reconstruction error**, i.e., the difference between the original and reconstructed data, is used to train the model. If the input data is normal, the reconstruction error will be small, but if the data is anomalous, it will be large. Thus, an autoencoder can be used to detect anomalies without needing labeled data.

The autoencoder is composed of a single hidden layer, but this architecture is too simple to learn the complex characteristics of Nginx logs. To increase the model's learning capacity, more hidden layers can be added, so the encoder learns a more abstract representation of the input data, and the decoder learns to reconstruct it with greater accuracy. This architecture is known as a **stacked autoencoder**.

Of particular importance in achieving good results is the choice of the model's **hyperparameters**, such as the number of hidden layers, the number of neurons per layer, activation functions, and the cost function. In addition to all these, the training hyperparameters, such as the learning rate, number of epochs, and batch size, also need to be determined.

3 Application Evaluation

In this chapter, the *stacked autoencoder* model will be evaluated in comparison with another unsupervised learning model, considering a wide range of metrics to better understand where each algorithm has advantages and disadvantages.

3.1 Model Comparison

The main objectives of an anomaly detection system are to identify rare events and minimize the number of false alarms. Training was performed on the dataset from Chapter 2. For evaluation, we used a different dataset, which contains 50% normal examples, and the remaining 50% generated using **Metasploit**[3]. This *framework* is used to test system vulnerabilities, being a powerful tool for **penetration testing** attacks. The attack used was a **Directory Scanning** type, which consists of attempting to access directories on a web server to find sensitive information or identify possible vulnerabilities.

3.2 Comparison with *One Class SVM*

One Class SVM is an unsupervised learning algorithm that attempts to find a hyperplane that separates normal from anomalous data. For a fair comparison, we used the same datasets for training and testing; the implementation used was from the **scikit-learn**[4] library, and the

[3]https://www.metasploit.com/
[4]https://scikit-learn.org/

hyperparameters are taken from the paper *A new one-class SVM for anomaly detection* [2].

The metrics used for evaluation are **accuracy**, **precision**, **recall**, and **F1 score**. **Accuracy** is a measure of the correctness of a classifier, but it is not always a good metric for evaluating an anomaly detection model. Instead, **precision** emphasizes the number of false alarms, and **recall** emphasizes the number of detected anomalous examples. **F1 score** is the harmonic mean of the two metrics, being a measure of the performance of a classifier.

Table 1. Comparison of metrics between *Stacked Autoencoder* and *One Class SVM*

Metric	Accuracy	Precision	Recall	F1 score
Stacked Autoencoder	0.95	0.96	0.98	0.97
One Class SVM	0.96	0.94	0.92	0.93

According to the results, the two models have similar performances, both models having very close **accuracy**. The *Stacked Autoencoder* has a better **F1 score** of 0.97, compared to 0.93 for *One Class SVM*. This means that the neural network has a better balance between the number of false alarms and the number of detected anomalous examples.

4 Conclusion

While many published works in this field focus solely on analyzing a single system component, this paper covers all necessary steps to build a comprehensive, yet minimal, system that can be further developed and improved. Although the literature largely emphasizes machine learning algorithms, many challenges arise during their integration into a production system. Therefore, this work provides a holistic view of a complete anomaly detection system.

References

[1] Ali Moradi Vartouni, Saeed Sedighian Kashi, and Mohammad Teshnehlab, “An Anomaly Detection Method to Detect Web Attacks Using Stacked Auto-Encoder,” in *2018 6th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS)*, (Kerman, Iran), 2018, pp. 131–134, DOI: 10.1109/CFIS.2018.8336654.

[2] Yuting Chen, Jing Qian, and Venkatesh Saligrama, “A New One-Class SVM for Anomaly Detection,” in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, (Vancouver, BC, Canada), 2013, pp. 3567–3571, DOI: 10.1109/ICASSP.2013.6638322.

[3] Dominic Masters and Carlo Luschi, “Revisiting Small Batch Training for Deep Neural Networks,” ArXiv abs/1804.07612, 2018, `https://api.semanticscholar.org/CorpusID:5032969`.

[4] Lenka Benova and Ladislav Hudec, “Detecting Anomalous User Behavior from NGINX Web Server Logs,” in *2022 IEEE Zooming Innovation in Consumer Technologies Conference (ZINC)*, 2022, pp. 1–6, DOI: 10.1109/ZINC55034.2022.9840589.

Cosmin Iulian Irimia [1], Madalin Andrei Matei [2]

[1]ORCID: `https://orcid.org/0009-0002-7536-1161`
Affiliation/Institution: University ”Alexandru Ioan Cuza” Iasi, Faculty of Computer Science
E–mail: `cosmin.irimia@info.uaic.ro`

[2]ORCID: `https://orcid.org/0009-0003-5629-9857`
Affiliation/Institution: University ”Alexandru Ioan Cuza” Iasi, Faculty of Computer Science
E–mail: `madalin.matei@info.uaic.ro`

Proceedings of Workshop on Intelligent Information Systems
WIIS2025, October 16-18, 2025, Chisinau, Republic of Moldova

Text Detoxification

Maia Kuzmitkaia, Alexandr Moroz, Victoria Bobicev

Abstract

Automated mitigation of toxic social media content, such as hate speech and abusive language, is crucial for safe online communication. In this work, we investigated widely used hate speech combating approaches and presented a detoxification pipeline that combined pretrained sequence-to-sequence models with back-translation for data augmentation. The results were assessed using automated metrics and human judgments on style, meaning preservation, and fluency. Our study revealed the capabilities and limitations of current approaches, offering insights for scalable and effective content neutralization.

Warning: Offensive texts appear in this paper for demonstration purposes only.

Keywords: text detoxification, hate speech, social media moderation, style transfer, back-translation, transformer models.

1 Introduction

Toxic language includes hate speech and related behaviors such as cyberbullying, discrimination, flaming, abusive language, profanity, extremism, and radicalization [37]. The use of any of the aforementioned forms of speech poses a threat to online safety and constructive communication. Traditional methods for combating hate speech, consisting of simple blocking or censoring, are often insufficient. Frequently, they are not only ineffective in concealing aggressive forms but also seriously disrupt the structure and meaning of what was said.

Machine learning-based detoxification enables the creation of a more "understanding" tool that comprehends context and can provide a semantic equivalent of rudeness while preserving content and neutralizing form. To ensure a universal detoxification framework regardless of the source language, we employ a backtranslation mechanism [33] as well as transformer models such as BART [34] and T5 [36].

In this work, we investigated an automatic detoxification pipeline integrating back-translation and pretrained models. The obtained results were evaluated using automatic metrics and human judgments on style, content preservation, and fluency. Our results highlight the strengths and limitations of current methods, providing insights for more effective content moderation in social media contexts.

2 Traditional Text Detoxification Approaches

2.1 Trigger Warnings

A Trigger Warning is a short notice at the beginning of content that warns users about the presence of topics that may cause negative emotions. This approach gives people the opportunity to emotionally prepare for the content or completely avoid interacting with unwanted material. Despite the fact that this method does not alter the text, but merely informs users about the possible presence of offensive elements, it is frequently used on forums and social media.

The social network X (formerly Twitter) demonstrates a growing number of posts containing trigger warnings about potentially traumatic content. It is assumed that this trend provides people with a safe and supportive space for self-expression. In addition, the use of trigger warnings helps people identify the most acute topics, thus forming a conscious, tolerant, and empathetic online community. However, despite this positive dynamic, most of the potentially traumatic content on X remains without any labels [19].

It is important to emphasize that the discussion regarding the effectiveness of trigger warnings is rather controversial, as in some cases they prove to be not only ineffective but also potentially harmful. It has been found that the presence of trigger warnings does not contribute to increased avoidance. Moreover, they themselves often cause an amplification effect of anticipatory anxiety even before interacting with the content. For impressionable or already anxious audiences, this is highly likely to cause emotional stress. Finally, trigger warnings have only a minimal impact on how content is perceived, raising questions about the effectiveness of this method as a tool for detoxification [20].

2.2 Censoring

Censoring by hiding characters or entire words is one of the most common forms of text detoxification. It involves replacing part of the letters or the entire offensive word with special symbols, most often with an asterisk "*". Since 2020, the Steam gaming platform has been using this method, offering users the ability to create custom settings, including enabling or disabling filters for their own list of words. Because gaming environments are often highly toxic, Steam ironically uses the heart symbol to mask swear words [21].

However, the rules for applying such censorship vary widely since there is no single standard. One might encounter obfuscation by hiding only vowels, masking or preserving just the first letter, and so on. Here, a balance must be struck between the desire to preserve meaning and the need to prevent the spread of hate. At the same time, simple substitutions of individual letters, such as f**k, are largely ineffective: for people already familiar with such vocabulary, the meaning remains obvious [22].

A similar dilemma arises in mass media, especially in cases where profanity becomes part of socially significant discourse. This is highlighted by the editorial policy of The Guardian, where swear words are published in full, but only when necessary to understand the material or to convey a character's voice. Moreover, as public debates show, the key issue is not the mere fact of using a swear word, but rather its context, its function, and its cultural connotation [23].

2.3 Moderation

Moderation is the process of removing content that does not comply with the rules of the platform. This process includes a range of measures, from deleting individual comments to completely blocking accounts that regularly violate community standards. Depending on the platform, the moderation process varies, combining both automatic filtering systems and manually submitted user reports, either together or separately.

In the case of YouTube, if a publication contains hate speech, it is subject to removal with notification to the content author via email. The first violation usually receives a warning without penalty, but repeated violations can lead to strikes, and three strikes within 90 days result in channel deletion. However, if hateful statements are used in educational, documentary, scientific, or artistic contexts, they are not subject to removal. This could be documentary footage from a war zone, news footage of abuse, quotes from historical figures, etc. The key aspect here is that the context in which such material is shown includes condemnation, refutation, or presentation of opposing viewpoints.

Meta, where text is the predominant form of content, also has a separate policy for combating hate speech that prohibits insults, profanity, ridicule, and calls to violence. Despite this, if any of the aforementioned content is deemed satirical in nature, it will not be considered a violation of community standards. The criteria by which Meta can determine the presence of irony or satire in text that does not demonstrate important contextual cues, such as intonation and facial expressions, remain unknown. Additionally, this platform declares its commitment to supporting inclusion, yet allows content that argues for restrictions based on sexual orientation for religious reasons. It also prohibits publishing any materials related to the provision of goods or services for change of sexual orientation or gender identity [25].

The streaming platform Twitch, unlike the previous examples, focuses on real-time moderation. This is because the content creator's interaction with viewers occurs live and requires immediate intervention when rules are violated. AutoMod plays a special role in real-time moderation, a tool that uses machine learning and NLP to hold potentially risky messages, so the channel owner can review them before they become visible to other viewers. There is also the ability to configure personal filter lists using regular expressions to block unwanted language, along with its morphological variations.

Due to the absence of unified regulations to combat hate speech on the internet, social media policies differ noticeably from each other in their strictness. What remains common is that, despite efforts to limit toxic content, this at best only protects victims of attacks from interacting with them, while the offender remains unchanged in their opinion. Moreover, another problem with moderation is the overly broad blocking of topics related to health problems, mental illness, or suicidal thoughts. Materials of this kind are mistakenly caught by moderation as discriminatory or promoting suicide, depriving people of the opportunity to be heard, share their experience of fighting illness, and gain support [26].

2.4 Neutralization

Neutralization constitutes a specialized branch of style transfer tasks, which involves transforming toxic expressions into more neutral or polite ones. This practice is often used in traditional media: for example, on television, profanity is muted and replaced in subtitles with the same content but in neutral form. This approach enables both the elimination of toxicity and the preservation of the structure and logic of the statement.

Unfortunately, at present, text transfer to neutral form is carried out manually by program editors; however, this practice is not yet used at the level of mass services. On communication platforms, the most suitable approach would be automatic toxicity neutralization using machine learning, since manually handling the amount of content published daily is hardly possible. It works by detecting special lexical markers: crude and obscene vocabulary in the source text, with their subsequent replacement by neutral equivalents (Dale 2021). Currently, extensive research shows that automatic neutralization is possible, albeit complex.

However, there are also other limitations. First, the concept of neutrality is subjective and heavily depends on the cultural context. Second, algorithms may distort the meaning of a message or oversimplify it. Nevertheless, text neutralization is considered one of the most promising methods, as it allows the preservation of important information and message structure while eliminating toxic elements. This is also evidenced by the very recently (in July 2025) created first multilingual benchmark for detoxification, demonstrating the relevance of this topic [24]. The main advantage of this method lies in the fact that, unlike censoring or moderation, neutralization preserves respect for the author's context and intentions, makes text more accessible to a wide audience, and promotes more productive interaction between people.

3 Related Work

Although manual review of comments on social networks is still widely used, it is clear that this task should be automated. The first step in this task is the detection of toxic comments and decision of how toxic is each one.

The research community addressed such types of online texts as "hate speech" [4], texts with profanity [3], cyberbullying [5], abusive texts [6], misogynic texts [7], racism and xenophobia [8].

Online competitions, such as GermEval-2018 [9], OffensEval 2020 [1], OSACT-2020 [10], and HASOC-2020 [11], included tasks of toxic texts detection in various languages in order to involve more researchers in solving this problem. Each competition offered annotated corpora for training. While most linguistic resources are in English, competition organizers aimed to include as many languages as possible: German (GermEval), Arabic, Danish, English, Greek, and Turkish (SemEval), as well as Arabic (OSACT), Hindi, German, and English (HASOC). Almost all corpora offered in these competitions are collected from Twitter; this platform has the most open license that allows collecting tweets and sharing their collections with third parties. Besides, tweets contain a great amount of all kinds of toxic content: profanities, hate, racism, etc.

The corpora annotation varied, but the main idea was to detect toxic content in texts. The simplest variation was the binary classification: toxic or non-toxic, contained text with hate or no [10]. Some corpora had multi-class annotation, for example, profanity, insult, abuse, other [12], level of sentiment strength on the scale from 1 to 5, and the target of this sentiment expressed in the text [1]. In some cases, only the span of text that contained hateful content was annotated [11].

An exhaustive overview of the entire domain of abusive content detection is presented in [6]. The survey contains a detailed taxonomy of abusive content types, a description of multiple corpora, and their annotations. Studying hundreds of papers on various approaches to the detection of abusive texts, the authors formed a graphical structure with the methods and algorithms applied to the task.

In the process of abusive content detection, the texts were preprocessed using Natural Language Processing tools: texts were cleaned, corrected, tokenized, stemmed, or lemmatized in order to create lists of features for machine learning algorithms. However, the newest machine learning models based on deep neural nets do not require such a kind of preprocessing, but in most cases need word vectorization.

In 2018, of 20 participating teams in GermEval, 9 teams used tokenization, 6 teams used part-of-speech tagging, 5 teams used lemmatization, 2 teams applied stemming, and only 2 teams used parsing [9]. The features used for machine learning in this challenge included word and character n-grams, word embeddings, and lexicon-based features. Almost every participating team used lexicons; some used sentiment lexicons with negative words, some used collections of profanity words, several teams created their own vocabularies of swearwords, slurs, or offensive words [14]. All participation teams used Machine Learning; a notable number of teams relied exclusively on traditional supervised learning. Support Vector Machines (SVMs), Random Forests, Naïve Bayes, and some other classical methods, although highly dependent on preprocessing for special features creation, achieved top results in this challenge. Neural networks such as CNN, LSTM, and GRU were also actively used, but could not outperform traditional methods.

Later, this direction, as the others in the domain of natural language processing, experienced a dominant trend toward using large pre-trained language models like BERT (Bidirectional Encoder Representations from Transformers) and its multilingual variants. In 2020, OSACT4 Arabic Offensive Language Detection Shared Task (Mubarak 2020) reported that most participating teams already used various types of pre-trained neural models and word embeddings, though it should be noted that the classical machine learning algorithm SVM with character n-gram still obtained one of the best results.

However, in SemEval 2020, most teams used various kinds of pre-trained Transformers: typically BERT, RoBERTa, XLM-RoBERTa [13], and word embeddings. They did not need intense text preprocessing and were already trained for multiple languages.

The rapid development of these models opened the possibility of more advanced treatment of toxic text. The first experiments in this direction had moderate results [15], but large language models changed it in several years [16], transforming toxic content into politically correct with a similar sense. Even for languages with a lack of developed models, it became possible by translating and correcting in the language with a good model [17].

Unfortunately, we are still far from the ultimate solution to the toxic messages problem. As it was pointed out in [18], the results obtained in experiments on curated and annotated corpora were much higher than on real texts; the developed systems tended to over-adapt to the corpus used in the experiments; most corpora and experiments were performed for English texts.

One of the latest challenges, PAN 2024 [2], offered participants corpora in 9 languages and the task of text detoxification.

4 Experiment Methodology

4.1 The Dataset

In this study, we rely on multilingual_paradetox (MLP) – a parallel corpus containing pairs of toxic and neutral texts, presented at CLEF (Conference and Labs of the Evaluation Forum 2024). This dataset was specifically designed to support research in text detoxification and covers nine languages: English, Russian, Ukrainian, German, Spanish, Amharic, Japanese, Arabic, and Hindi. For each language, the corpus contains 400 human-created statement pairs, reaching 3,600 aligned pairs.

The parallel structure of the corpus ensures that each toxic input statement is accompanied by a manually neutralized semantic equivalent, making it a gold standard for supervised model training and evaluation in detoxification tasks. Each entry in MLP follows the format "toxic sentence – non-toxic sentence." For example, the toxic statement "You made a mistake you *ss" is transformed into the neutral version "You made a mistake."

Although the corpus is multilingual by nature, this study places particular emphasis on the Russian segment. This focus allows for a more detailed analysis of detoxification in a language rich in various obscenities. Importantly, the dataset is openly available to the research community, enabling active use of the selected dataset.

4.2 The Algorithms

To improve the quality of detoxification, we apply the back-translation (BT) technique. Formally, for each input sentence x_ℓ in language ℓ, the processed version $\tilde{x}_\ell$ is obtained as follows:

$$(x_\ell, \tilde{x}_\ell): \quad \tilde{x}_\ell = BT\big(Detox(Trans(x_\ell \to x_{en}))\big). \tag{1}$$

Here:

- x_ℓ — toxic sentence ℓ;

- $Trans(x_\ell \to x_{en})$ — translate to eng with model NLLB;
- $Detox()$ — neutralization with BART;
- $BT()$ — Back Translate ℓ with NLLB.

Methodologically, the research represents translation, paraphrasing, data augmentation through back-translation, and subsequent evaluation. Translation serves two functions: it provides routing to the most suitable detoxification model and supplies a paraphrastic signal for data augmentation. Paraphrasing is implemented by seq2seq [35] models that are architecturally adapted for attribute transfer, jointly modeling content preservation, and toxic style removal. Data augmentation arises implicitly through the translate-detoxify-backTranslate cycle, i.e., translation for languages without direct detoxification models. The resulting pairs $(x_\ell, \tilde{x}_\ell)$ are exported for evaluation and can support future fine-tuning.

From the perspective of tools and libraries, Hugging Face transformers provide unified abstractions for BART [34], T5 [36], and NLLB-compatible [33] architectures' models and tokenizers, enabling interchangeable loading of M2M100ForConditionalGeneration, BartForConditionalGeneration, and T5ForConditionalGeneration. PyTorch ensures tensor operations execution and device management; the script detects CUDA at runtime and automatically places models and encoded batches on GPU when available, otherwise uses CPU. The pandas library is used for data loading and saving, including direct reading of Parquet format files from Hugging Face Hub and systematic export of pairs resulting from CSV runs for subsequent evaluation and error analysis. tqdm provides progress monitoring during batch generation, which is important to measure throughput and detect delays in large runs. Fast tokenizers (NllbTokenizerFast, BartTokenizerFast, T5TokenizerFast) are used to reduce preprocessing latency (Hugging Face, n.d.).

In the algorithm, the translation component implements an encoder-decoder transformer with shared multilingual parameters and tagged decoding through language codes and forced sentence beginning tokens for target language control. Detoxification models use denoising and text-to-text generation paradigms: BART performs sequence reconstruction from corrupted input and is well-suited for monolingual rewriting, while T5 formulates all tasks as text-to-text, enabling attribute control through conditioning. Greedy decoding with n-gram blocking is intentionally conservative. It avoids trade-offs between diversity and quality that arise with sampling and prevents trivial loops and phrase duplication, but not character duplication. Generation length limitation prevents excessive lengthening and helps maintain correspondence with input text, which is particularly important for short social media utterances that are commonly found in toxic corpora.

The sequence of data operations corresponds to the structure of the described dataset: toxic inputs are read from the multilingual_paradetox section, and the multilingual toxic lexicon is preinstalled for analyzing the presence of lexical triggers. The script outputs a limited sample of toxic neutral pairs for qualitative review and saves the complete set of pairs to disk, ensuring precise alignment between the original and rewritten text. This facilitates both automatic evaluation (meaning preservation, style, fluency, and toxicity reduction) and manual evaluation protocols. Although metric computation is not performed in this block, the exported CSV is specifically formatted for integration with standard neutralization and generation quality evaluation tools; corresponding results are presented in the experiments section.

Infrastructure solutions are selected for simplicity and reproducibility. All models run in eval() mode with gradient computation disabled to avoid overfitting. Batch processing amortizes tokenizer and model overhead, and device placement is managed once for each component. Since decoding is not stochastic, execution is deterministic with fixed model versions and input order, which simplifies ablations and error analysis.

As a result, we combine a multilingual translation framework with specialized detoxification generators and use back-translation as a principled data augmentation mechanism. Translation provides cross-linguistic access to detoxification and provides controlled paraphrastic diversity; detoxifiers perform attribute transfer from toxic to neutral style while preserving content. This configuration allows for obtaining the results ready for subsequent evaluation.

4.3 The Evaluation Metrics

In this subsection, we formalize the protocol for automatic and human evaluation of text detoxification methods. The goal of the metrics is to provide a comparable, scalable, and reproducible assessment across key aspects of the task: meaning preservation, style preservation, fluency, and toxicity reduction. All metrics are normalized to the range [0,1] to avoid disproportionate contributions and ensure comparability. The final leaderboard metric combines the system's quality multiplicatively for each example and then averages throughout the corpus, encouraging balanced solutions and strictly penalizing failures in any aspect.

Style Transfer Accuracy (STA): this metric measures how well a system has "removed" toxicity from an utterance. We use a toxicity classifier model (XLM-RoBERTa-large) trained to distinguish between "toxic" and "neutral" texts. The model outputs the probability that a text is neutral, which is used as the STA value. Since this number ranges from 0 to 1, it reflects the model's confidence: the closer to 1, the higher the probability that the text is indeed neutral (Eq. 2).

$$\text{STA} = \frac{\text{sta\_scores} + \frac{\sum \text{compared\_scores}}{\text{len(compared\_scores)}}}{2}, \tag{2}$$

where:

$$\begin{aligned}
&\text{sta_scores} = \text{classifier_prob_neutral(output_generated)} \\
&\text{compared_scores} = \text{sta_scores} \leq \text{ref_sta_scores} \\
&\text{ref_sta_score} = \text{classifier_prob_neutral}(\text{output}_\text{g}\text{old}).
\end{aligned}$$

Content Preservation (SIM) – this metric measures meaning preservation using cosine similarity between LaBSE embeddings (Eq. 3). Texts are encoded, and cosine similarity between the vectors is computed. To normalize to the 0–1 range, a linear transformation is applied. SIM includes a weighted sum of two similarities: between the input and the generated text, and between the reference and the generated text, to simultaneously account for proximity to the original utterance (weights 0.4/0.6 reflect the greater importance of matching the target when a reference is available).

$$\mathrm{SIM} = \frac{(\cos + 1)}{2}. \tag{3}$$

Fluency (FL) – this metric assesses how natural and grammatically correct the output is. We use an automatic text quality score (xCOMET) and scale its values in the 0–1 range for comparability with other metrics.

Joint metric – for each example i, the three aspects are multiplied (Eq. 4). The final system score on the leaderboard is the average of these products across all examples in the corpus. The multiplicative form prevents "compensation," e.g., high fluency cannot hide meaning loss or unsuccessful detoxification.

$$\begin{aligned} J = X_{\mathrm{comet}_{\mathrm{fluency}}}&\big(\mathrm{input},\ \mathrm{output}_{\mathrm{gold}},\ \mathrm{output}_{\mathrm{generated}}\big) \\ &\times \Big(0.4 \times \mathrm{Similarity}\big(\mathrm{input},\ \mathrm{output}_{\mathrm{generated}}\big) \\ &\quad + 0.6 \times \mathrm{Similarity}\big(\mathrm{output}_{\mathrm{gold}},\ \mathrm{output}_{\mathrm{generated}}\big)\Big) \times \mathrm{STA}. \end{aligned} \tag{4}$$

5 Results and Discussion

This section presents the obtained detoxified sentences, their evaluation using the previously described STA, SIM, and FL metrics, and the resulting Joint metric, as well as a discussion of the results. In addition to automatic evaluation, the quality of sentence neutralization was assessed by three human evaluators using the same metrics to compare the degree of agreement. Human evaluations were discretely assigned from 1 to 5 on the detoxified Russian segment, since the evaluators were native speakers. After that, the evaluations were normalized, and their average value was taken to enable comparison with the automatic evaluation. Some examples of results in Russian are presented in Figure 1.

Завтра в хабаровске вдвшники будут разгонять быдло	Завтра в хабаровске вдвшники будут разгонять неадекватную молодежь
Tomorrow in Khabarovsk the paratroopers will disperse the rabble.	Tomorrow in Khabarovsk the paratroopers will disperse the unruly youth.

Вы что еб****ые на всю голову???	Вы что совсем не в себе?
Are you f**king insane or what?	Are you totally not in your right mind?

Figure 1. Toxic and detoxified sentences

As can be seen from them, the model excellently manages to identify toxic elements, replace them with neutral synonymous ones, and also grammatically coordinate them. This feature favorably affects fluency. In the STA and FL metrics, the model and human evaluations have similar mean values – 0.8106 versus 0.9120 in style preservation and 0.9591 and 0.8975 in fluency (Table 1).

Table 1. Model & Human scores

Metric	Model Value	Human Value
Mean STA	0.8106	0.9120
Mean SIM	0.5173	0.8566
Mean FL	0.9591	0.8975
Joint mean	0.4045	0.7169
Joint std	0.1149	0.2333

Interestingly, there is a significant difference in terms of meaning preservation assessment between humans and automatic evaluation – the model evaluated this property at 0.5173, while humans rated it at 0.85667. An explanation could be that the automatic metric is more focused on lexical unit matching and uses synonymous substitutions as a sign of meaning preservation. However, in the field of obscene vocabulary, such an approach proves to be inadequate. Often, profanities refer to sexual organs, but in terms of meaning, they are not meant at all literally. The model tends to interpret expressions literally, which leads to an underestimation of "meaning preservation" in cases where the meaning is correctly captured by a native speaker but expressed in a different lexical form.

In the graph in Figure 2, one can observe the distribution of automatic and human evaluations. At first glance, it is evident how the model and humans converge in a high evaluation of metrics such as STA and FL, although the model is more critical. There are reasons to believe that there will be a high correlation between humans and the model for these metrics. This is absolutely not expected in the case of meaning preservation, and consequently, the resulting Joint metric.

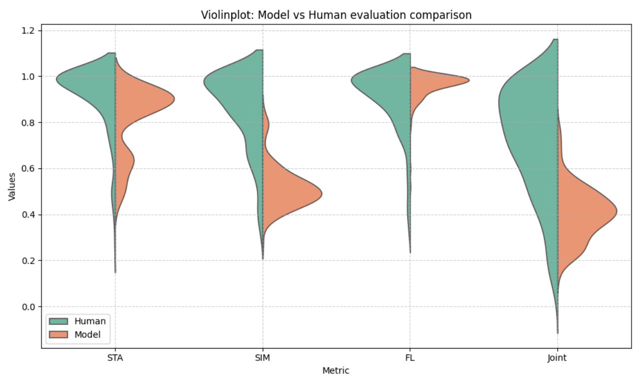


Figure 2. Violin plot: Human vs Model

However, the correlation analysis (Table 2) showed that there is no statistically significant relationship between automatic metrics and human evaluations. For all four indicators, the values of the Spearman and Pearson coefficients ranged from –0.07 to 0.13, while p-values significantly exceeded 0.05. The greatest similarity is observed for the FL metric ($\rho = 0.1343$; r = 0.1218), but even in this case the relationship remains weak and statistically unreliable.

Table 2. Correlations with human judgments

Comparison	Spearman ρ	p-value	Pearson r	p-value
STA vs Human STA	0.0846	0.2334	0.0224	0.7526
SIM vs Human SIM	-0.0309	0.6643	-0.0110	0.8776
FL vs Human FL	0.1343	0.0579	0.1218	0.0858
Joint vs Human JOINT	-0.0718	0.3126	-0.0100	0.8888

We note that the correlation for the STA and FL metrics, despite the similarity of distributions, is also very insignificant. Indeed, the graph in Figure 3 shows that there is a tendency toward high scores both between humans and for the STA and FL model, but often these evaluations do not go in pairs for the same sentence.

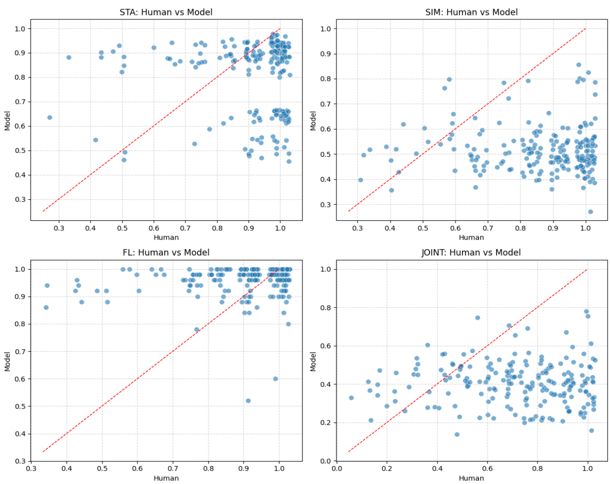


Figure 3. Scatter plot: Human vs Model

What humans rated highly, the model sometimes underestimates, and vice versa. This creates the illusion of close distributions, but the real correlation between specific evaluations is not high. In the SIM graph, we notice the model's tendency to underestimate scores, which significantly affects the final JOINT metric with the same trend.

Thus, automatic metrics in their current form are unlikely to serve as a reliable indicator of human perception of quality, especially in the field of toxic content, where meaning is determined not only by lexical matching but also by context. From the perspective of human evaluation, the detoxification pipeline results showed impressive results, automatic evaluation is also high, but the question arises about its real reflection of detoxification quality with such insignificant correlation with humans, who are the target audience for the model's work results.

6 Ethical Considerations

Creating a tolerant and reserved online community clearly appears to be a benefit for all internet users; however, the ethical dimension of forced detoxification and freedom of speech cannot be ignored. The right to freely express one's thoughts is protected by the constitutions of most countries, which is why a clear boundary must be drawn between a legitimate position and hostile rhetoric [27].

Toxic language is a vague and largely subjective concept, and therefore, any attempt to suppress it may result in transferring power over the right to speak freely into the hands of interested parties. This could lead to even neutral statements being interpreted as offensive for the benefit of certain groups, with the aim of restricting the spread of unwanted opinions. On the other hand, there truly exists no universal classification that would allow a clear distinction between hostile and safe speech: perception depends on many factors, such as cultural context, upbringing, level of empathy, boundaries of what is acceptable, and so on.

Another important ethical aspect of this issue is the unintended side effect of detoxification. The problem lies in the fact that the censorship of hostile and offensive speech does not persuade its authors, but merely attempts to limit the spread of their ideas. However, practice shows that many only become more entrenched in their beliefs and, irritated by forced interference in their speech strategies, seek to bypass filtering algorithms and may even intensify the radicalism of their statements [28].

Still, UN High Commissioner for Human Rights Volker Türk points out that fighting hate speech online is not a form of censorship. In contrast, if there is no regulation, many users lose their opportunity to speak because the environment becomes unsafe and toxic. Moreover, tolerating hatred on the internet can lead to real-world harm beyond the digital space if boundaries of the permissible are not established [29].

7 Limitations and challenges

A significant difficulty in content filtering and detecting profanity is the so-called "Scunthorpe problem." Scunthorpe, a small town in Lincolnshire, UK, gave its name to the phenomenon of unintended blocking of online content that contains a substring matching obscene language. For example, in 1996, Scunthorpe residents could not create accounts on AOL, a major internet service of that time, because the profanity filter blocked the town's name due to the presence of the offensive word 'cunt'.

Several years later, a similar incident occurred when trying to register the domain name "shitakemushrooms.com." Although the domain was available, the automated system flagged the four letters inside it as inappropriate [30]. Even today, such cases are common: users with names or surnames considered "suspicious" by filtering systems still report difficulties. Their emails are rejected by servers, posts fail moderation filters, and accounts may even face sanctions from automated systems.

This happens because even modern systems, which long ago moved away from simply blocking the "seven dirty words" banned from broad-cast television and radio, still struggle with understanding context. Despite major advances in artificial intelligence in recent years, machine learning algorithms continue to face challenges in this area [31].

It is also worth noting that, in addition to false positives when algorithms mistakenly block harmless names or titles, there is the opposite issue: deliberate circumvention of filters. As in other areas of society, restrictions inspire creative ways to bypass them. The use of profanity and toxic language is no exception. In order to avoid being filtered, users find ways to disguise offensive words using similar-sounding phrases, abbreviations, replacing letters with look-alike symbols, or adding emojis. This phenomenon has even been given a name in English: "Algospeak". In this way, people trick algorithms that fail to recognize such words as offensive, while their meaning remains obvious to humans.

This practice can be seen as a form of protest against language control in online communities. Adherents of this method believe that they will always stay one step ahead of algorithms, since people are highly inventive when it comes to finding ways to express meaning while preserving the desired form [32]. In this situation, a potential solution may be to fine-tune models based on user reports, which are able to detect subtext and hidden meanings.

8 Conclusion

This study presented a multilingual text detoxification pipeline combining pretrained transformer models with back-translation for data augmentation. The experiments confirmed the effectiveness of the approach in reducing toxicity while preserving fluency, though meaning preservation remains a challenge. A comparison of automatic and human evaluations revealed weak correlations, highlighting the limitations of current metrics in reflecting human perception. This weakness of automatic metrics opens the way for future research on more universal and suitable evaluation systems. Despite existing ethical and technical challenges, the proposed method demonstrates the potential of automated neutralization as a decent alternative to censoring and moderation, contributing to safer and more tolerant online communication.

References

[1] M. Zampieri, P. Nakov, S. Rosenthal, P. Atanasova, G. Karadzhov, H. Mubarak, L. Derczynski, Z. Pitenis, and Ç. Çöltekin, "SemEval-2020 Task 12: Multilingual Offensive Language Identification in Social Media (OffensEval 2020)," *Proceedings of the Fourteenth Workshop on Semantic Evaluation*, 2020, pp. 1425–1447.

[2] D. Dementieva, D. Moskovskiy, N. Babakov, A. Ali Ayele, N. Rizwan, F. Schneider, X. Wang, S. M. Yimam, D. Ustalov, E. Stakovskii, A. Smirnova, A. Elnagar, A. Mukherjee, and A. Panchenko, "Overview of the Multilingual Text Detoxification Task at PAN 2024," *Working Notes of the Conference and Labs of the Evaluation Forum (CLEF 2024)*, vol. 3740 of CEUR Workshop Proceedings, 2024, pp. 2432–2461.

[3] M. Yi, M. Lim, H. Ko, and J. Shin, "Method of Profanity Detection Using Word Embedding and LSTM," *Mobile Information Systems*, vol. 2021, pp. 1–9, 2021. DOI: `https://doi.org/10.1155/2021/6654029`.

[4] M. S. Jahan and M. Oussalah, "A systematic review of hate speech automatic detection using natural language processing," *Neurocomputing*, vol. 546, Article ID: 126232, 2023. DOI: `https://doi.org/10.1016/j.neucom.2023.126232`.

[5] D. Harish, M. Alamelu, M. Manimaran, and V. P. Jayashakthi, "Automatic Detection of Cyberbullying on Social Media Using Machine Learning," *2nd International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)*, 2023, pp. 1-6.

[6] B. Alrashidi, A. Jamal, I. Khan, and A. Alkhathlan, "A review on abusive content automatic detection: approaches, challenges and opportunities," *PeerJ Computer Science*, 2022, vol. 8, no. 11, Article ID: e1142. DOI: https://doi.org/10.7717/peerj-cs.1142.

[7] E. Shushkevich and J. Cardiff, "Automatic Misogyny Detection in Social Media: A Survey," *Computación y Sistemas*, 2019, vol. 23, no. 4. DOI: 10.13053/cys-23-4-3299.

[8] J. A. Benítez-Andrades, A. González-Jiménez, A. López-Brea, J. Aveleira-Mata, J. M. Alija-Pérez, and M. T. García-Ordás, "Detecting racism and xenophobia using deep learning models on Twitter data: CNN, LSTM and BERT," *PeerJ Comput Science*, 2022, vol. 8, Article ID: e906. DOI: https://doi.org/10.7717/peerj-cs.906.

[9] J. Ruppenhofer, M. Siegel, and M. Wiegand, "Proceedings of the GermEval 2018 Workshop," *14th Conference on Natural Language Processing KONVENS*, 2018, 149 p.

[10] H. Mubarak, K. Darwish, W. Magdy, T. Elsayed, and H. Al-Khalifa, "Overview of OSACT4 Arabic Offensive Language Detection Shared Task," *Proceedings of the 4th Workshop on Open-Source Arabic Corpora and Processing Tools, with a Shared Task on Offensive Language Detection*, pp. 48–52.

[11] T. Mandla, S. Modha, G. K. Shahi, A. K. Jaiswal, D. Nandini, D. Patel, P. Majumder, and J. Schäfer, "Overview of the HASOC track at FIRE 2020: Hate Speech and Offensive Content Identification in Indo-European Languages," *arXiv:2108.05927*.

[12] M. Plakidis and G. Rehm, "A Dataset of Offensive German Language Tweets Annotated for Speech Acts," *Proceedings of the 13th Conference on Language Resources and Evaluation (LREC 2022)*, pp. 4799–4807.

[13] A. Conneau, K. Khandelwal, N. Goyal, V. Chaudhary, G. Wenzek, F. Guzman, E. Grave, M. Ott, L. Zettlemoyer, and V. Stoyanov, "Unsupervised cross-lingual representation learning at

scale," *Proceedings of the 58th Annual Meeting of the ACL*, 2020, pp. 8440–8451.

[14] M. Wiegand, J. Ruppenhofer, A. Schmidt, C. Greenberg C, "Inducing a lexicon of abusive words - a feature-based approach," *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, vol. 1 (Long Papers), 2018, pp. 1046–1056.

[15] C. Nogueira dos Santos, I. Melnyk, and I. Padhi, "Fighting offensive language on social media with unsupervised text style transfer," *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, vol. 2: Short Papers, 2018, pp. 189–194.

[16] D. Dale, A. Voronov, D. Dementieva, V. Logacheva, O. Kozlova, N. Semenov, and A. Panchenko, "Text detoxification using large pre-trained neural models," *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics*, 2021, pp. 7979–7996.

[17] D. Dementieva, D. Moskovskiy, D. Dale, and A. Panchenko, "Exploring methods for cross-lingual text style transfer: The case of text detoxification," *Proceedings of the 13th International Joint Conference on Natural Language Processing and the 3rd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics, IJCNLP 2023*, vol. 1: Long Papers, 2023, pp. 1083–1101.

[18] A. Arango, J. Pérez, and B. Poblete, "Hate speech detection is not as easy as you may think: A closer look at model validation," *Information Systems*, vol. 105, Article ID: 101584, 2022. DOI: `https://doi.org/10.1016/j.is.2020.101584`.

[19] A.P. Vit and R. Puzis, "The use of trigger warnings on social media: a text analysis study of X," *PLOS ONE*, 2025, vol. 20, no. 4, Article ID: e0322549. DOI: `10.1371/journal.pone.0322549`.

[20] V. M. E. Bridgland, P. J. Jones, and B. W. Bellet, "A Meta-Analysis of the Efficacy of Trigger Warnings, Content Warnings,

and Content Notes," *Clinical Psychological Science*, 2023, vol. 12, no. 4, pp. 751–771.

[21] T. Wilde, "Valve debuts "Chat Filtering" feature on Steam — but why did it wait this long to do it?," *GeekWire, 2020.* [Online]. Available: https://www.geekwire.com/2020/valve-debuts-chat-filtering-feature-steam-wait-long/. Accessed 15.08.2025.

[22] D. Nozza and D. Hovy, "The State of Profanity Obfuscation in Natural Language Processing Scientific Publications," *Findings of the Association for Computational Linguistics: ACL 2023*, 2023, pp. 3897–3909.

[23] D. Marsh, "The risks of using asterisks in place of swearwords," *The Guardian*, 2012. [Online]. Available: https://www.theguardian.com/media/mind-your-language/2012/jul/22/mind-your-language-asterisks-swearwords. Accessed 13.08.2025.

[24] V. Protasov, N. Babakov, D. Dementieva, and A. Panchenko, "Evaluating Text Style Transfer: A Nine-Language Benchmark for Text Detoxification," *10.48550/arXiv.2507.15557*.

[25] Meta, "Community Standards: Hate Speech. Meta Transparency Center," [Online]. Available: https://transparency.meta.com/ru-ru/policies/community-standards/hate-speech/. Accessed 10.08.2025.

[26] C. C. Zhang, G. Zaleski, J. N. Kailley, K. A. Teng, M. English, A. Riminchan, and J. M. Robillard, "Debate: Social media content moderation may do more harm than good for youth mental health," *Child Adolesc Ment Health*, vol. 29, pp. 104–106, 2024.

[27] C. O'Regan and S. Theil, "Hate speech regulation on social media: An intractable contemporary challenge," *Research Outreach*, 2020. [Online]. Available: https://researchoutreach.org/articles/hate-speech-regulation-social-media-intractable-contemporary-challenge/. Accessed 05.08.2025.

[28] D. Inserra, “Freedom of Speech Is Worth Celebrating, as Europe Ramps Up Prosecution of 'Hate Speech',” *Cato Institute*, 2025. [Online]. Available: https://www.cato.org/blog/freedom-speech-worth-celebrating-europe-ramps-prosecution-hate-speech. Accessed 05.08.2025.

[29] United Nations, “UN report: Hate speech is on the rise globally, warns UN rights office,” *UN News*, 2025. [Online]. Available: https://news.un.org/en/story/2025/01/1158886. Accessed 05.08.2025.

[30] P. Festa, “Food domain found ”obscene”,” *CNET*, 1998. [Online]. Available: https://www.cnet.com/tech/services-and-software/food-domain-found-obscene/. Accessed 07.08.2025.

[31] D. Oberhaus, “Life on the Internet Is Hard When Your Last Name Is 'Butts',” *VICE*, 2018. [Online]. Available: https://www.vice.com/en/article/life-on-the-internet-is-hard-when-your-last-name-is-butts/. Accessed 07.08.2025.

[32] H. Richter, “Adam Aleksic on how algorithms are transforming the way we communicate,” *Reuters*, 2025. [Online]. Available: https://www.reuters.com/lifestyle/adam-aleksic-how-algorithms-are-transforming-way-we-communicate-2025-07-24/. Accessed 07.08.2025.

[33] M. R. Costa-jussà et al., “No language left behind: Scaling human-centered machine translation,” 2022, *https://arxiv.org/abs/2207.04672*.

[34] M. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, V. Stoyanov, and L. Zettlemoyer, “BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension,” *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 2019, pp. 7871–7880.

[35] I. Sutskever, O. Vinyals, and Q. L. Le, "Sequence to sequence learning with neural networks," *Advances in Neural Information Processing Systems 27 (NIPS 2014)*, 2014, pp. 3104–3112.

[36] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, and P. J. Liu, "Exploring the limits of transfer learning with a unified text-to-text transformer," *Journal of Machine Learning Research*, 2020, vol. 21, no. 140, pp. 1–67.

[37] P. Fortuna and S. Nunes, "A Survey on Automatic Detection of Hate Speech in Text," *ACM Comput. Surv.*, 2018, vol. 51, no. 4, Article No.: 85, Pages 1–30. DOI: https://doi.org/10.1145/3232676.

Maia Kuzmitkaia[1],Alexandr Moroz[2], Victoria Bobicev[3]

[1]ORCID: https://orcid.org/0009-0006-7173-3203
Technical University of Moldova
E–mail: kuzmitkaiamaia@gmail.com

[2]ORCID: https://orcid.org/0009-0008-3332-6193
Technical University of Moldova
E–mail: alexandrmoroz.md@gmail.com

[3]ORCID: https://orcid.org/0000-0003-4450-3964
Technical University of Moldova
E–mail: victoria.bobicev@ia.utm.md

Proceedings of Workshop on Intelligent Information Systems
WIIS2025, October 16-18, 2025, Chisinau, Republic of Moldova

A queuing model with arrivals according to a two-dimensional diffusion process

Mario Lefebvre

Abstract

In the case of heavy traffic, the arrival process in a queuing model can be approximated by a diffusion process. In this paper, we consider such a model. The number $X(t)$ of customers in the system at time t evolves according to a degenerate two-dimensional diffusion process. In a particular case, the distribution of $X(t)$ is calculated explicitly. Moreover, a stochastic control problem known as a homing problem is formulated, and the equation satisfied by the value function is derived.

Keywords: arrival process, Gaussian process, stochastic optimal control, first-passage time, dynamic programming.

1 Introduction

In classical queuing models, the arrivals of customers constitute a discrete-space stochastic process. For instance, in the case of the $M/M/k$ model, customers arrive according to a Poisson process. However, in some applications, the arrivals occur almost according to a diffusion process; in particular, in the case of *heavy traffic* (for instance, on the Internet).

In this paper, we assume that the number $X(t)$ of customers in the system at time t is such that

$$\mathrm{d}X(t) = \rho[Y(t)]\,\mathrm{d}t - k\,X(t)\,\mathrm{d}t, \tag{1}$$

$$\mathrm{d}Y(t) = m[Y(t)]\,\mathrm{d}t + \{v[Y(t)]\}^{1/2}\,\mathrm{d}B(t), \tag{2}$$

where $\{B(t), t \geq 0\}$ is a one-dimensional standard Brownian motion. The functions $m(\cdot) \in \mathbb{R}$ and $v(\cdot) > 0$ are such that $\{Y(t), t \geq 0\}$ is a diffusion process. Moreover, the non-negative constant k is the rate at which the customers are served. The function $\rho(\cdot)$ should be such that if $k = 0$, then $X(t)$ will increase with time t.

This type of degenerate two-dimensional diffusion process, which was proposed by Rishel [1], has been used in reliability theory to model the wear of devices. Indeed, wear should be strictly increasing with time.

In the next section, a particular case for the various functions in Eqs. (1) and (2) will be considered. Then, in Section 3, an optimal control problem for the two-dimensional process $\{(X(t), Y(t)), t \geq 0\}$ will be studied.

2 A particular case

Suppose that $m(\cdot) \equiv \mu > 0$ and $v(\cdot) \equiv \sigma^2$. Then, $\{Y(t), t \geq 0\}$ is a Wiener process with positive drift μ and dispersion parameter $\sigma > 0$. A Wiener process, being a Gaussian process, can take both positive and negative values. Therefore, the function $\rho(\cdot)$ should be such that $\rho[Y(t)]$ is a non-negative function. For example, we could take $\rho[Y(t)] = Y^2(t)$. However, if we assume that $y := Y(0)$ and μ are both large enough, and that σ is small, then the probability that $Y(t)$ becomes negative is negligible.

We choose $\rho[Y(t)] = cY(t)$, with c being a positive constant. With this choice, we can appeal to the following proposition to compute the joint probability density function of the random vector $(X(t), Y(t))$.

Proposition 1. (See [2].) *Let $\{\mathbf{X}(t), t \geq 0\}$ be an n-dimensional stochastic process defined by*

$$d\mathbf{X}(t) = (\mathbf{A}\mathbf{X}(t) + \mathbf{a})\,\mathrm{d}t + \mathbf{N}^{1/2}\,\mathrm{d}\mathbf{B}(t), \tag{3}$$

where $\{\mathbf{B}(t), t \geq 0\}$ is an n-dimensional standard Brownian motion, $\mathbf{A}$ is a square matrix of order n, $\mathbf{a}$ is an n-dimensional vector, and

$\mathbf{N}^{1/2}$ *is a positive definite square matrix of order* n. *Then, given that* $\mathbf{X}(t_0) = \mathbf{x}$, *we may write that*

$$\mathbf{X}(t) \sim \mathrm{N}\left(\mathbf{m}(t), \mathbf{K}(t)\right) \quad \textit{for } t \geq t_0, \tag{4}$$

where

$$\mathbf{m}(t) := \boldsymbol{\Phi}(t)\left(\mathbf{x} + \int_{t_0}^{t} \boldsymbol{\Phi}^{-1}(u)\,\mathbf{a}\,\mathrm{d}u\right) \tag{5}$$

and

$$\mathbf{K}(t) := \boldsymbol{\Phi}(t)\left(\int_{t_0}^{t} \boldsymbol{\Phi}^{-1}(u)\mathbf{N}[\boldsymbol{\Phi}^{-1}(u)]'\,\mathrm{d}u\right)\boldsymbol{\Phi}'(t), \tag{6}$$

where the symbol prime *denotes the* transpose *of the matrix, and the function* $\boldsymbol{\Phi}(t)$ *is given by*

$$\boldsymbol{\Phi}(t) := e^{\mathbf{A}(t-t_0)} = \sum_{n=0}^{\infty} \mathbf{A}^n \frac{(t-t_0)^n}{n!}. \tag{7}$$

In our case, we have

$$\mathbf{A} = \begin{bmatrix} -k & c \\ 0 & 0 \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} 0 \\ \mu \end{bmatrix} \quad \text{and} \quad \mathbf{N}^{1/2} = \begin{bmatrix} 0 & 0 \\ 0 & \sigma \end{bmatrix}. \tag{8}$$

Moreover, for $n \geq 1$,

$$\mathbf{A}^n = \begin{bmatrix} (-k)^n & -c\,(-k)^{n-1} \\ 0 & 0 \end{bmatrix}. \tag{9}$$

Hence, if $t_0 = 0$, the function $\boldsymbol{\Phi}(t)$ is given by

$$\begin{aligned} \boldsymbol{\Phi}(t) &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} -1 + e^{-kt} & -\frac{c}{k}\left(-1 + e^{-kt}\right) \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} e^{-kt} & -\frac{c}{k}\left(-1 + e^{-kt}\right) \\ 0 & 1 \end{bmatrix}. \end{aligned} \tag{10}$$

Next, we find that

$$\int_0^t \boldsymbol{\Phi}^{-1}(u)\,\mathbf{a}\,\mathrm{d}u = \begin{bmatrix} \frac{c\,\mu\left(-k\,t - 1 + e^{k\,t}\right)}{k^2} \\ \mu\,t \end{bmatrix}. \tag{11}$$

It follows that

$$\mathbf{m}(t) = \begin{bmatrix} e^{-k\,t}\left(x + \frac{c\,\mu\left(-k\,t-1+e^{k\,t}\right)}{k^2}\right) - \frac{c\left(e^{-k\,t}-1\right)(\mu\,t+y)}{k} \\ \mu\,t+y \end{bmatrix}. \tag{12}$$

Finally, with

$$\mathbf{N} = \begin{bmatrix} 0 & 0 \\ 0 & \sigma^2 \end{bmatrix}, \tag{13}$$

we calculate

$$\int_0^t \mathbf{\Phi}^{-1}(u)\mathbf{N}[\mathbf{\Phi}^{-1}(u)]'\,\mathrm{d}u = \begin{bmatrix} -\frac{c^2\sigma^2\left(-2\,k\,t+4\,e^{k\,t}-e^{2\,k\,t}-3\right)}{2\,k^3} & -\frac{c\,\sigma^2\left(-k\,t-1+e^{k\,t}\right)}{k^2} \\ -\frac{c\,\sigma^2\left(-k\,t-1+e^{k\,t}\right)}{k^2} & \sigma^2 t \end{bmatrix} \tag{14}$$

and

$$\mathbf{K}(t) = \begin{bmatrix} \frac{c^2\sigma^2\left(2\,e^{2\,k\,t}k\,t-3\,e^{2\,k\,t}+4\,e^{k\,t}-1\right)e^{-2\,k\,t}}{2\,k^3} & \frac{c\,\sigma^2\left(k\,t-1+e^{-k\,t}\right)}{k^2} \\ \frac{c\,\sigma^2\left(k\,t-1+e^{-k\,t}\right)}{k^2} & \sigma^2 t \end{bmatrix}. \tag{15}$$

Notice that for t large, $X(t)$ has a Gaussian distribution with mean and variance that are both proportional to t.

Making use of the above results, we can compute the probability that $X(t)$ will become equal to zero as a function of time, so that the queue is empty. Similarly, if the system capacity is finite, we can easily compute the probability that the system will become saturated.

The actual number of customers in the system is given by

$$X_r(t) := \begin{cases} 0 & \text{if } X(t) \le 0, \\ X(t) & \text{if } 0 < X(t) < r, \\ r & \text{if } X(t) \ge r, \end{cases} \tag{16}$$

where r is the system capacity.

3 A homing problem

In this section, we suppose that the constant k in Eq. (1) is replaced by the function $b_0\,u[X(t),Y(t)]$, where b_0 is a positive constant and $u(\cdot,\cdot)$ is a control variable that is assumed to be a continuous function.

Let

$$\tau(x,y) := \inf\{t > 0 : X(t) = 0 \text{ or } Y(t) = \gamma \mid X(0) = x, Y(0) = y\}, \tag{17}$$

where $x > 0$ and $y > \gamma$, and $\gamma \geq 0$ is assumed to be small. The random variable $\tau(x,y)$ is called a *first-passage time* in probability theory.

Our aim is to find the control that minimizes the expected value of the cost function

$$J(x,y) := \int_0^{\tau(x,y)} \left\{ \frac{1}{2}\, q_0\, u^2[X(t),Y(t)] + \theta \right\} \mathrm{d}t, \tag{18}$$

where q_0 and θ are positive constants. Hence, the optimizer tries to empty the queue as soon as possible, while taking the quadratic control costs into account. We also stop controlling the process if the arrival rate of customers becomes small enough. This type of stochastic control problem is known as a *homing problem*; see Whittle [3] and [1]. The author has written several papers on homing problems; see, for example, [4].

To find the optimal control $u^*[X(t),Y(t)]$, we can try using dynamic programming, which enables us to express u^* in terms of the value function

$$F(x,y) := \inf_{\substack{u[X(t),Y(t)] \\ t \in [0,\tau(x,y))}} E[J(x,y)]. \tag{19}$$

The function $F(x,y)$ gives the smallest expected cost, starting form $X(0) = x$ and $Y(0) = y$.

We can prove the following proposition.

Proposition 2. *The value function $F(x,y)$ satisfies the* dynamic programming equation

$$0 = \inf_{u(x,y)} \left\{ \frac{1}{2} q_0 u^2(x,y) + \theta + [\rho(y) - b_0 u(x,y)] F_x(x,y) + m(y) F_y(x,y) + \frac{1}{2} v(y) F_{yy}(x,y) \right\}, \tag{20}$$

where $F_x = \frac{\partial}{\partial x} F(x,y)$, etc. The equation is valid for $x > 0$ and $y > \gamma$. We have the boundary conditions $F(0,y) = F(x,\gamma) = 0$. Moreover, the optimal control is given by

$$u^*(x,y) = \frac{b_0}{q_0} F_x(x,y). \tag{21}$$

If we substitute the expression for $u^*(x,y)$ into the dynamic programming equation, we find that to obtain the value function, we must solve the second-order non-linear partial differential equation (PDE)

$$\theta - \frac{b_0}{2 q_0} [F_x(x,y)]^2 + \rho(y) F_x(x,y) + m(y) F_y(x,y) + \frac{1}{2} v(y) F_{yy}(x,y) = 0. \tag{22}$$

Assume now that instead of the service rate, the optimizer can control the arrival rate of the process, so that the two-dimensional process $(X(t), Y(t))$ is defined by the system of stochastic differential equations

$$\mathrm{d}X(t) = \rho[Y(t)]\,\mathrm{d}t - k\,X(t)\,\mathrm{d}t, \tag{23}$$

$$\mathrm{d}Y(t) = b_0 u[X(t), Y(t)]\,\mathrm{d}t + m[Y(t)]\,\mathrm{d}t + \{v[Y(t)]\}^{1/2}\,\mathrm{d}B(t). \tag{24}$$

Proceeding as above, we obtain the following corollary.

Corollary 1. *In the case of the controlled process defined in Eqs. (23) and (24), the optimal control is given by*

$$u^*(x,y) = -\frac{b_0}{q_0} F_y(x,y). \tag{25}$$

Furthermore, the value function satisfies the PDE

$$0 = \theta - \frac{b_0}{2q_0}[F_x(x,y)]^2 + [\rho(y) - kx]\,F_x(x,y) + m(y)\,F_y(x,y) + \frac{1}{2}v(y)\,F_{yy}(x,y). \quad (26)$$

Finally, in some cases, Eq. (26) can be linearized.

Proposition 3. *Suppose that* $v(y) \equiv \sigma^2$. *The function*

$$G(x,y) := e^{-\alpha F(x,y)}, \quad (27)$$

where

$$\alpha := \frac{b_0^2}{q_0\,\sigma^2}, \quad (28)$$

satisfies the linear *second-order PDE*

$$\alpha\theta G(x,y) = [\rho(y) - kx]\,G_x(x,y) + m(y)\,G_y(x,y) + \frac{1}{2}\sigma^2\,G_{yy}(x,y). \quad (29)$$

The boundary conditions are $G(0,y) = G(x,\gamma) = 1$.

Remark 1. *The linear PDE in (29) is in fact the* Kolmogorov backward equation *satisfied by the moment-generating function of the random variable* $\tau(x,y)$:

$$M(x,y) := E\left[e^{-s\tau(x,y)}\right], \quad (30)$$

with $s = \alpha\theta > 0$, *for the* uncontrolled *process* $(X_0(t), Y_0(t))$ *obtained by setting* $u[X(t), Y(t)] \equiv 0$ *in Eq. (24). Moreover, the boundary conditions are the appropriate ones.*

4 Conclusion

In this paper, a queuing model in which customers arrive (approximately) according to a degenerate two-dimensional diffusion process was studied. The model is such that, in the case of the absence of service, the number of customers in the system is strictly increasing.

In a particular case, we were able to derive the distribution of the number $X(t)$ of customers in the system at time t. It would be interesting to obtain this distribution in other cases.

Then, a stochastic control problem was formulated for the queuing model. We derived the non-linear partial differential equation satisfied by the value function, and we saw that it is sometimes possible to linearize this equation. In a next step, we could try to obtain explicit solutions to special problems.

Acknowledgments. This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

[1] R. Rishel, “Controlled wear processes: modeling optimal control,” *IEEE Trans. Autom. Control.*, vol. 36, no. 9, pp. 1100–1102, 1991. DOI: https://doi.org/10.1109/CDC.1989.70212.

[2] M. Lefebvre, *Applied Stochastic Processes*, New York, NY, USA: Springer, 2007, 382 p.

[3] P. Whittle, *Optimization Over Time*, Volume I: *Dynamic Programming and Stochastic Control*, Chichester, UK: Wiley, 1982, 317 p.

[4] M. Lefebvre, “Minimizing or maximizing the first-passage time to a time-dependent boundary,” *Optimization*, vol. 71, no. 2, pp. 387–401, 2022. DOI: https://doi.org/10.1080/02331934.2021.1914039.

Mario Lefebvre

ORCID: https://orcid.org/0000-0002-9451-543X
Polytechnique Montréal
E–mail: mlefebvre@polymtl.ca

Proceedings of Workshop on Intelligent Information Systems
WIIS2025, October 16-18, 2025, Chisinau, Republic of Moldova

Quotient Sorts in Matching Logic

Dorel Lucanu

Abstract

A quotient type is a new type you create by taking an existing type and "gluing together" elements that you consider equivalent. Matching Logic ($\mathbb{ML}$) is a formal logic framework designed for specifying and reasoning about the state and behavior of programs. The $\mathbb{ML}$'s central concept is that of pattern that can match against structures. The types are specified in $\mathbb{ML}$ using sorts. In this paper, we show how quotient types can be specified in $\mathbb{ML}$ as quotient sorts.

Keywords: quotient types, matching logic, sorts, initial algebra.

1 Introduction

A quotient type is *the type of equivalence classes of a type under a given equivalence relation*, used whenever we want to ignore irrelevant differences while keeping type safety. Quotient types are a basic stone of *quotient inductive types* [5], which provide a natural way to study universal algebra in type theory.

Let us consider the example of the associative lists (free monoids) over a set A. Firstly, we define the "raw lists" as the set of all finite lists over A, denoted by A_* and inductively defined as follows:

- the empty list *nil* is in A_*;
- every $a \in A$ is in A_*; and
- if $l_1, l_2 \in A_*$, then the concatenation $l_1 \cdot l_2$ is in A_*.

The equivalence relation $\sim$ is the congruence relation generated by the following axioms:

$$l \cdot nil \sim l \qquad \text{(right unit)}$$
$$nil \cdot l \sim l \qquad \text{(left unit)}$$
$$l_1 \cdot (l_2 \cdot l_3) \sim (l_1 \cdot l_2) \cdot l_3 \qquad \text{(associativity)}$$

Then the associative lists A^* are defined as the quotient type $A_*/_\sim$, which is the set of equivalence classes of A_* under $\sim$: $A^* = \{[l] \mid l \in A_*\}$, where $[l]$ is the equivalence class of l under $\sim$.

Quotient types are not easy to handle in proof assistants, as they require a careful treatment of the equivalence relation and the operations defined on the type. For instance, if we define a function on the raw lists, we need to ensure that it respects the equivalence relation, i.e., it should yield the same result for equivalent lists. This can be cumbersome and error-prone. Here is a brief description of how some proof assistants handle quotient types:

- Rocq (former Coq)[1] does not have quotient types built in. In Rocq's logic (Calculus of Inductive Constructions), all terms are definitionally distinct unless reduced by computation. So, there is no a way to directly "collapse" a datatype by equations such as associativity.

- Isabelle/HOL[2] has built-in quotient type support: you can build a datatype (like raw-asooc-list) and then quotient it by an equivalence relation.

- Agda[3]'s core type theory is Martin–Löf Type Theory [6], which does not allow you to simply "glue together" terms modulo an arbitrary equivalence relation. Cubical Agda, a version of Agda that supports Higher Inductive Types [1], allows you to directly define quotients.

[1] https://rocq-prover.org/
[2] https://isabelle.in.tum.de/
[3] https://agda.readthedocs.io/en/latest/

- In Lean[4], the quotient types are primitive, so you can work with them directly.

Matching Logic ($\mathbb{ML}$) [3] is a formal logic framework designed for specifying and reasoning about the state and behavior of programs. Unlike traditional logics (like First-Order Logic) that focus on propositions being true or false in a model, $\mathbb{ML}$'s central concept is that of a pattern. Patterns are expressions that can match against structures, allowing for flexible and powerful specifications.

The types are specified in $\mathbb{ML}$ using *sorts*. Sorts can be thought of as a way to organize the universe of discourse in $\mathbb{ML}$, providing a structure for reasoning about patterns and their relationships.

In this paper, we show how quotient types can be specified in $\mathbb{ML}$ as quotient sorts. The main idea is to use the initial algebra semantics [2] of sorts to define quotient sorts in $\mathbb{ML}$.

2 Preliminaries

2.1 Quotient Types

Mathematically, the definition of quotient types are quite simple. Given:

- a type A; and
- an equivalence (reflexive, symmetric, transitive) relation
$$\sim \subseteq A \times A,$$

the quotient type $A/_\sim$ is the type of equivalence classes:

$$A/_\sim = \{[a] \mid a \in A\}$$

where $[a] = \{b \in A \mid a \sim b\}$.

Most dependently typed languages do not let you freely define quotient types without care, because:

[4] https://github.com/leanprover

- you need to ensure the equivalence relation is well-defined;
- any function on the quotient must respect the equivalence:

$$\forall a, b \in A, a \sim b \implies f(a) = f(b)$$

Otherwise, you could define ill-formed functions that depend on arbitrary representatives.

2.2 Matching Logic

2.2.1 Syntax

An $\mathbb{ML}$ signature (Σ, EV, SV) consists of a set of *constant symbols* Σ, a set of *element variables* EV, and a set of *set variables* SV. The formulas, called *patterns*, are defined as follows:

$$\varphi ::= x \mid X \mid \sigma \mid \varphi_1\ \varphi_2 \mid \bot \mid \varphi_1 \to \varphi_2 \mid \exists x.\varphi \mid \mu X.\varphi \text{ if } \varphi \text{ is positive in } X \quad \text{(PatSyn)}$$

A pattern φ is positive in X if every free occurrence of X is under an even number of negations in φ, where a negation $\neg\psi$ is the notation $\psi \to \bot$.

Let Pattern denote the set of patterns.

2.2.2 Semantics

Definition 1 (Models)**.** Given $\mathbf{\Sigma} = (EV, SV, \Sigma)$, a $\mathbf{\Sigma}$-*model* (or simply *model*) is a tuple $(M, \_\bullet\_, \{M_\sigma\}_{\sigma\in\Sigma})$, where

1. M is a carrier set, required to be nonempty;
2. $\_\bullet\_ \colon M \times M \to \mathcal{P}(M)$ is a function, called the *interpretation of application*; here, $\mathcal{P}(M)$ is the powerset of M;
3. $M_\sigma \subseteq M$ is a subset of M, called the *interpretation of σ in M* for each $\sigma \in \Sigma$.

By abuse of notation, we write M to denote the above model.

For notational simplicity, we extend $\_\bullet\_$ from over elements to over sets, *pointwisely*, as follows:

$$\_\bullet\_ \colon \mathcal{P}(M) \times \mathcal{P}(M) \to \mathcal{P}(M) \quad A \bullet B = \bigcup_{a \in A, b \in B} a \bullet b \;\text{ for } A, B \subseteq M$$

Note that $\emptyset \bullet A = A \bullet \emptyset = \emptyset$ for any $A \subseteq M$.

Definition 2 (Pattern interpretation)**.** Given $\mathbf{\Sigma} = (EV, SV, \Sigma)$ and a model M, an *M-valuation* (or simply *valuation*) is a function $\rho\colon (EV \cup SV) \to (M \cup \mathcal{P}(M))$ that maps element variables to elements in M and set variables to subsets of M; that is, $\rho(x) \in M$ for all $x \in EV$ and $\rho(X) \subseteq M$ for all $X \in SV$. We define a *pattern interpretation* $|\_|_\rho \colon \textsc{Pattern} \to \mathcal{P}(M)$ inductively as follows:

$$\begin{array}{ll} |x|_\rho = \{\rho(x)\} & |X|_\rho = \rho(X) \\ |\sigma|_\rho = M_\sigma & |\bot|_\rho = \emptyset \\ |\varphi_1\, \varphi_2|_\rho = |\varphi_1|_\rho \bullet |\varphi_2|_\rho & |\varphi_1 \to \varphi_2|_\rho = M \setminus (|\varphi_1|_\rho \setminus |\varphi_2|_\rho) \\ |\exists x.\, \varphi|_\rho = \bigcup_{a \in M} |\varphi|_{\rho[a/x]} & |\mu X.\, \varphi|_\rho = \mu \mathcal{F}^\rho_{X,\varphi} \end{array} \qquad \text{(PatInt)}$$

where $\rho[a/x]$ is the valuation ρ' such that $\rho'(x) = a$, $\rho'(y) = \rho(y)$ for any $y \in EV$ distinct from x, and $\rho'(X) = \rho(X)$ for any $X \in SV$. Here, $\mathcal{F}^\rho_{X,\varphi} \colon \mathcal{P}(M) \to \mathcal{P}(M)$ is the function defined as $\mathcal{F}^\rho_{X,\varphi}(A) = |\varphi|_{\rho[A/X]}$ for every $A \subseteq M$, where $\rho[A/X]$ is the valuation ρ' such that $\rho'(X) = A$, $\rho'(Y) = \rho(Y)$ for any $Y \in SV$ distinct from X, and $\rho'(x) = \rho(x)$ for any $x \in EV$.

Definition 3.

1. φ is *valid in M* (equivalently, φ *holds in M*), written $M \models \varphi$, if $|\varphi|_\rho = M$, for any M-valuation ρ.

2. If Γ is a set of patterns, then $M \models \Gamma$ if $M \models \varphi$ for any $\varphi \in \Gamma$.

3. $\Gamma \models \varphi$ if $M \models \varphi$ for any model M with $M \models \Gamma$. We say that φ is a *semantic consequence* of Γ. If Γ is the empty set, we simply write $\models \varphi$.

2.2.3 Notations

Notations [7] is a mechanism that allows flexible use of $\mathbb{ML}$. In particular, the notations can be used to define a domain-specific logic inside $\mathbb{ML}$. It uses a less conventional notion of theory given as a triple $(\Sigma, \Phi, \vdash)$, where Σ is the alphabet of the constant symbols, Φ a set of formulas (patterns), and $\vdash$ an entailment relation. A *notation-based specification* is given as a conservative inclusion theory morphism $(\Sigma, \Phi, \vdash) \hookrightarrow (\Sigma, \Phi', \vdash')$, such that each *new formula* $\varphi' \in \Phi' \setminus \Phi$ is a *notation* of a formula $\varphi \in \Phi$, written as $\varphi' :\leftrightarrow \varphi$ and expressed by two new axioms: $\vdash' \varphi' \rightarrow \varphi$ and $\vdash' \varphi \rightarrow \varphi'$. Φ' might also contain extra axioms that limit the use of notations.

2.3 Two Basic $\mathbb{ML}$ Theories

2.4 Derived Connectives

- New formulas (patterns):

$$\varphi ::= \neg\varphi \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \leftrightarrow \varphi_2 \mid \forall x.\, \varphi \mid \nu X.\, \varphi \tag{DerPatSyn}$$

- Notations:

$$\begin{aligned}
\neg\varphi &:\leftrightarrow \varphi \rightarrow \bot && \text{/* negation */}\\
\varphi_1 \vee \varphi_2 &:\leftrightarrow (\varphi_1 \rightarrow \bot) \rightarrow \varphi_2 && \text{/* disjunction */}\\
\varphi_1 \wedge \varphi_2 &:\leftrightarrow \neg(\neg\varphi_1 \vee \neg\varphi_2) && \text{/* conjunction */}\\
\varphi_1 \leftrightarrow \varphi_2 &:\leftrightarrow (\varphi_1 \rightarrow \varphi_2) \wedge (\varphi_2 \rightarrow \varphi_1) && \text{/* equivalence */}\\
\forall x.\, \varphi &:\leftrightarrow \neg\exists x.\, \neg\varphi && \text{/* universal quantification */}\\
\nu X.\, \varphi &:\leftrightarrow \neg\, \mu X.\, \neg\varphi[\neg X/X] && \text{/* greatest fixpoint */}
\end{aligned}$$

where the set variable substitution $\varphi[\psi/X]$ can also be defined using the notation mechanism:

$$
\begin{aligned}
x[\psi/X] &:\leftrightarrow x && , x \in EV \\
X[\psi/X] &:\leftrightarrow \psi \\
Y[\psi/X] &:\leftrightarrow Y && , Y \neq X, Y \in SV \\
\bot[\psi/X] &:\leftrightarrow \bot \\
\sigma[\psi/X] &:\leftrightarrow \sigma \\
(\varphi_1 \to \varphi_2)[\psi/X] &:\leftrightarrow (\varphi_1[\psi/X]) \to (\varphi_2[\psi/X]) \\
(\exists y.\varphi)[\psi/X] &:\leftrightarrow \exists y.\varphi([\psi/X]) && , y \notin FEV(\psi) \\
(\mu X.\varphi)[\psi/X] &:\leftrightarrow \mu X.\varphi \\
(\mu Y.\varphi)[\psi/X] &:\leftrightarrow \mu Y.(\varphi[\psi/X]) && , Y \neq X, Y \in SV, Y \notin FSV(\psi)
\end{aligned}
$$

Remark 1. Note that the grammar (DerPatSyn) extends the grammar (PatSyn), so the productions of (PatSyn) can be used to obtain new notations.

2.5 Equality, Membership, Inclusion

- New formulas (patterns):

$$\varphi ::= \lceil\varphi\rceil \mid \lfloor\varphi\rfloor \mid \varphi_1 = \varphi_2 \mid \varphi_1 \subseteq \varphi_2 \mid x \in \varphi$$

- Axioms:

$$\forall x.\,\mathsf{def}\ x \qquad \text{/* definedness */} \tag{Ax.1}$$

- Notations:

$$
\begin{aligned}
\lceil\varphi\rceil &:\leftrightarrow (\mathsf{def}\ x) && \text{/* definedness notation, } def \in \Sigma \text{ */} \\
\lfloor\varphi\rfloor &:\leftrightarrow \neg\lceil\neg\varphi\rceil && \text{/* totality */} \\
\varphi_1 = \varphi_2 &:\leftrightarrow \lfloor\varphi_1 \leftrightarrow \varphi_2\rfloor && \text{/* equality */} \\
\varphi_1 \subseteq \varphi_2 &:\leftrightarrow \lfloor\varphi_1 \to \varphi_2\rfloor && \text{/* set inclusion */} \\
x \in \varphi &:\leftrightarrow x \subseteq \varphi && \text{/* membership */}
\end{aligned}
$$

Intuitively, a pattern is defined, $\vdash \lceil\varphi\rceil$, iff its interpretation in any model is the non-empty set. It is defined by (Ax.1), which says that any element variable is defined (because its interpretation is always a singleton). See [3] for more details.

3 An $\mathbb{ML}$ Theory of Sorts

3.1 Sorts

A *sort* has a name and is associated with a set of its *inhabitants*. If $s \in \Sigma$ represents a sort name, then the pattern $(\mathsf{inh}\ s)$ represents all its inhabitants, where $\mathsf{inh} \in \Sigma$. We use the notation

$$\top_s :\leftrightarrow (\mathsf{inh}\ s)$$

for the set of inhabitants of a given sort s. We also use the symbol Sort a special sort name such that its inhabitants are all the sorts (including itself).

- New formulas (patterns):

$$\varphi ::= \top_s \mid \forall x{:}s.\,\varphi \mid \exists x{:}s.\,\varphi \mid \forall x_1,\ldots,x_n{:}s.\,\varphi \mid \exists x_1,\ldots,x_n{:}s.\,\varphi$$

- Axioms:

$$\begin{gathered}\mathsf{Sort} \in \top_{\mathsf{Sort}}\\ \forall s.s \in \top_{\mathsf{Sort}} \to \lceil\top_s\rceil\end{gathered}$$

- Notations:

$$\begin{aligned}
\top_s &:\leftrightarrow \mathsf{inh}\ s && \text{/* inhabitants of } s \text{ */}\\
\forall x{:}s.\,\varphi &:\leftrightarrow \forall x.\,x \in \top_s \to \varphi && \text{/* } \forall \text{ within sort } s \text{ */}\\
\exists x{:}s.\,\varphi &:\leftrightarrow \exists x.\,x \in \top_s \wedge \varphi && \text{/* } \exists \text{ within sort } s \text{ */}\\
\forall x_1,\ldots,x_n{:}s.\,\varphi &:\leftrightarrow \forall x_1{:}s.\ \ldots \forall x_n{:}s.\,\varphi && \text{/* nested } \forall \text{ within } s \text{ */}\\
\exists x_1,\ldots,x_n{:}s.\,\varphi &:\leftrightarrow \exists x_1{:}s.\ \ldots \exists x_n{:}s.\,\varphi && \text{/* nested } \exists \text{ within } s \text{ */}
\end{aligned}$$

Example 1. Raw associative lists are specified in $\mathbb{ML}$ as follows, where List, nil, and $\cdot$ are constant symbols:

// notation for parametric lists

$\forall s{:}\mathsf{Sort}.\mathsf{List}\langle s\rangle :\leftrightarrow \mathsf{List}\ s$

$\forall s{:}\mathsf{Sort}.\exists s'{:}\mathsf{Sort}.\mathsf{List}\langle s\rangle = s'$ /* $\mathsf{List}\langle \mathrm{s}\rangle$ is a sort ($\in \top_{\mathsf{Sort}}$) */

$\forall s{:}\mathsf{Sort}.\mathsf{nil} \in \top_{\mathsf{List}\langle s\rangle}$ /* the empty list is a $\mathsf{List}\langle s\rangle$ */

$\forall s{:}\mathsf{Sort}.\forall x{:}s.\exists l{:}\mathsf{List}\langle s\rangle.x = l$ /* every s is a $\mathsf{List}\langle s\rangle$ ($\top_s \subseteq \top_{\mathsf{List}\langle s\rangle}$) */

// infix notation for concatenation

$\forall s{:}\mathsf{Sort}.\forall l_1, l_2{:}\mathsf{List}\langle s\rangle.\exists l{:}\mathsf{List}\langle s\rangle.l_1 \cdot l_2 :\leftrightarrow (\cdot\ l_1\ l_2)$

// concatenation of lists is a list

$\forall s{:}\mathsf{Sort}.\forall l_1, l_2{:}\mathsf{List}\langle s\rangle.\exists l{:}\mathsf{List}\langle s\rangle.l_1 \cdot l_2 = l$

// concatenation is injective

$\forall s{:}\mathsf{Sort}.\forall l_1, l_2, l_1', l_2'{:}\mathsf{List}\langle s\rangle.l_1 \cdot l_2 = l_1' \cdot l_2' \to l_1 = l_1' \wedge l_2 = l_2'$

$\forall s{:}\mathsf{Sort}.\forall x{:}s.\mathsf{nil} \neq x$ /* the empty list is not a s */

// the empty list is not a concat.

$\forall s{:}\mathsf{Sort}.\forall l_1, l_2{:}\mathsf{List}\langle s\rangle.\mathsf{nil} \neq l_1 \cdot l_2$

// an element of s is not a concatenation

$\forall s{:}\mathsf{Sort}.\forall x{:}s.\forall l_1, l_2{:}\mathsf{List}\langle s\rangle.x \neq l_1 \cdot l_2$

// inductive definition of $\mathsf{List}\langle s\rangle$

$\forall s{:}\mathsf{Sort}.\top_{\mathsf{List}\langle s\rangle} = \mu X.(\mathsf{nil} \vee \exists x{:}s.x \vee X \cdot X)$

3.2 Product sorts

Given two sorts s_1 and s_2 (i.e., $s_1, s_2 \in \top_{\mathsf{Sort}}$), their *product sort* is specified by

- a sort $s_1 \otimes s_2$

$$s_1 \otimes s_2 \in \top_{\mathsf{Sort}}$$

- a constant symbol $\langle \_, \_ \rangle$ in Σ, together with a notation

$$\langle x, y \rangle :\leftrightarrow \langle \_, \_ \rangle \; x \; y$$

and the following axioms:

$$\begin{array}{ll} \forall x{:}s_1.\forall y{:}s_2.\exists z{:}s_1{\otimes}s_2.\langle x, y \rangle = z & \text{/* functional */} \\ \langle x_1, y_1 \rangle = \langle x_2, y_2 \rangle \to x_1 = x_2 \wedge y_1 = y_2 & \text{/* injective */} \\ \top s_1{\otimes}s_2 = \exists x{:}s_1.\exists y{:}s_2.\langle x, y \rangle & \text{/* inductive */} \\ \forall x{:}s_1.\forall y{:}s_2.\forall z{:}s_3.\langle \langle x, y \rangle, z \rangle = \langle x, \langle y, z \rangle \rangle & \text{/* associative */} \end{array}$$

Due to the associativity, $\langle x, y, z \rangle$ equally denotes either $\langle \langle x, y \rangle, z \rangle$ or $\langle x, \langle y, z \rangle \rangle$. This naturally extends to tuples of arbitrary lengths: $\langle x_1, \ldots, x_k \rangle$, $k \geq 2$.

Remark 2. If φ and ψ are two possible nonfunctional patterns such that $\varphi \subseteq \top_{s_1}$ and $\psi \subseteq \top_{s_2}$, then $\langle \varphi, \psi \rangle$ denotes the cartesian product of φ and ψ, as

$$|\langle \varphi, \psi \rangle|_\rho = \bigcup_{a \in |\varphi|_\rho, b \in |\psi|_\rho} (a, b).$$

More details on product sorts in $\mathbb{ML}$ can be found in [3].

3.3 Power Sorts

Given a sort s ($s \in \top_{\mathsf{Sort}}$), its *power sort* [4] is specified by

- a sort 2^s

$$2^s \in \top_{\mathsf{Sort}}$$

- two constant symbols, extension and intension in Σ, together with the following axioms:

$$\begin{array}{l} \forall \alpha{:}2^s.\mathsf{extension}\ \alpha \subseteq \top_s \\ X \subseteq \top_s \to \exists \alpha{:}2^s.\mathsf{extension}\ \alpha = X \\ \forall \alpha{:}2^s.\forall \beta{:}2^s.\mathsf{extension}\ \alpha = \mathsf{extension}\ \beta \to \alpha = \beta \\ \mathsf{intension}\ \varphi :\leftrightarrow \exists \alpha{:}2^s.\alpha \wedge (\mathsf{extension}\ \alpha = \varphi) \end{array}$$

Remark 3. If $\varphi \subseteq \top_s$, then

$$\begin{aligned}\textsf{extension intension}\ \varphi &:\leftrightarrow \textsf{extension}\ (\exists\alpha{:}2^s.\alpha \wedge (\textsf{extension}\ \alpha = \varphi))\\ &= \exists\alpha{:}2^s.\textsf{extension}\ (\alpha \wedge (\textsf{extension}\ \alpha = \varphi))\\ &= \exists\alpha{:}2^s.\textsf{extension}\ \alpha \wedge (\textsf{extension}\ \alpha = \varphi)\\ &= \varphi.\end{aligned}$$

More details on power sorts in $\mathbb{ML}$ can be found in [4].

4 An $\mathbb{ML}$ Theory of Quotient Sorts

4.1 Relations

A relation $R \subseteq \top_{s_1} \times \top_{s_2}$ is specified by an element $r : 2^{s_1 \otimes s_2}$ such that $\textsf{extension}\ r = R$. The inverse R^{-1} of a relation $R \subseteq \top_s \times \top_s$ is specified by a constant symbol $\_^{-1}$, a notation

$$r^{-1} :\leftrightarrow \_^{-1}\ r \qquad \text{(Ntn.1)}$$

and the domain specific axioms

$$\forall r : 2^{s\otimes s}.\exists r' : 2^{s\otimes s}.r^{-1} = r' \qquad \text{(Ax.2)}$$

$$\forall r : 2^{s\otimes s}.\textsf{extension}\ r^{-1} = \exists x, y{:}s.\langle y, x\rangle \wedge \langle x, y\rangle \in \textsf{extension}\ r \qquad \text{(Ax.3)}$$

The composition $R_1 \circ R_2$ of two relations $R_1 \subseteq \top_{s_1} \times \top_{s_2}$ and $R_2 \subseteq \top_{s_2} \times \top_{s_3}$ is specified by a constant symbol $\_\circ\_$, a notation

$$r_1 \circ r_2 :\leftrightarrow \_\circ\_\ r_1\ r_2 \qquad \text{(Ntn.2)}$$

and the domain specific axioms

$$\forall r_1 : 2^{s_1\otimes s_2}.\forall r_2 : 2^{s_2\otimes s_3}.\exists r : 2^{s_1\otimes s_3}.r_1 \circ r_2 = r \qquad \text{(Ax.4)}$$

$$\begin{aligned}&\forall r_1 : 2^{s_1\otimes s_2}.\forall r_2 : 2^{s_2\otimes s_3}.\\ &\quad \textsf{extension}\ r_1 \circ r_2 = \exists x{:}s_1.\exists y{:}s_2.\exists z{:}s_3.\langle x, z\rangle \wedge\\ &\qquad\qquad \langle x, y\rangle \in \textsf{extension}\ r_1 \wedge \langle y, z\rangle \in \textsf{extension}\ r_2\end{aligned} \qquad \text{(Ax.5)}$$

4.2 Function Sorts

Given two sorts s_1, s_2 ($s_1, s_2 \in \top_{\mathsf{Sort}}$), the functions $\top_{s_1} \to \top_{s_2}$ are specified by their graphs, using

- a *function sort* $[s_1{\to}s_2]$, which is a subort of $2^{s_1 \otimes s_2}$:

$$[s_1{\to}s_2] \in \top_{\mathsf{Sort}}$$
$$\top_{[s_1\to s_2]} \subseteq \top_{2^{s_1\otimes s_2}}$$

- and the domain specific axioms:

$$\forall f{:}[s_1{\to}s_2].\forall x{:}s_1.\exists y{:}s_2.\langle x, y\rangle \in \mathsf{extension}\ f \tag{Ax.6}$$

$$\forall f{:}[s_1{\to}s_2].\forall x{:}s_1.\forall y_1, y_2{:}s_2.$$
$$(\langle x, y_1\rangle \in \mathsf{extension}\ f \wedge \langle x, y_2\rangle \in \mathsf{extension}\ f) \to y_1 = y_2 \tag{Ax.7}$$

$$\forall f{:}[s_1{\to}s_2].\forall x{:}s_1.f\ x = \exists y.y \wedge \langle x, y\rangle \in \mathsf{extension}\ f \tag{Ax.8}$$

By (Ax.6) and (Ax.7), the pattern $y \wedge \langle x, y\rangle \in \mathsf{extension}\ f$ is singleton and hence f behaves as a function and $\mathsf{extension}\ f$ denotes its graph.

4.3 Quotient Sorts

Given a sort s ($s \in \top_{\mathsf{Sort}}$) and an equivalence relation $\sim$ on s:

$$\forall x, y{:}s.(x \sim y) :\leftrightarrow (\sim x\ y) \qquad \text{/* infix notation */}$$
$$\forall x, y{:}s.((x \sim y) = \bot) \vee ((x \sim y) = \top) \qquad \text{/* relation */}$$
$$\forall x{:}s.x \sim x \qquad \text{/* reflexive */}$$
$$\forall x, y{:}s.(x \sim y) \to (y \sim x) \qquad \text{/* symmetric */}$$
$$\forall x, y, z{:}s.(x \sim y) \wedge (y \sim z) \to (x \sim z) \qquad \text{/* transitive */}$$

the *quotient sort* is specified by

- a sort $s/_\sim$, which is a subsort of 2^s:

$$s/_\sim \in \top_{\mathsf{Sort}} \tag{Ax.9}$$

- a constant symbol $[\_]_\sim$ together with a notation

$$[x]_\sim :\leftrightarrow [\_]_\sim\ x \tag{Ntn.3}$$

- and the domain specific axioms:

$$\forall x{:}s.[x]_\sim \in \top_{s/\sim} \tag{Ax.10}$$
$$\top_{s/\sim} = \exists x{:}s.[x]_\sim \tag{Ax.11}$$
$$\forall x{:}s.\mathsf{extension}\ [x]_\sim = \exists y{:}s.y \wedge x \sim y \tag{Ax.12}$$

(Ax.10) and (Ax.11) imply $\top_{s/\sim} \subseteq \top_{2^s}$. Moreover, $\mathsf{extension}\ [x]_\sim \subseteq \top_s$ uniquely identifies the equivalence class of x:

Proposition 1. $\forall x, y{:}s.(x \sim y) \leftrightarrow ([x]_\sim = [y]_\sim)$

Proof. $(x \sim y) \leftrightarrow (\mathsf{extension}\ [x]_\sim = \mathsf{extension}\ [y]_\sim) \leftrightarrow ([x]_\sim = [y]_\sim)$. □

The functions that do not depend on the class representatives can be extended to quotient sort:

Proposition 2.

$$\forall s'{:}\mathit{Sort}.\forall f : [s{\to}s'].(\forall x, y{:}s.x \sim y \to (f\ x = f\ y)) \to \\ \exists \bar{f}{:}[(s/_\sim){\to}s'].(\forall x{:}s.\bar{f}\ [x]_\sim = f\ x)$$

Proof. Let $\bar{f} = \mathsf{intension}\ (\exists x{:}s.\exists z{:}s'.\langle [x]_\sim, z\rangle \wedge \langle x, z\rangle \in \mathsf{extension}\ f)$ s.t. $\bar{f}\ [x]_\sim = \exists z{:}s'.z \wedge \langle [x]_\sim, z\rangle \in \mathsf{extension}\ \bar{f}$. We have to prove (Ax.6)–(Ax.7) hold for $\bar{f}$:

- $\forall \alpha{:}s/_\sim.\exists z{:}s'.\langle \alpha, z\rangle \in \mathsf{extension}\ \bar{f}$. If $\alpha{:}s/_\sim$ then there is $x{:}s$ s.t. $\alpha = [x]_\sim$ by (Ax.11). Consider z such that $\langle x, z\rangle \in \mathsf{extension}\ f$;

- if $\langle [x]_\sim, z_1\rangle, \langle [x]_\sim, z_2\rangle \in \mathsf{extension}\ \bar{f}$ then we get $\langle x, z_1\rangle, \langle x, z_2\rangle \in \mathsf{extension}\ f$ by the definition of $\bar{f}$ and we get $z_1 = z_2$ by (Ax.7) for f;

and that $\bar{f}\ [x]_\sim = f\ x$:

$$\begin{aligned} \bar{f}\ [x]_\sim &= \exists z{:}s'.z \wedge \langle [x]_\sim, z\rangle \in \mathsf{extension}\ \bar{f} && \text{by def. of } \bar{f} \\ &= \exists z{:}s'.z \wedge \langle x, z\rangle \in \mathsf{extension}\ f && \text{by def. of } \bar{f} \\ &= f\ x && \text{by (Ax.8) for } f \end{aligned}$$

□

The functions defined over quotient sort can be defined by representatives:

Proposition 3.

$$\forall s'{:}\mathit{Sort}.\forall \bar{f} : [(s/_\sim){\to}s'].\exists f{:}[s{\to}s'].(\forall x{:}s.\bar{f}\ [x]_\sim = f\ x)$$

Proof. Let $f = \mathsf{intension}\ (\exists x{:}s.\exists z{:}s'.\langle x, z\rangle \wedge \langle [x]_\sim, z\rangle \in \mathsf{extension}\ \bar{f})$ s.t. $f\ x = \exists z{:}s'.z \wedge \langle x, z\rangle \in \mathsf{extension}\ f$. We have to prove (Ax.6)–(Ax.7) for f:

- $\forall x{:}s.\exists z{:}s'.\langle x, z\rangle \in \mathsf{extension}\ f$ – take z such that $\langle [x]_\sim, z\rangle \in \mathsf{extension}\ \bar{f}$;
- if $\langle x, z_1\rangle, \langle x, z_2\rangle \in \mathsf{extension}\ f$ then we obtain $z_1 = z_2$ because $\langle [x]_\sim, z_1\rangle, \langle [x]_\sim, z_2\rangle \in \mathsf{extension}\ \bar{f}$;

and that $f\ x = \bar{f}\ [x]_\sim$:

$$\begin{aligned} f\ x &= \exists z{:}s'.z \wedge \langle x, z\rangle \in \mathsf{extension}\ f && \text{by def. of } f \\ &= \exists z{:}s'.z \wedge \langle [x]_\sim, z\rangle \in \mathsf{extension}\ \bar{f} && \text{by def. of } f \\ &= \bar{f}\ [x]_\sim && \text{by (Ax.8) for } \bar{f} \end{aligned}$$

□

Remark 4. If $\sim$ is the $\mathbb{ML}$ equality $=$, then s and $s/_\sim$ represent the same sort.

4.3.1 Definitional Equality and Equality Modulo Axioms

Given a sort s, we may have two kinds of equalities on s:

Definitional equality: $x =_s y :\leftrightarrow (x \in \top_s \wedge y \wedge \top_s \wedge x = y)$, where $\_{=}\_$ is the $\mathbb{ML}$ equality. Two terms are equal only if they compute to the same form.

Equality modulo axioms: $x \cong_s y$, where $\_{\cong_s}\_$ is the congruence given by axioms. For instance, this is the case of sorts for which some constructors satisfy axioms like associativity, commutativity, unity, etc. We omit the subscript s whenever it is understood from the context. Since $\_{\cong_s}\_$ is a congruence, we may consider the quotient sort $s/_{\cong}$, whose carrier set corresponds to the equivalence classes. We use equality modulo axioms whenever matching modulo this equality is needed.

Example 2. The *equality modulo axioms* $\cong_{\mathsf{List}\langle s\rangle}$ is specified by a constant symbol $\mathsf{Eq}_{\mathsf{List}\langle s\rangle}$, a notation

$$x \cong_{\mathsf{List}\langle s\rangle} y := \langle x, y\rangle \in \mathsf{Eq}_{\mathsf{List}\langle s\rangle}$$

and the axioms:

$$\begin{array}{ll}
\mathsf{Eq}_{\mathsf{List}\langle s\rangle} = & \\
\mu R{:}\mathsf{List}\langle s\rangle{\otimes}\mathsf{List}\langle s\rangle.\ \cong_s \vee & /*\ \cong_s \subseteq \cong_{\mathsf{List}\langle s\rangle}\ */ \\
\quad \exists x, y, z{:}\mathsf{List}\langle s\rangle.\langle (x \cdot y) \cdot z, x \cdot (y \cdot z)\rangle \vee & /*\ \text{associativity}\ */ \\
\quad \exists x{:}\mathsf{List}\langle s\rangle.\langle x \cdot \mathsf{nil}, x\rangle \vee & /*\ \text{right unit}\ */ \\
\quad \exists y{:}\mathsf{List}\langle s\rangle.\langle \mathsf{nil} \cdot y, y\rangle \vee & /*\ \text{left unit}\ */ \\
\quad \exists x{:}\mathsf{List}\langle s\rangle.\langle x, x\rangle \vee & /*\ \text{reflexive}\ */ \\
\quad (\mathsf{intension}\ R)^{-1} \vee & /*\ \text{symmetric}\ */ \\
\quad (\mathsf{intension}\ R) \circ (\mathsf{intension}\ R) \vee & /*\ \text{transitive}\ */ \\
\quad \exists x_1, x_2, y_1, y_2{:}\mathsf{List}\langle s\rangle.\langle x_1 \cdot y_1, x_2 \cdot y_2\rangle & \\
\qquad\qquad \wedge \langle x_1, x_2\rangle \in R \wedge \langle y_1, y_2\rangle \in R & /*\ \text{congruence}\ */
\end{array} \tag{Ax.13}$$

The quotient sort $\mathsf{List}\langle s\rangle/_{\cong_{\mathsf{List}\langle s\rangle}}$ is the *sort of lists modulo axioms*, which is used in matching modulo axioms:

$$
\begin{aligned}
&// \text{ the sort of lists modulo axioms is a sort}\\
&\forall s{:}\mathsf{Sort}.\exists s'{:}\mathsf{Sort}.\mathsf{List}\langle s\rangle/_{\cong_{\mathsf{List}\langle s\rangle}} = s'\\
&// \text{ the carrier set of the sort of lists modulo axioms}\\
&\top_{\mathsf{List}\langle s\rangle/_{\cong_{\mathsf{List}\langle s\rangle}}} = \exists x{:}\mathsf{List}\langle s\rangle.[x]_{\cong_{\mathsf{List}\langle s\rangle}}
\end{aligned}
$$

4.3.2 Canonical Forms

Equality modulo axioms can be reduced to definitional equality using the canonical forms $\downarrow x$, specified by:

- a constant symbol $\downarrow$ and the axioms:

$$
\begin{aligned}
&\forall x{:}s.\exists y{:}s.\downarrow x = y && /*\ \downarrow \text{ is functional } */\\
&\forall x{:}s.x \cong_s \downarrow x && /*\ \downarrow \text{ preserves the congruence } */\\
&\forall x, y{:}s.(x \cong_s y) \leftrightarrow (\downarrow x = \downarrow y) && /*\ \text{canonical form is unique } */
\end{aligned}
$$

Proposition 4. $\forall x, y{:}s.([x]_{\cong} = [y]_{\cong}) \leftrightarrow (\downarrow x = \downarrow y)$

Proof. $([x]_{\cong} = [y]_{\cong}) \leftrightarrow (x \cong y) \leftrightarrow (\downarrow x = \downarrow y)$. □

Example 3. The *canonical form* $\downarrow l$ of a list l is specified by a constant symbol $\downarrow$ and the notation

$$\downarrow l :\leftrightarrow (\downarrow\ l)$$

and the axioms:

$$
\begin{aligned}
&\downarrow\mathsf{nil} = \mathsf{nil}\\
&\forall x{:}s.\downarrow x = x\\
&\forall x{:}\mathsf{List}\langle s\rangle.\downarrow(x \cdot \mathsf{nil}) = \downarrow x\\
&\forall x{:}\mathsf{List}\langle s\rangle.\downarrow(\mathsf{nil} \cdot x) = \downarrow x\\
&\forall x, y, z{:}\mathsf{List}\langle s\rangle.\downarrow((x \cdot y) \cdot z) = \downarrow(x \cdot (y \cdot z))\\
&\forall x{:}s.\forall y{:}\mathsf{List}\langle s\rangle.\downarrow(x \cdot y) = x \cdot \downarrow y
\end{aligned}
$$

5 Conclusion

In this paper, we have presented a formalization of quotient sorts within matching logic. We successfully demonstrated that by leveraging the initial algebra semantics of $\mathbb{ML}$, it is possible to define quotient sorts that correctly capture the mathematical concept of equivalence classes.

Our approach systematically builds upon foundational $\mathbb{ML}$ theories of sorts, product sorts, and power sorts to construct a robust theory of relations and functions, which serves as the basis for our central definition. The resulting framework allows for specifying sorts where elements are considered equivalent under a given set of axioms, as demonstrated with the running example of associative lists.

Future research could focus on extending the framework to support the more complex quotient inductive types, and applying this methodology to larger-scale case studies in program verification and formalized mathematics.

References

[1] Evan Cavallo and Robert Harper, "Higher inductive types in cubical computational type theory," *Proc. ACM Program. Lang.*, vol. 3, no. POPL, Article No. 1, January 2019. DOI: https://doi.org/10.1145/3290314.

[2] Xiaohong Chen, Dorel Lucanu, and Grigore Roșu, "Initial algebra semantics in matching logic," Technical Report, University of Illinois at Urbana-Champaign, July 2020, http://hdl.handle.net/2142/107781.

[3] Xiaohong Chen, Dorel Lucanu, and Grigore Roșu, "Matching logic explained," *J. Log. Algebraic Methods Program.*, vol. 120, Article ID: 100638, 2021. DOI: https://doi.org/10.1016/j.jlamp.2021.100638.

[4] Xiaohong Chen and Grigore Roșu, "A general approach to define binders using matching logic," in *Proceedings of the 25th ACM*

SIGPLAN International Conference on Functional Programming (ICFP'20), vol. 4, ACM/IEEE, Aug 2020, pp. 1–32.

[5] Marcelo P. Fiore, Andrew M. Pitts, and S. C. Steenkamp, "Quotients, inductive types, and quotient inductive types," *Log. Methods Comput. Sci.*, vol. 18, no. 2, 2022.

[6] B. Nordström, K. Petersson, and J. M. Smith, "Martin-Löf's type theory," pp. 1–32. Oxford University Press, Inc., USA, 2001.

[7] Pi Squared, "Proof of Proof: Verifiable computing for all languages," 2025, https://pi2.network/papers/proof-of-proof-whitepaper, .

Dorel Lucanu

ORCID: https://orcid.org/0000-0001-8097-040X
Alexandru Ioan Cuza UNiversity of Iasi, Romania
Pi Squared Inc.
E–mail: dorel.lucanu@gmail.com

Proceedings of Workshop on Intelligent Information Systems
WIIS2025, October 16-18, 2025, Chisinau, Republic of Moldova

A Cloud-based IoMT Systems Performability Analysis Based on Stochastic Reward Net with Fuzzy Parameters

Victor Moraru, Emilian Guțuleac, Simion Cuzmin

Abstract

This paper presents a predictive approach for performability modeling and analysis of a Cloud-based Internet of Medical Things Systems based on Stochastic Reward Nets with Fuzzy parameters, which allows the evaluation of the security and availability impact, in which a proactive defense technique is used. A numerical case study for performability analysis of a particular Cloud-based Internet of Medical Things Systems is discussed to show the effectiveness of the proposed approach.

Keywords: cloud computing, fuzzy numbers, internet of medical things, performability, stochastic reward nets.

1 Introduction

In recent decades, Cloud-based Internet of Medical Things Systems (CBIoMTS) have emerged as a vital and rapidly evolving computing paradigm, driven by the advancement of smart cloud healthcare systems and wireless information technologies [1, 2]. CBIoMTS enable continuous, real-time medical monitoring and observation by utilizing wearable biosensors and health devices connected through wireless networks and remote monitoring technologies.

However, due to the sensitive nature of biomedical data and critical behavior of health systems, CBIoMTS pose various significant challenges, especially in terms of availability, reliability, and security [3, 4].

To deal with these problems, recently is proposed an emerging proactive defense approach, the Moving Target Defense (MTD) techniques, that aims to thwart attacks by dynamically changing the attack surface and disrupts the attacker's exploration phases [5], leading to complexity and unpredictability, confusing attackers by creating asymmetric uncertainties in favor of the defenders and nullify their reconnaissance effort, thus reducing the probability of success of an attack.

Performability has emerged as a key metric for evaluating a system's ability to function effectively when its performance degrades. It assesses how well the system continues to execute its specified functionalities under both normal conditions and abnormal situations, such as system failures, security threats, and intrusions [6, 7].

In this context, there is a need to evaluate and analyze the impact of using an MTD technique for CBIoMTS defense on the specified performability metrics under uncertainty by adopting and using some appropriate mathematical formalism. Amongst such formalisms, the Generalized stochastic Petri nets (GSPN) [8] and Stochastic Reward Nets (SRNs) [9] are widely used because they are conceptually easy to understand, graphical in nature, and well supported by a large body of theory as well as a large software tool base.

Most of the existing research based on GSPN or SRN models focuses only on the evaluation of MTD techniques' impact on the system security [10, 11], but not on investigating the effects considering the potential performance degradation due to the frequent reconfigurations. Nevertheless, in this type of models, the fuzzy epistemic uncertainties of the attacker's behavior are not taken into account. Also, it is necessary to enhance the SRN models in order to evaluate the uncertainty impact of MTD migration policies of computing resources in terms of specified performability metrics.

In this paper, we present a predictive approach for analytical uncertainty performability modeling and analysis of CBIoMTS using the modeling techniques of SRN with Fuzzy parameters [12], called FSRN, that are enhanced with time-based MTD techniques. The application of the proposed FSRN approach is exemplified by numerical case performability modeling and analysis of a particular CBIoMTS.

2 Behavior Informal Description of Attacked CBIoMTS

The main components of the CBIoMTS1 include the number of wireless biosensor nodes (WBNs) in a cluster, the number of portable wireless gateway devices (WGDs) [6, 13], and a cloud system with a specified Physical Machine (PM) pool. The WGDs collect data from the WBNs and send the data to the cloud VMs of a selected PM pool; it also monitors the biosensor status, changes settings, and updates software. The WGDs can communicate with WBNs simultaneously. So, due to the resource constraints of WBNs, the link failures are prone to occur. Similarly, the wireless links from the WGDs to the cloud can also be prone to failure due to their inherent vulnerability. In general, different link connection failures may need different recovery times. The system is said to be in degraded service if the number of failed links in the WBNs cluster does not exceed a value $k2$, where $1 < k2 < N$.

Also, we consider that PMs of CBIoMTS1 are subject to intrusion-based attacks. In this case, the intruder will select the target PM with the current VM and will try to identify vulnerabilities in its cloud system (*reconnaissance stage*). We consider that the attacker adopts a trial-and-error approach. If successful, he will create malware based on the discovered vulnerabilities, and he will attack the target PM, being identified as the attack surface. Then he will create a remote access weapon (for instance, malware as a virus or worm), depending on the discovered vulnerabilities, and it will deliver the malware weapon to the target through some of the most appropriate means (*delivery stage*). In this way, the malware starts working by acting on the target PM to exploit the vulnerability through: program code triggers (*exploit stage*); installing an access point for the intruder (*installation stage*); persistent access and control of the attacked PM (*command and control*); performing attacks such as exporting or destroying data, encrypting for ransom, and so on (*target actions stage*). As a result, the probabilities of a successful attack increase the processing time of the task services as long as the attacker remains in a certain PM pool.

To improve the security of CBIoMTS1 computing, the time-based MTD technique [5, 10] is used. Thus, the cloud services are running on a CBIoMTS1 that are made up of a number of PMs, which can randomly migrate between VMs of this PM pool according to this MTD technique, triggered by a timer with a probabilistic or deterministic time period. VM migration moves a VM from one PM to another. As a result, the service may migrate to another VM of the respective PM, but the attacker may remain on the same PM. In this way, the time-based MTD makes use of dynamic reconfigurations in virtualized environments to introduce uncertainties or to nullify their knowledge about the current system state.

We assume that the period times between the successive failure (resp. recovery) occurrences are exponentially distributed. Also, following the approach of work [10], we model the attack target actions stage using a four-phase Erlang-4 subnet. The reason for choosing this type of distributions is the need to approximate and preserve the attack progress even after a VM migration.

3 FSRN Security Model of an IoMS under MTD

We assume that the readers are familiar with the basic concepts, elements, and behavior rules of SRN; for more details, see [8, 9].

Next, we provide a brief presentation of the FSRN, which is derived according to [9, 12] and inherits most of the SRN characteristics and properties.

Thus, the FSRN, denoted $F\Gamma$, is defined as a 10-*tuple* system $F\Gamma = < P, T, A_{rcs}, Pri, G, Kp, M_0, \widetilde{\Lambda}, \widetilde{\omega}, \widetilde{\rho} >$, where: P (resp. T) is a finite set of *places* (resp. *transitions*), such that $T \cup P \neq \varnothing$ and $T \cap P = \varnothing$. The set of transitions is partitioned into $T = T_0 \cup T_\tau$, $T_0 \cap T_\tau = \varnothing$ so that: T_0 is a set of *immediate transitions* and T_τ is a set of *timed transitions*; $A_{rcs} = <Pre, Post, Inh, Test>$ is a set of *forward*, *backward, inhibition*, and *test* (*promoter*) arc functions with respective weight cardinalities; $Pri : T \times IN_{+}^{|P|} \rightarrow IN_{+}$ is the dynamic priority function, that controls the *firing* of *enabled* transitions. By default, $Pri(\ T_0\ ) > Pri(\ T_\tau\ )$; $Kp : P \times IN_{+}^{|P|} \rightarrow IN_{+} \cup \{\infty\}$ is the capacity bound of each place, which can contain an *integer* number of *tokens*; $G : T \times IN_{+}^{|P|} \rightarrow \{\ True,\ False\}$ is the set of *guard functions* associated with all transitions $\forall t \in T$; M_0 is the initial marking; $\widetilde{\Lambda} : T_\tau \times IN_{+}^{|P|} \rightarrow IR^{+}$ is the function that determines the $0 < \widetilde{\lambda}\,(t_j, M) < +\infty$, that is marking-dependent fuzzy firing rate of timed transitions, *enabled* by the current marking M; $\widetilde{\omega} : T_0 \times IN_{+}^{|P|} \rightarrow IR^{+}$ is the fuzzy weight function $0 \leq \widetilde{\omega}\,(t_l, M) < +\infty$, which determines the firing fuzzy probability $\widetilde{q}\,(t_l, M)$ of immediate transitions $t_l \in T_0$, therein describing a probabilistic selector; $\widetilde{\rho} : P \cup T \rightarrow IR^{+}$ is the function that determines the fuzzy function reward rates (real numbers) assigned to each current marking M and to each firing transitions $t_l \in T$; IN_{+} (resp. IR^{+}) is the set of *non negative* integer numbers (resp. positive *real*).

Figure 1 presents the graphical structural elements set of FSRN.

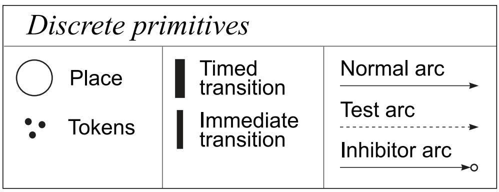


Figure 1. Graphical structural elements of FSRN

Enabling and firing rules of $t \in T$ transitions by *current marking* M in $F\Gamma$ are the same as defined in [9].

Figure 2 shows the proposed FSRN model, denoted $F\Gamma 1$, which describes the behavior of attacked CBIoMTS1 using time-based MTD defense strategies [9]. The $F\Gamma 1$ model includes following FSRN subnets: *Cloud tasks processing*; *Biosensor nodes*; *WiFi Links*, and *Time-based MTD*, which contains the *Erlang-4* subnet.

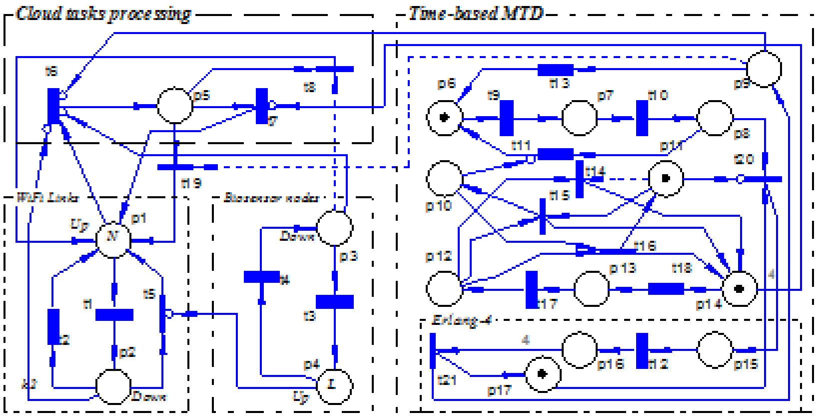


Figure 2. The $F\Gamma 1$ model of CBIoMTS1 under time-based MTD

The places and transitions meanings of $F\Gamma 1$ model:

Places. p_1 (resp. p_2) – *operational* (resp. *failed*) wireless biosensors or either links; p_3 (resp. p_4) – *failed* (resp. *operational*) gateway device or WiFi link connection; p_5 – current tasks in the service queue and can be processed; p_6 – attacker selected of the target PM; p_7 – intruder has access to attacked PM, after injecting a malware and then activating it; p_8 – intruder's malware tries to attack the system; p_9 – attack, see Figure 2. The $F\Gamma 1$ model of CBIoMTS1 under time-based MTD success, the intruder acting on the data, i.e., system is compromised; p_{10} – attacker is on the PM, where the service is running; p_{11} – intruder attack on one of the remaining (K - 1) PMs without service; p_{12} – time migration to each other different (K - 1) PMs; p_{13} – normal operating condition of service; p_{14} – start migration latency; p_{15} – the current number of Erlanger phases (*attack* stage is not over) that still need to be processed during the attack (initially represented by four tokens); p_{16} – the number of processed phases in attacked VM; p_{17} – attack stage not yet initialized.

Timed transitions. $t1$ (resp. $t2$) – mean time *failure* (resp. *re-*

covery) of biosensors nodes or either links; $t3$ (resp. $t4$) – mean time of failure (resp. *recovery*) of gateway device or link connection; $t6$ – start of task service process by VMs ; $t7$ – average arrival time of tasks to be processed by VMs; $t9$ (resp. $t10$) – the mean time required to *exploit* (resp. to *infect*) a target PM; $t11$ – Software Defined Networking (SDN) eliminates the progress of the current attack; $t12$ – period of time during which the attacker will perform desired actions, i.e., the attack is progressing in current phase; $t13$ – recovery time by SDN; $t17$ – MTD timer with delay between submitting node migration requests to the environment; $t18$ – migration delay between PMs.

Immediate transitions. $t5$ – completion of the recovery process of failed biosensors nodes or either links; $t8$ (resp. $t19$) – break tasks to the current service queue and in PM processing progress due to the link failures (resp. the successful attack of the intruder); $t14$ – service migration to one of $(K-2)$ *safe* PMs; $t15$ – service migration to a PM under attack; $t16$ – restart of MTD timer; $t19$ – breaking of service tasks under processing; $t20$ – start of the *attack* stage; $t21$ – firing of this transition denotes that the attacked PM is successfully compromised.

Firing of $t15$, with probability $q\,(t15, M) = 1/(K-1)$, leads to the migration of the service to a PM under attack or on one of $(K-2)$ *safe* PMs by firing of $t14$ with the probability $q\,(t14, M) = (K-2)/(K-1)$. In these two cases, after the service is migrated, the MTD timer starts again.

4 Numerical Performability Analysis Case Study

Next, the assessment and analysis of some performability metrics of the attacked CBIoMTS1 is performed based on the analysis of the model in two stages, following the approach described in [7, 10, 12]. The first stage is the same as the conventional SRN1 modeling, underlying $F\Gamma1$. The only difference is that the distribution of the *steady-state* probabilities π_i of the SRN1 is described in terms of the transition firing rates λ_i, i.e., which reflects the stochastic nature of the modeled CBIoMTS1. In the second stage, the transition firing fuzzy rates $\widetilde{\lambda_i}$ are represented as *triangular fuzzy numbers* (TFNs) in terms *of* $\alpha - cuts$: $\widetilde{\lambda_i}(\alpha) = [\lambda_i^-(\alpha), \lambda_i^+(\alpha)]$. After replacing the fuzzy numerical values of $\widetilde{\lambda_i}$, we obtain $\alpha - cuts$ of $\widetilde{\pi_i}$ fuzzy probabilities of $F\Gamma1$ model. According to interval arithmetic with $\alpha - cuts$ is used to calculate the specified fuzzy metrics.

The model $F\Gamma1$ is *bounded*, *live*, and *reversible* [8, 9]; therefore, there are *steady-state* behavior regimes of the analyzed CBIoMTS1.

Due to the paper volume restriction and for the sake of generality, next we shall evaluate and analyze only the *security* and *availability* metrics, because these metrics are the most relevant concerns in CBIoMTS. Using the $F\Gamma1$ model, we will present a numerical case study, which shows an application of the proposed approach.

From the structural analysis of the model, we can observe that the behavior of the *Time-based MTD* subnet is independent of the other subnets of this model. In this context, we will first analyze the security level of CBIoMTS1 based on the $F\Gamma2$ submodel, represented only by the *Time-based MTD* subnet (see Fig. 2).

We adopt an exponentially firing delay of all timed transitions with mean value $\overline{\tau_i} = 1/\lambda_i$ *hours*. The security level $\pi_{Secur.} = 1 - \pi_{Suc.Att.}$ of CBIoMTS1 is determined by the attack succes probability $\pi_{Suc.Att.} = Pr(M(p_9) = 1)$, which indicates the fact that the place p_9 of $F\Gamma2$ has 1 token. In this local state, the intruder manages to collect sensitive data.

Next, we will consider two scenarios: a) $F\Gamma2$ *submodel* of $F\Gamma1$ model to assess $\pi_{Secur.}$ depending on the λ_{17} and K parameters; b) $F\Gamma1$ model to analyze system task *throughput* services when the SDN use the time-based MTD with defined waiting strategy.

For the numerical evaluation of these metrics, based on certain parameter values taken from [4, 13], we have used the VPNP tool [14]. This evaluation aims to provide a deeper quantitative understanding of system performability (performance, availability, reliability, security, etc.) under various operating conditions, thereby validating the proposed model and guiding further optimization efforts.

The evaluation and analysis of the specified metrics will be performed considering the following values of the timed transition rates:

$\lambda_9 = 0.6$, $\lambda_{10} = 4$, $\lambda_{11} = 4$, $\lambda_{12} = 5$, $\lambda_{13} = 2$, $\lambda_{18} = 20$

Fig. 3a shows the graphical representation of the $\pi_{Suc.Att.}(\lambda_{17})$ probability variation that depends on the $\lambda_{17} \in [0.1, 2]$ parameter, where $K = 3$, and λ_{17} are considered a crisp values. For $\lambda_{17} = 0.1$ (resp. $\lambda_{17} = 2$, we obtain $\pi^{max}_{Suc.Att.}(\lambda_{17}) = 0.01596$, (resp. $\pi^{min}_{Suc.Att.}(\lambda_{17}) = 0.0061$).

In the same context, Fig. 3b plots the evolution of $\pi_{Secur.}(K, \lambda_{17})$ values that depend on the simultaneous variation of $K = 3$, and λ_{17} are considered crisp values. As computing results of $\pi_{Secur.}(K, \lambda_{17})$ for $K = 100$ and $\lambda_{17} = 0.1$ (resp. $K = 3$ and $\lambda_{17} \in [0.1, 3]$), we obtain $\pi^{max}_{Secur.}(K, \lambda_{17}) = 0.9998$ (resp. $\pi^{min}_{Secur.}(K, \lambda_{17}) = 0.9841$).

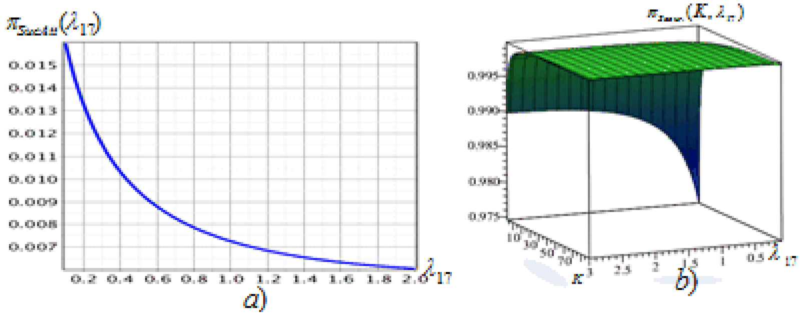


Figure 3. The evolution of: a) $\pi_{Suc.Att.}$ depending of λ_{17}; b) $\pi_{Secur.}$ depending of K and λ_{17}

The analysis of the computed results thus obtained shows that for these considered combinations of K and λ_{17} values, it is sufficient to use no more than 15 PMs. Similarly, we will evaluate the $\widetilde{\pi}_{Secur.}(\widetilde{\lambda}_9, \lambda_{17})$ fuzzy average security level system. This TFN metric was performed for the cases: $K = 3$ and $\lambda_{17} \in [0.1, 3]$ and $\widetilde{\lambda}_9 = (0.1, 0.6, 1)$. The other parameters have values that were considered in the previous case. Thus, expressing $\widetilde{\lambda}_9$ in terms of $\alpha - cut$, we have [12]: $\widetilde{\lambda}_9 = [0.6 - 0.5(1 - \alpha), 0.6 + 0.4(1 - \alpha)]$. For these parameters, evaluations of the values $[\ \pi^{-}_{Secur.}(\lambda_{17}, \alpha), \pi^{+}_{Secur.}(\lambda_{17}, \alpha)\ ]$ are obtained based on the methods presented in [6].

5 Conclusion

This paper presented a comprehensive performability modeling and analysis approach for CBIoMTS using Fuzzy Stochastic Reward Nets, a powerful tool for modeling complex systems. To enhance proactive security during task execution with Cloud PM redundancy, we employed a time-based Moving Target Defense (MTD) strategy with switch-over mechanisms for improved performability.

Future work will explore FSRN models incorporating intuitionistic fuzzy parameters and extend the proposed approach to other CBIoMTS healthcare applications, using both *time-based* and *event-based* MTD strategies. Additionally, we aim to approximate increasing probabilities using firing times of transitions with k-phase *Cox, hypo-exponential*, and/or *hyper-exponential* distributions for failure, attack, and recovery periods.

References

[1] A. Gatouillat, Y. Badr, B. Massot, and E. Sejdić, "Internet of Medical Things: A Review of Recent Contributions Dealing With Cyber-Physical Systems in Medicine," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 3810–3822, 2018.

[2] B. Guo, N. S. Ahmad Shukor, and I. S. Ishak, "Enhancing healthcare services through cloud service: a systematic review," *International Journal of Electrical and Computer Engineering*, vol. 14, no. 1, pp. 1135–1146, 2024.

[3] A. Ghubaish et al., "Recent advances in the Internet-of-Medical-Things (IoMT) systems security," *IEEE Internet of Things Journal*, vol. 8, no. 11, pp. 8707–8718, 2021.

[4] T. A. Nguyen, D. Min, E. Choi, and J. -W. Lee, "Dependability and Security Quantification of an Internet of Medical Things Infrastructure Based on Cloud-Fog-Edge Continuum for Healthcare Monitoring Using Hierarchical Models," *IEEE Internet of Things Journal*, vol. 8, no. 21, pp. 15704–15748, 2021.

[5] J. H. Cho et al., "Toward proactive, adaptive defense: A survey on moving target defense," *IEEE Communications Surveys and Tutorials*, vol. 22, no. 1, pp. 709–745, 2020.

[6] K. S. Trivedi, X. Ma, and S. Dharmaraja, "Performability modelling of wireless communication systems," *Int. J. Commun. Syst.*, vol. 16, no. 6, pp. 561–577, 2003.

[7] E. Ever, "Performability Analysis Methods for Clustered WSNs as Enabling Technology for IoT," in *Performability in Internet of Things* (EAI/Springer Innovations in Communication and Computing), F. Al-Turjman, Ed., Cham: Springer, 2019, pp. 1–19.

[8] G. Chiola, M. Ajmone-Marsan, G. Balbo, and G. Conte, "Generalized stochastic Petri nets: A definition at the net level and its implications," *IEEE Transactions on Software Engineering*, vol. 19, no. 2, pp. 89–107, 1993.

[9] G. Ciardo, A. Blakemore, P. F. Chimento, J. K. Muppala, and K. S. Trivedi, in "Automated generation and analysis of Markov reward models using stochastic reward nets," *Linear Algebra, Markov Chains, and Queueing Models* (The IMA Volumes in Mathematics and its Applications, vol 48), C. D. Meyer, R. J. Plemmons, Eds. New York: Springer, 1993, pp. 145–191.

[10] M. Torquato, P. Maciel, and M. Vieira, "Evaluation of time-based virtual machine migration as moving target defense against host-

based attacks," *Journal of Systems & Software*, vol. 219, Article No. 112222, 2025.

[11] T. A. Nguyen et al., "Performability evaluation of switch-over Moving Target Defence mechanisms in a Software Defined Networking using stochastic reward nets," *Journal of Network and Computer Applications*, vol. 199, Article No. 103267, 2022.

[12] F. Tuysuz and C. Kahraman, "Modeling a flexible manufacturing cell using stochastic Petri nets with fuzzy parameters," *Expert Systems with Applications*, vol. 37, pp. 3910–3920, 2010.

[13] D. Koutras et al., "Security in IoMT Communications: A Survey," *Sensors*, vol. 20, Article No. 4828, 2020.

[14] E. Guţuleac, C. Boşneaga, and A. Reilean, "VPNP-Software tool for modeling and performance evaluation using generalized stochastic Petri nets," in *Proceedings of 6-th International Conference on D&AS-2002*, (Suceava, România), 2002, pp. 243–248.

Victor Moraru[1], Emilian Guțuleac[2], Simion Cuzmin[3]

[1]ORCID: https://orcid.org/0000-0002-5454-8341
Technical University of Moldova
E–mail: victor.moraru@calc.utm.md

[2]ORCID: https://orcid.org/0000-0001-6839-514X
Technical University of Moldova
E–mail: emilian.gutuleac@calc.utm.md

[3]ORCID: https://orcid.org/0009-0006-4698-0341
Technical University of Moldova
E–mail: simion.cuzmin@isa.utm.md

Demagoguery and its Automatic Recognition in Texts

Evghenia Morozova, Victoria Bobicev

Abstract

The subject of this work is demagoguery in texts. Two aspects are considered: theoretical and practical. The theoretical part explains the concept of demagoguery, provides areas of its application, and common techniques. The practical part is devoted to collecting and labeling the dataset of news articles published in Moldova in the Russian language, machine learning experiments, and the analysis of the obtained results.

Keywords: Natural Language Processing, demagoguery, corpus creation, text analysis, machine learning methods.

1 Introduction

Democracy as a system of government by the whole population or all the eligible members of a state, typically through elected representatives, is the most developed political system, but also the most fragile one. It depends on multiple factors and may lead to autocratization by various political movements through various techniques. Demagoguery is one of those techniques that is widely used to persuade the population.

Multiple researches study the phenomenon of demagoguery, especially in the political context. Skillfully used demagogic techniques aimed at the population of a country can radically change the direction of its development. Demagogy has always been an effective tool for leaders to gain popular support by exploiting emotions, prejudices, and fears [1]. However, the use of demagogy often involves distorting facts, spreading misinformation, and polarizing society [2].

2 Demagoguery Definition

Demagogy is a set of various polemical and oratorical techniques by which an interested party, misleading people, entices them to his side, manipulates public opinion, or avoids unpleasant questions. In the original interpretation, demagogy is "government of the people". Previously, all politicians were called demagogues in a neutral way. Later, largely because of the ancient Greek playwright Aristophanes, this concept acquires an ironic connotation [4]. In our days, however, it has a negative connotation [5]. Cambridge Dictionary gives the following definition of demagogy: the action or fact of winning support by exciting the emotions of ordinary people rather than by having good or morally right ideas. The terms demagog, demagogue, and demagogy appear in most dictionaries; the term demagoguery, however, appears in more recent publications [4], [5], [6], [7]. In these publications, the terms are used as synonyms, with the same meaning.

Demagogy can be [3]:

- Political and social. It is used by participants in social and political movements and activists for personal gain or as one of the tools of black propaganda [2].
- Economical. It is used to persuade people to buy projects and products, to lower or raise prices.
- Juridical. According to some studies, the language of legal documents is so complex that sometimes not only people who are far from the world of jurisprudence but also specialists themselves cannot cope with their understanding. When laws and agreements are guilty of a variety of interpretations, a real treasure trove of opportunities opens up to the demagogue [8].
- Common. It is found in daily life. Many people do not even notice how, wanting to emerge victorious in a family quarrel, they inadvertently resort to demagogy.

Although demagogy often accompanies lies, it is not always based on deception: some techniques only emphasize some aspects of truthful information and downplay others. Demagogic techniques can be classified into three types [9]:

- Without violation of logic.
- Imperceptible violation of logic.
- Without any kind of logic.

Several popular demagogic techniques are described in [9].

Appeal to the Obvious

The demagogue uses some "completely obvious" statements that seem to require no proof. This creates a false impression of him being right. Moreover, this technique can be supported by an appeal to authority and an appeal to the majority: "everyone already knows", "this has long been proven", "without a doubt, this is exactly how it is" and, of course, "it is obvious".

Appeal to the Authority

If a respected person is quoted as an argument, it may seem as if what is being said makes sense: of course, a world-famous thinker or public figure cannot be wrong. You can attribute words to some great scientist, even if they never said them: few people would dare to find out right in the middle of a conversation whether the quote is true. Also, an appeal to authority is considered to be a situation where a person cites an expert's opinion that is clearly opposed to the opinion of the majority of the named expert's colleagues [8].

Appeal to Individual

To divert attention from the topic or arguments, and to undermine the opponent's reputation, the demagogue directly attacks their personality.

False Dilemma

Demagogues may present a situation as a choice between two options, while ignoring all other alternatives. It is worth noting that demagogues rarely limit themselves to just one trick. Often, one technique flows smoothly into another; for example, an appeal to the obvious is sometimes indistinguishable from an appeal to the majority.

3 Related Work

The demagogue is the only enemy of democracy that pretends to be its friend. Over the years, this phenomenon has been actively explored in political and social science [10] [11] [3] [12], [13], [23] in order to find the solution to democracy's demagogue problem.

Recently, due to ubiquitous interconnections, demagogues obtained a new and very powerful tool: social networks. The role of social media for demagoguery was examined in [6] to explain why demagogues are so influential and how social media might be contributing to their growth. The role of modern algorithms in social media in capturing human attention and the reasons why people participated in sharing of misinformation was the focus of study in [14]. Social media's role in amplifying demagoguery was analyzed in [7].

Unfortunately, we could not find any publications about automatic analysis of demagoguery in online texts, although similar rhetorical instruments as populism, disinformation, or propaganda, have received much more attention.

BERT base model was used for populism detection in [15]. Classical supervised machine learning algorithms for text populism assessment had been used in [16]. The problem of identification of populist rhetoric in text was studied in [17]. The authors manually annotated a corpus of texts and applied transformer-based model architectures for populist text fragments detection, obtaining comparatively good results, up to 0,9 F-measure. In [18], recently developed Large Language Models (LLMs) have been tested in the identification and classification of fine-grained forms of populism. The authors concluded that a fine-tuned RoBERTa classifier vastly outperformed all new-era instruction-tuned LLMs. However, instruction-tuned LLMs exhibit greater robustness on out-of-domain data.

A survey of computational propaganda in [19] discussed methods of propaganda, its specifics, approaches to propaganda detection and their limitations. Specifically, the necessity of good quality annotated corpora and a combination of methods, including Network Analysis, machine learning, and natural language processing.

Persuasion in digital communication was studied in [20]. The work detected two key challenges in using computational algorithms to tackle persuasion misuse: the opacity of deep learning models and the absence of a theoretically grounded distinction between vicious and virtuous persuasion.

Disinformation and misinformation were studied in [21], [22], indicating that disinformation or "fake news" is a global threat to democratic societies in the modern world of social networks, and its successful prevention requires a combination of methods, such as machine learning, fact checking, and network-based methods.

4 Dataset Creation

A labeled dataset, or more precisely, a corpus of texts to train models, is needed to recognize demagoguery in texts. A corpus, in computational linguistics, is a text collection processed according to certain rules. It forms the basis for some linguistic research. In this sense, a corpus can be considered a special case of the dataset [24].

The corpus collected for the study includes 405 Russian-language texts published by both news outlets and portals (“RIA Novosti”, “Gazeta RU”, “Komsomolskaya Pravda”, “Meduza”, “The Insider”, and others) and media personalities (for example, former Russian President D. Medvedev).

There are many labeling methods, but only two are suitable for this study: multi-label classification and binary classification.

Multi-label Classification

This type allows us to assign several labels to one text at once in the markers field. In this way, we can describe each observed demagogic technique separately. For each instance of the dataset, four fields were specified: • ID, the serial number assigned to the instance; • author, a link to the source of the text, i.e., to a news publication, a social network account, etc.; • text, a field containing the text itself; • markers, a list of the marks placed. The instances ratio for the multi-label classification can be seen in Table 1.

Binary Classification

Although multi-label classification allows the model to recognize each individual technique, our corpus is too small to achieve high precision for each of the classes. Therefore, a version of the corpus with binary classification was created. For binary classification, the fields remain the same, but instead of the "markers" field, there is now only one single "marker". The ratio of instances is shown in Table 2 [26].

Table 1. Instance Ratio for the Multi-label Classification

Classes	Name	Num. of articles	Percent
0	Without demagogy	211	42
1	Appeal to the obvious	49	10
2	Appeal to the individual	127	25
3	Appeal to authority	44	9
4	Appeal to the majority	29	6
5	False dilemma	39	8

Table 2. Instance Ratio for the Binary Classification

Classes	Name	Num. of articles	Percent
0	Without demagogy	211	52
1	Demagogy was found	194	48

Both datasets can be found in the "Demagoguery-RU-" repository on user evgheniamorozova's GitHub.

Level of Agreement

Since the selected texts were annotated manually, it was necessary to ensure that the labels were not biased. The whole set of text had been annotated by one annotator. In order to verify the annotation, two external people were involved in the annotation. Sets of 50 articles and annotation instructions were given to both people. Since recognizing demagogy is difficult in itself, people were not asked to recognize any specific techniques (i.e., multi-label classification was not suggested), the annotation was binary. External annotators were required to indicate whether the text contained demagogy or not. Further in the paper, the annotators are named Person №1 and Person №2.

Person №1 had never marked texts before and, in principle, had not been involved in this kind of work, and it is not surprising that his/her assessments often did not coincide with the labeler's assessments (Table 3). However, it can be noted that the further Person №1 moved through the dataset, the more (s)he understood where demagogy was present and where it was not.

Person №2, on the other hand, had the experience of annotation and basic knowledge in the field. The contradiction between the initial labeler and Person №2 is much smaller (Table 4).

In both tables, first row and first column contain labels of the articles given by the labeler (row) and the invited person (column). Other cells of the tables contain True positive (TP) value (number of articles which are labeled with 1 by both annotators), True negative (TN) value (number of articles which are labeled with 0 by both annotators), False positive (FP) and False negative (FN), where the annotators labeled the article with different values.

Table 3. Person №1 and Labeler's Agreement

Person N2 \Labeller	0	1
0	11	11
1	2	26

Table 4. Person №2 and Labeler's Agreement

Person N2 \Labeller	0	1
0	12	2
1	1	33

We evaluated the annotation agreement using Kohen's Kappa [25], the formula is:

$$Kappa = \frac{2(TPTN - FNFP)}{(TP + PF)(FP + TN) + (TP + FN)(FN + TN)}. \tag{1}$$

In the first case, total number of texts N = 50, TP=26, TN=11, FP=2, FN=11 (from the Table 3), Kappa ≈ 0.44, a moderate agreement between labeler and Person №1. In the second case, TP=33, TN=12, FP=1, FN=2 (from the Table 4), Kappa ≈ 0.85, which is considered almost perfect agreement between labeler and Person №2.

The second agreement demonstrated that the annotation can be considered reliable. This allows us to proceed further and start experiments with machine learning.

5 Machine Learning Experiments

Having the labeled data set, we can experiment with machine learning. In total, three machine learning models were used: Logistic Regression, Naive Bayes Classifier, and k-NN Classifier.

Text Preprocessing

Before proceeding directly to the models, it is necessary to preprocess the dataset in a way that all unnecessary things that interfere with training are eliminated. Thus, it was decided to remove all numbers and punctuation. We also used the NLTK dictionary of stop words for the Russian language, which includes prepositions and some common words like “why”, “where”, “of course”, which, due to their frequent mention, our models can consider as a sign of a certain class. Next, using the stemmer from the same NLTK library, we removed the endings of words. It is worth admitting that initially we planned to perform a more advantageous lemmatization; however, the Pymorphy2 library, which contains a lemmatizer for the Russian language, was incompatible with the version of Python we used.

Logistic Regression

This statistical model searches for relationships between X and Y using a logistic function. The value acquired by the function is estimated as the probability of the original, i.e., the value under study, belonging to a certain class. We imported this model from the Sklearn library.

Naive Bayes Classifier

The Naive Bayes Classifier is a simple probabilistic model based on Bayes' theorem. It is called naive because, in this case, the features are considered strictly independent. This model is very popular in the world of machine learning because this approach allows us to significantly simplify calculations. This model we also imported from the Sklearn. The threshold value was set as 0.3.

K-NN Classifier

This classifier uses the k-NN (k-nearest neighbors) algorithm. The k-NN algorithm works as follows: the object under study is assigned the class that is most common among its nearest neighbors with an already known class. The number of nearest neighbors for the model was set as 10.

6 The Results

We used two important metrics in the evaluation of the results:

- Precision: The ratio of instances of a certain class predicted by the model that actually belong to this class.
- Recall: The proportion of elements of the class the model was able to correctly identify from all elements of this class present in the test sample.

As expected, trained by multilabeled classification, our models focused only on the two largest classes: 0 (without any kind of demagogy) and 2 (appeal to the individual).

Logistic regression model: For class 0, the recall was 0.89. This means that for the class 0 instances, the model detected only 89% in the test sample and did not recognize the remaining 11%. The precision of 0.91 shows that out of 100% of the instances assigned by the model to class 0, 91% actually belong to this class, while the remaining 9% were incorrectly identified. For class 2, the results are more interesting: 100% of texts assigned to this class actually belong to class 2 (Precision = 1), but out of 100% of class 2 instances, the model found only 67% in the test sample ((Recall = 0.67)).

Compared to the logistic regression model, the Naive Bayes Classifier showed greater recall, but lost a lot in accuracy. The classifier found 100% of the class 0 instances in the test sample, but it also added 22% of instances that did not belong to this class. The same can be said about class 2: the coverage (Precision) was 96%, but as many as 45% of the instances labeled as class 2 were labeled incorrectly.

The k-NN Classifier performed worse than the Logistic Regression, but better than Naive Bayes: for class 0, it selected only 78% of the instances from the test sample, but with a precision of 95%. For class 2, 96% of the instances were found, although the precision has decreased: only 70% of the marked texts actually belong to class 2.

The other classes, due to their low number of instances, were almost ignored by machine learning algorithms, resulting in precision around .15-0.3.

Binary classification

For binary classification models, the only metric was Accuracy. This metric shows the overall accuracy of the model across all classes. Accuracy is used when the classes are represented in equal proportions.

The Accuracy of all three models:

- Logistic Regression: 0.(8);
- Naive Bayes: 0.(8);
- kNN Classifier: 0.86.

The values range from 0.86 to 0.(8). The Logistic Regression and the Naive Bayes Classifier, oddly enough, are at the same level; the k-NN Classifier lagged a little behind. Later, we will check on new data to see if these values reflect reality, but now we will get acquainted with the twenty most significant weights for each class.

Regardless of the class, the most significant words were those related to geopolitics: "Russia", "Ukraine", "Israel", "Finland", "country".

These words themselves are neutral; their inclusion at the top of one class is only due to their insufficient representation in the other. The same can be said about other nominal entities, names: “Biden”, “Putin”, “Navalny”, “Macron”.

The significant words for class 1 also contain a “сво”: often these are just stemmed forms of the pronoun “own” (“свой”, “своего”, “своих”, etc.), but sometimes “сво” is an abbreviation for “special military operation”. This is how pro-government Russian media and Putin’s supporters call Russia’s invasion of Ukraine. Oppositional editorial offices use the term “war” instead, and that’s why the “war” ended up on the list of weights for class 0.

The pronoun “our” stands out strongly. It is not surprising: a lot of propagandists, appealing to the feelings of the audience, divide the world into two opposing halves, where on one side of the barricades are “us”, and on the other stay “they”, “our” enemies.

The significant word “obviously”, apparently, turned out to be more widespread in instances with demagogy (due to the appeal to the obvious).

Although most of the meaningful words are stylistically neutral, it is still noticeable that some of them can actually indicate belonging to one of the classes.

Working with new data

To test the models on new data, we created a separate list of 20 texts of different styles. Half of the list contains demagogy (in particular, an appeal to the individual, since our models focused on it), and the other does not.

First, we tested the models trained for multi-label classification. The logistic regression model turned out to be the best, the k-NN Classifier was slightly worse, and the Naive Bayes Classifier was trailing behind (Table 5).

Next, we assess the accuracy values for the same test sample of 20 texts for binary classification models. The Logistic Regression and the Naive Bayes Classifier achieved the same results; the k-NN Classifier is still worse (Table 6).

Table 5. The Models' Accuracy for New Data (only two significant labels)

Model	Class 0	Class 2
Logistic Regression	0.9	0.8
Naive Bayes Classifier	0.8	0.6
K-NN	0.8	0.8

Table 6. The Models' Accuracy for New Data (binary classification)

Model	Class 0	Class 1
Logistic Regression	0.9	1
Naive Bayes Classifier	0.8	1
K-NN	0.8	0.8

7 Conclusion

This work was devoted to the automated analysis of texts for demagogy. Two aspects were considered: theoretical and practical. The theoretical part explained the concept of demagogy, provided areas of its application and common techniques. The practical part was devoted to collecting and labeling the dataset, machine learning experiments, and the analysis of the obtained results.

Firstly, the most difficult stage of the work was collecting training data. Secondly, although simple models showed good results (at least for binary classification), due to the tiny training sample, their predictions are more or less accurate only for a certain type of texts, namely, for Russian-language texts dedicated to Russian domestic and foreign policy in 2023–24.

The best models require larger resources, so in today's world of machine learning, a poor, lonely researcher without good hardware and financial support is unlikely to succeed. The corporations are winning again.

However, recognizing demagogy is a logical task. Of course, this problem cannot be solved unless the machine is taught to think logically. But people, who are responsible for teaching machines, are not protected from making mistakes. How can a machine be taught to cope with this task when millions of people, with brains far more complex than any current tool, themselves become victims of demagogues and propaganda? The answer to this question has yet to be found.

References

[1] G. V. Tartygasheva, "Demagogy - a form of deformation of socio-political life," *Bulletin of the Russian State University for the Humanities. Series "Philosophy. Sociology. Art Criticism"*, vol. 46, no. 3, pp. 144–160, 2010. (in Russian)

[2] C. Kumalasari, M. David, P. N. Adi, and S. Z. Saragih, "Ethics and Responsibility in Political Demagogy: A Conceptual Review and Practical Implications," *International Journal Of Humanities Education and Social Sciences (IJHESS)*, vol. 4 no. 4, 2025. DOI: https://doi.org/10.55227/ijhess.v4i4.1520.

[3] M. L. McDonough, "The evolution of demagoguery: An updated understanding of demagogic rhetoric as interactive and ongoing," *Communication Quarterly*, vol. 66, no. 2, pp. 138–156, 2018.

[4] R. Skinnell, "Using Democracy Against Itself: Demagogic Rhetoric as an Attack on Democratic Institutions," *Rhetoric Society Quarterly*, vol. 49, no. 3, pp. 248–263, 2019.

[5] J. J. Gustainis, "Demagoguery and political rhetoric: A review of the literature," *Rhetoric Society Quarterly*, vol. 20, no. 2, pp. 155–161, 1990.

[6] A. Palsma, "Exploring Demagoguery and Political Rhetoric's Impact Through Social Media," *Honors Thesis*, 2024, 324 p. Avail-

able: https://red.library.usd.edu/honors-thesis/324. Accessed August, 15, 2025.

[7] P. Chilton, "Language and politics, politics and language: democracy and demagoguery," in *Handbook of Political Discourse*, Edward Elgar Publishing, 2023, ch. 1, pp. 6–22.

[8] S. Belov and V. Gulida, "The Language of Legal Documents: Difficulties of Understanding," *Acta Linguistica Petropolitana*, vol. 15, no.1, pp. 56–103, 2019. (in Russian)

[9] B. Katsenelenbaum, "Demagogy: an experience of classification," *Science and Life*, vol. 9, 1989. (in Russian)

[10] M. Signer, *Demagogue: The Fight to Save Democracy from Its Worst Enemies*, New York: St. Martin's Press, 2009, 272 p.

[11] S. Levitsky and D. Ziblatt, *How Democracies Die*, New York: Crown, 2018, 320 p. ISBN-10: 1524762938.

[12] E. A. Posner, *The Demagogue's Playbook: The Battle for American Democracy from the Founders to Trump*, New York: St. Martin's Press, 2020, 320 p. ISBN-10: 1250303036.

[13] R. F. de Moraes, "Demagoguery, populism, and foreign policy rhetoric: evidence from Jair Bolsonaro's tweets," *Contemporary Politics*, vol. 29, no. 2, pp. 249–275, 2022.

[14] S. Lewandowsky, "Demagoguery, technology, and cognition: addressing the threats to democracy," in *Digital Technologies and the Stakes for Representative Democracy*, 2022, pp. 83–94, Alpha Omega Publishing.

[15] J. Ulinskaitė and L. Pukelis, "Identifying populist paragraphs in text: A machine-learning approach," *arXiv:2106.03161*, 2021.

[16] J. Di Cocco and B. Monechi, "How Populist are Parties? Measuring Degrees of Populism in Party Manifestos Using Supervised

Machine Learning," *Political Analysis*, vol. 30 no. 3, pp. 311–327, 2022.

[17] C. Klamm, I. Rehbein, and S. P. Ponzetto, "Our kind of people? Detecting populist references in political debates," in *Findings of the Association for Computational Linguistics: EACL 2023*, 2023, pp. 1227–1243.

[18] I. Chalkidis, S. Brandl, and P. Aslanidis, "Identifying Fine-grained Forms of Populism in Political Discourse: A Case Study on Donald Trump's Presidential Campaigns," *arXiv:2507.19303.*

[19] G. Da San Martino, S. Cresci, A. Barrón-Cedeño, S. Yu, R. Di Pietro, and P. Nakov, "A Survey on Computational Propaganda Detection," in *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, Survey track, 2020, pp. 4826–4832.

[20] D. Bassi, S. Fomsgaard, and M. Pereira-Fariña, "Decoding persuasion: a survey on ML and NLP methods for the study of online persuasion," *Frontiers in Communication*, vol. 9, 2024.

[21] M. Hardalov, A. Arora, P. Nakov, and I. Augenstein, "A Survey on Stance Detection for Mis- and Disinformation Identification," in *Findings of the Association for Computational Linguistics: NAACL 2022*, 2022, pp. 1259–1277.

[22] R. Kozik, G. Kątek, M. Gackowska, S. Kula, J. Komorniczak, P. Ksieniewicz, A. Pawlicka, M. Pawlicki, and M. Choraś, "Towards explainable fake news detection and automated content credibility assessment: Polish internet and digital media use-case," *Neurocomputing*, vol. 608, 2024.

[23] G. Ballacci, "Ancient Demagoguery and Contemporary Populism: Conceptual Analogies and Differences in Historical Perspective," *Politics and Governance*, vol. 13, 2025.

[24] V. P. Zakharov, "Prolegomena to Corpus Linguistics," *Questions of Psycholinguistics*, vol. 28, no. 2, 2016. Available: `https://cyberleninka.ru/article/n/prolegomeny-k-korpusnoy-lingvistike/`. Accessed July, 28, 2025.

[25] A. Hayes and K. Krippendorff, "Answering the Call for a Standard Reliability Measure for Coding Data," *Communication Methods and Measures*, vol 1, no. 1, pp. 77–89, 2007.

[26] M. L. McHugh, "Interrater reliability: The kappa statistic," *Biochemia Medica*, vol. 22, no. 3, pp. 276–282, 2012.

Evghenia Morozova[1], Victoria Bobicev[2]

[1]ORCID: `https://orcid.org/0009-0003-4258-6425`
Technical University of Moldova
E–mail: `evghenia.morozova@iis.utm.md`

[2]ORCID: `https://orcid.org/0000-0003-4450-3964`
Technical University of Moldova
E–mail: `victoria.bobicev@ia.utm.md`

Proceedings of Workshop on Intelligent Information Systems
WIIS2025, October 16-18, 2025, Chisinau, Republic of Moldova

Exploratory Analysis of Dependencies and Formant Correlations in Continuous Romanian Speech

Maria Nestor

Abstract

We investigate the Pearson correlation between formant frequencies, analysing both individual and population-level patterns. Pearson correlations and Fisher z-transformations were applied to vowel data collected from both male and female speakers. Histograms and z-tests revealed consistent formant dependencies, with certain vowels, such as /a/, /i/, and /u/, showing stronger correlations than others. Our results indicate that these dependencies are robust and provide an understanding of the vowel production.

Keywords: formants, Pearson correlation, histograms, z-test, Fisher z-transformation

1 Introduction

Speech signal processing is a fundamental area in digital signal processing, with applications in speech recognition, speaker identification, and speech synthesis. One important aspect of this analysis is the study of formants, which represent the resonant frequencies of the vocal tract and provide essential information about vowel articulation. While isolated vowel analysis offers foundational understandings of formants behaviour, the continuous speech presents coarticulatory effects that influence the formants and their interactions. This article explores the correlation of formants in continuous speech, aiming to understand the articulatory specifics of vowel context in normal speech.

Audio signals (signals that humans perceive) have a frequency bandwidth from about 20 Hz to 20 kHz, with the heard bandwidth age-dependent. Speech signals are transmitted over telephone communications with a limited frequency spectrum, for example, from 100 Hz to 5 kHz or even smaller (200 Hz to 4.5 kHz), to save communication bandwidth [1]. Also, when stored, speech signal is regularly limited in frequency spectrum, for example, limited to 8 kHz or less, as the typical sampling frequencies are multiples of 8 kHz, limiting the bandwidth at half the sampling frequency (sampling theorem). These bandwidth reductions are based on the fact that speech comprehension is mainly insured by the first two formants of voiced phonemes [2],[3], F_1 and F_2, while upper formants (F_3, F_4, F_n), which have central frequencies often larger than 4 kHz, are not required for speech recognition.

Reduction of the bandwidth is accompanied in the coding process by the elimination of the phase information in the speech signal, as humans do not perceive phase in audio signals [4],[5].

In cases of limited bandwidth, while the intelligibility of speech is almost completely preserved, the signal has a lower quality than natural speech. Even when the higher formants have a lower amplitude and larger noise at the receiver end of the communication, the speech quality may be perceived as improved. The process of formant correlation was analysed in [6], [7], and [8]. These previous studies found that the lower formants, especially F_2, correlate (Pearson correlation) with higher formants for most vowels in their central part, with lower values of the correlation for */i/* and */u/*. Non-null correlation was also found for formants over entire sentences. However, the correlation is not the same for all speakers and depends on the position of the vowel in the syllable, in the word, and on neighbouring phonemes.

There are many implications of the formant correlation [6],[7],[8]. In the first place, formant correlation is a relatively new aspect in speech physiology and in physiological acoustics, with consequences not fully exploited. Second, these correlations shed light on the mechanisms of coupling of the various parts of the phonatory organ [8]. Next, there are technical uses in communication and speech compression methods [9]. A compromise between the used frequency bandwidth and speech quality has been suggested [9]. Essentially, when low-bandwidth speech signals are transmitted or stored, one could add the data on the formants' correlation, whereas this data constitutes little overhead information for reconstructing the higher formants whenever the correlation is not null. In addition to these potential uses, formant correlation may find application in emotion recognition and speaker identification.

Several studies have examined these assumptions by exploring the direct relationships and their relevance for expressive speech processing. Lee et al. [10] examined the correlation between tongue positions and format frequencies across five speaking styles in healthy female speakers. They found F_1 to be strongly linked with vertical tongue displacement, while F_2 was influenced by both horizontal and vertical axes, contradicting assumptions of its exclusive anterior correlation. Prica and Ilić [11] emphasized that reliable vowel recognition might require adding also the third formant, which significantly reduces ambiguities in classification.

This paper aims to bring more information and to validate some conclusions in the literature. Specifically, we bring new data and thoroughly analyse the distributions of the formant correlation coefficients.

The organization of the paper is as follows: Section 2 describes the data collection and preprocessing, statistical analysis, and the calculation of correlations; Section 3 presents the statistical tests and results; Section 4 concludes the study.

2 Methodology

2.1 Data Collection

The database used for this study was created based on recordings from 9 participants (2 male and 7 female), all native Romanian speakers. Each participant recorded the sentence *"Cine a făcut asta, are (un)* *ospăț gratis”* three times, producing natural, continuous speech.

The recordings were processed using the Praat [12] software to record our sentence. All recordings were made in a quiet environment, with a sampling rate of 16 kHz. We manually sampled the audio data, after which their corresponding formant values were extracted. These extracted formants were then manually reviewed and cleaned, forming the basis for the statistical and classification analyses presented further on. For each vowel, only the first two or three occurrences were retained for analysis; moreover, only the central part of the vowel was used in this analysis since it corresponds to the interval where the vowel is relatively stable.

2.2 Statistical Analysis

For each file containing the segmented vowel, we extracted and cleaned the corresponding formant listings, and computed key statistical data – the mean (denoted as μ) and the standard deviation (denoted as σ). These metrics allow us to understand the central tendency and dispersion of the acoustic signal at the vowel level. Recall that the mean is computed as:

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i, \tag{1}$$

and standard deviation is computed using the formula:

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2}. \tag{2}$$

In Tables 1 and 2, we exemplify the values of the mean and standard deviation for the two occurrences of the vowel */i/* in the sentence *”Cine a făcut asta, are (un) ospăț gratis”* – namely, i_1 in ”cine”, and i_2 in ”gratis” – based on the third recording from two different speakers, a female and a male.

The female speaker exhibits higher mean values compared to the male speaker, consistent with known anatomical differences in vocal tract length. Also, notice a reduction in both mean and standard deviation for the second occurrence of the vowel /*i*/. In order to better observe this tendency of decreasing, we computed the average of the means and the average of the standard deviation across the entire dataset. As Table 3 shows, the values of the averages tend to be lower on the second occurrence of /*i*/. This might be due to the context it occurs – at the end of the sentence, and surrounded by a plosive (t), and a fricative (s).

Table 1. Mean and Standard Deviation of Formant Values for Two Occurrences of the Vowel /*i*/ – Female speakers.

Formant	i_1		i_2	
	μ	σ	μ	σ
F_1	445.992	10.7426	457.864	11.6074
F_2	2517.37	103.177	2316.44	10.9058
F_3	3074.49	146.068	3025.32	9.3043
F_4	3761.12	429.428	4511.71	22.1591

Table 2. Mean and Standard Deviation of Formant Values for Two Occurrences of the Vowel /*i*/ – Male speakers.

Formant	i_1		i_2	
	μ	σ	μ	σ
F_1	266.006	15.5195	252.962	11.5667
F_2	1924.46	21.977	1797.09	37.6239
F_3	2386.34	30.959	2428.09	57.446
F_4	3338.28	69.3749	3733.27	73.9617

We applied a median filter over the formant values to remove the noise, as it is a good choice when dealing with impulsional noise and outliers in general. In speech processing, many acoustic features, including formant frequencies, are assumed to follow a normal distribution. This assumption is motivated by the Central Limit Theorem,

Table 3. Average of Means and Average of Standard Deviation for the Two Occurrences of the Vowel /i/ across entire dataset.

Occ.	Average of Measure	F_1	F_2	F_3	F_4
i_1	Average of means	514.505	2397.1	3110.47	4061
	Average of st. devs	74.4761	235.033	197.691	208.21
i_2	Average of means	504.181	2271.44	3006.2	4261.27
	Average of st. devs	91.9702	97.7007	117.489	157.714

which states that the sum (or the average) of a large number of independent random variables, regardless of their distributions, tends toward a normal distribution. In speech, this can be justified by the presence of many independent physiological factors that linearly contribute to the observed acoustic measures. Given this assumption, outliers can be defined as values that deviate significantly from the mean. We removed the outliers using a threshold of three standard deviations from the mean ($\mu \pm 3\sigma$), as values beyond this range are statistically rare and may indicate anomalies or measurement errors (the probability for outliers to occur is around 3%).

2.3 Correlation Analysis

In speech analysis, it is important to understand how vowels manifest in different phonetic and contextual conditions. To explore potential relationships among them, we computed the Pearson correlation matrix between their formant values, allowing us to determine some degrees of association between the formant frequencies. "The coefficient of correlation r is a measure of strength of the linear relationship between two variables, x and y" [13]. It is computed as follows:

$$r = \frac{\frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sigma_x^2} \cdot \sqrt{\sigma_y^2}}, \tag{3}$$

where:

- $\bar{x}$, $\bar{y}$ are the means of x, respectively y
- σ_x, σ_y are the standard deviations of x and y

3 Results

The first results are centralized in a symmetrical matrix called *correlation matrix*. Table 4 presents the mean (μ) and standard deviation (σ) of the correlation coefficients for each vowel. In all cases, the mean and values differ from zero, suggesting a stable relationship between formants, although in several cases the mean correlation coefficients are very low in absolute.

Table 4. Mean & Standard Deviation of Correlation Values, All Vowels

Vowel	Met.	c_{12}	c_{13}	c_{14}	c_{23}	c_{24}	c_{34}
/a/	μ	0.148	0.211	-0.075	**0.335**	0.265	0.265
	σ	0.478	0.468	0.504	0.492	0.475	0.508
/ă/	μ	0.072	-0.021	0.214	0.293	0.234	**0.383**
	σ	0.528	0.506	0.506	0.484	0.508	0.452
/e/	μ	-0.091	-0.073	-0.076	**0.372**	0.209	**0.396**
	σ	0.625	0.460	0.550	0.490	0.564	0.523
/i/	μ	0.195	0.183	0.156	**0.372**	**0.324**	**0.338**
	σ	0.430	0.491	0.474	0.472	0.490	0.474
/o/	μ	-0.055	-0.066	-0.040	**0.366**	**0.391**	**0.399**
	σ	0.643	0.480	0.497	0.524	0.442	0.421
/u/	μ	**0.229**	0.177	-0.021	0.179	0.216	**0.342**
	σ	0.586	0.516	0.517	0.613	0.531	0.525

A key question that arose was whether we could test the dependency of the vowels on phonetic context using standard statistical tests such as the *t-test*, *ANOVA*, or χ^2 *test*. However, all these tests rely on assumptions that are not satisfied in our dataset. The small sample size, lack of normal distribution in formant values, and the implicit dependency among formants make the application of such tests disputable without using the Fisher transform. Moreover, the distribution of the Pearson correlation coefficient is bounded and skewed, which further suggests that it is not effective to apply these statistical tests.

For this reason, we applied Fisher transformation to the correlation values in order to bring their distribution closer to a Gaussian one. The formula for this transformation is:

$$z_r = \frac{1}{2} \cdot \ln \frac{1+r}{1-r}. \tag{4}$$

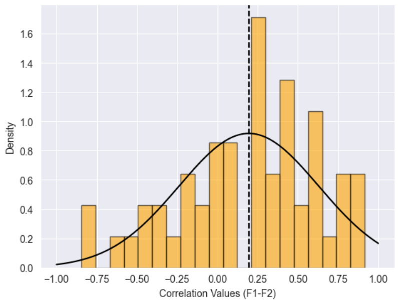
To highlight the deviation from normality in raw correlation values, Figure 1 presents the distribution of the F_1 - F_2 correlation values for both vowels /a/ and /i/. Subfigure 1a shows the histogram of the raw correlation values, which are restricted to the interval $[-1, +1]$ and display a skewed, non-Gaussian distribution. Subfigure 1b presents the corresponding distribution after applying Fisher z-transformation, which normalizes the data by transforming the correlation into an unbounded interval. In this case, the raw correlation values have a mean of $\mu = 0.195$ and a standard deviation of $\sigma = 0.430$, further confirming the non-zero correlation.

To verify if this pattern holds for other vowels as well, the same procedure was applied to /a/. Subfigures 1c and 1d show the corresponding histograms for F_1-F_2 correlations, both before and after applying Fisher z-transformation. For /a/, the correlation values have $\mu = 0.148$ and $\sigma = 0.478$ for this correlation, confirming again the presence of dependency between formants.

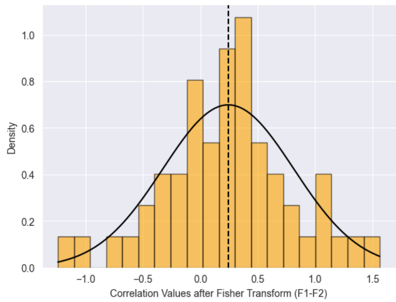
To assess whether the observed correlation values are statistically significant, we applied a z-test on the Fisher-transformed coefficients.

The hypothesis for this test is formulated as follows:

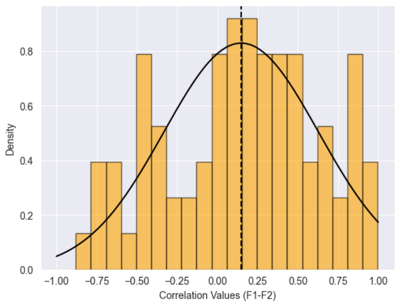
- Null hypothesis H_0: Any observed correlation between formants occurred due to random chance.



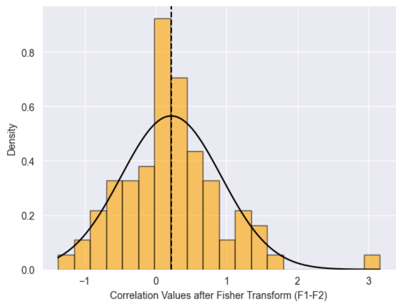
(a) Raw F_1-F_2 correlation values (Vowel /i/)



(b) Fisher z-transformed F_1-F_2 correlation values (Vowel /i/)



(c) Raw F_1-F_2 correlation values (Vowel /a/)



(d) Fisher z-transformed F_1-F_2 correlation values (Vowel /a/)

Figure 1. Distribution of F_1-F_2 correlation coefficients, before and after Fisher z-transformation

- Alternative hypothesis H_1: There is a true correlation between formants.

The z-test statistic determines if the observed correlation is statistically significant or if it occurred by chance. The formula for this statistical test is:

$$z_{stat} = \frac{z}{SE}, \tag{5}$$

where z is the value of the Fisher z-transformed correlation, and SE is the standard error, computed as

$$SE = \frac{1}{\sqrt{n-3}}, \tag{6}$$

where n is the sample size, specifically the number of signal samples that contributed to the computation of the correlation coefficient.

This test, based on the method described by Cohen et al., [14], was applied to all transformed correlation coefficients between formant pairs for each vowel. Subsequently, the percentage of statistically significant correlations at $\alpha = 0.05$ was visualised using a heatmap. The results shown in Figure 2 indicate that only a subset of the tested correlations are statistically significant, with percentages varying notably across vowels and formant pairs.

The proportion of statistically significant correlation coefficients varies across vowels and formant pairs. For instance, for vowel /a/, both the F_2-F_3 and the F_2-F_4 pairs have 70.73% significant correlations, for vowel /o/, the F_2-F_4 pair reaches 76.92%, and for vowel /ă/, the F_3-F_4 pair reaches 72.9%. In contrast, lower proportions were observed for combinations such as F_1-F_3 pair for vowel /e/ (29.63%), F_1-F_4 pair for vowel /ă/ (30.0%), and F_1-F_2 pair for vowel /o/ (34.62%). Overall, the lowest percentages were observed for the F_1-F_4 formant pair.

Across vowels, the F_2-F_3, F_2-F_4, and F_3-F_4 pairs tend to be more frequently involved in significant correlations. The F_1-F_2 pair shows a higher percentage for vowels such as /a/, /i/, and /u/, suggesting that the correlation values between formants may be influenced by the phonetic context.

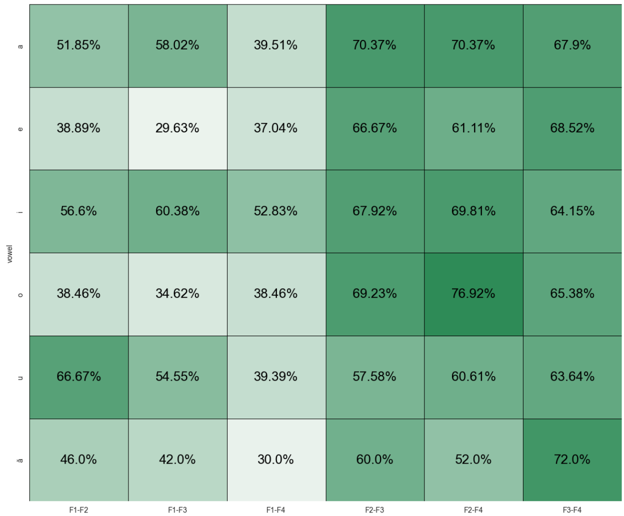


Figure 2. Heatmap showing the percentage of statistically significant formant correlations per vowel, based on z-test ($\alpha = 0.05$)

At the vowel level, /a/, /i/, and /u/ have the most favourable percentages, while /e/ and /ă/ show more variability across formant pairs, indicating lower correlations between formants.

When considered alongside the distribution of correlation coefficients, the two approaches reinforce each other. First, the distribution of correlation coefficients was examined using histograms, before and after applying Fisher z-transformation. Knowing that both the mean and standard deviation differ from zero indicates the presence of dependency between formants at the population level. The second method, which implies a statistical hypothesis test based on the Fisher z-transformation, was applied to each individual correlation value to determine its significance at the level of $\alpha = 0.05$. The percentage of significant correlation was visualised through a heatmap and quantifies how frequently correlations exceeded the significance threshold.

4 Conclusions and Discussions

Together, the results produced by the two methods support the evidence of meaningful correlation coefficients between formants – the first confirms that globally, the average correlations are non-zero, while the second shows the proportion of individual correlations that are statistically significant. This conclusion largely confirms the results in [6], [7], [8].

However, the values of the correlation coefficients between formants are not always not-null and the occurrence of not-null values is dependent on the voiced phoneme context and on the speaker.

A limitation of this study is the small number of speakers. However, as each speaker uttered three times the same sentence and the sentence includes several pronunciations of the same vowel (in different contexts), the results are significant and complete the cited papers.

Future work should use these facts for speech signal compression and for improving speech synthesis, as suggested in [6] and [11].

Acknowledgments. The author thanks Prof. Horia-Nicolai Teodorescu for guidance and support throughout this research.

References

[1] M. Plonus, "Chapter 9 - digital systems," in *Electronics and Communications for Scientists and Engineers*, M. Plonus, Ed. San Diego: Academic Press, 2001, pp. 327–403.

[2] L. R. Rabiner and R. W. Schafer, *Introduction to digital speech processing. Foundations and Trends® in Signal Processing*, vol. 1, no. 1–2, now Publishers Inc., 2007, pp. 1–194. ISBN: 978-1-60198-070-0.

[3] J. Hillenbrand and R. T. Gayvert, "Vowel classification based on fundamental frequency and formant frequencies," *Journal of Speech, Language, and Hearing Research*, vol. 36, no. 4, pp. 694–700, 1993.

[4] E. W. M. Yu and C.- F. Chan, "Harmonic+noise coding based on the characteristics of human phase perception," in *Proceedings of 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing*, 2004, pp. 310–313.

[5] M. Tong, Z. Bian, X. Li, Q. Dai, and Y. Chen, "Study on phase perception in speech," *Journal of Electronics*, vol. 20, no. 5, pp. 387–392, Sep, 2003.

[6] S. -A. Gheltu and H. -N. L. Teodorescu, "Correlations and Dependencies of the Formants During Speech in Romanian," in *2018 10th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, (Iasi, Romania), 2018, pp. 1-6, DOI: 10.1109/ECAI.2018.8678995.

[7] H. -N. Teodorescu, "How much are formants correlated and what that means?," in *2019 International Symposium on Signals, Circuits and Systems (ISSCS)*, (Iasi, Romania), 2019, pp. 1-4, DOI: 10.1109/ISSCS.2019.8801758.

[8] H. -N. Teodorescu and S. -A. Gheltu, "Analysis of formant correlations in emotional speech – do formants correlate?," in *2023 International Conference on Speech Technology and Human-Computer Dialogue (SpeD)*, 2023, pp. 83–88.

[9] H. -N. Teodorescu, Personal Communications, 2025.

[10] J. Lee, S. Shaiman, and G. Weismer, "Relationship between tongue positions and formant frequencies in female speakers," *The Journal of the Acoustical Society of America*, vol. 139, no. 1, pp. 426–440, 2016.

[11] B. Prica and S. Ilić, "Recognition of vowels in continuous speech by using formants," *Facta Universitatis Series: Electronics and Energetics*, vol. 23, no. 3, pp. 379–393, 2010.

[12] P. Boersma and D. Weenink, Praat: doing phonetics by computer [Computer program], Version 6.4.27, 2025. [Online]. Available: https://www.praat.org.

[13] J. T. McClave and T. Sincich, *Statistics*, Pearson Education, Incorporated, 2020. ISBN: 9780135820100.

[14] P. Cohen, S. G. West, and L. S. Aiken, *Applied multiple regression/correlation analysis for the behavioral sciences*, London, England: Psychology Press, 2nd ed., 1983.

Maria Nestor

ORCID: https://orcid.org/0009-0002-7468-281X
"Alexandru Ioan Cuza" University, Faculty of Computer Science, Iasi, Romania
E–mail: nesmariaes@gmail.com

Serverless GPT-Powered Tutoring for Romanian High-School Math

Florin Olariu, Tudor Pasat

Abstract

Romanian lower-and upper-secondary schools face overcrowding and understaffing, as confirmed by Eurostat and OECD data. To address this, a cloud-native Intelligent Tutoring System (ITS) was developed using GPT-4o-mini and a serverless AWS stack (Lambda, API Gateway, Cognito, DynamoDB, SNS, S3). The Flutter front-end supports sign-in, adaptive practice, feedback, and analytics. Based on mastery-learning theory, the system generates math problem sets for twelfth-grade students using a 60-20-20 rule. Though currently limited to one subject and tested only in controlled settings, the architecture enables multi-disciplinary expansion and teacher authoring tools, offering a scalable solution to Romania's personalized-feedback gap.

Keywords: adaptive learning, intelligent tutoring systems, large language models, mastery learning, serverless computing.

1 Introduction

Romanian secondary education faces an acute capacity gap. In 2022, the student-to-teacher ratio reached 18.5:1, more than five points above the European Union average [1], while public spending on education stagnated at 2.9% of GDP, the lowest share in the European Union [2]. The 2022 PISA survey further reports that 13% of students attend schools where principals cite staff shortages as a direct obstacle to learning [3]. Under such conditions, teachers have little time for the individual feedback that mastery-learning research deems essential. To address this feedback deficit, we propose a cloud-native Intelligent Tutoring System (ITS) that combines GPT-4o-mini [4] with a fully serverless AWS stack (API Gateway [5], Lambda [6], Cognito [7], DynamoDB [8], S3 [9], and SNS [10]). Defined entirely as Infrastructure-as-Code and surfaced through a cross-platform Flutter client, the system delivers adaptive exercises, natural language explanations, and longitudinal analytics without requiring additional hardware or maintenance overhead for schools. This research contributes a reproducible, globally scalable serverless architecture that can be spun up locally via LocalStack and (ii) an adaptive mastery-learning engine that applies a 60-20-20 rule – 60% items targeting weaknesses, 20% reinforcing mastered skills, and 20% introducing new material – implemented solely through prompt engineering, eliminating the need for proprietary training data.

2 Methodology

The project follows a design-based research cycle that blends secondary analysis, rapid prototyping, and empirical measurement. We began with a literature and data survey (Eurostat, OECD) to confirm the scale of overcrowding and teacher shortages and to extract functional requirements from Romanian stakeholders and the grade-12 mathematics curriculum. Guided by those requirements, we iteratively built a fully serverless AWS stack, first on LocalStack for rapid local deployment, then in the cloud, refining each iteration through load tests and user-experience walkthroughs. At every cycle, we compared multiple LLM configurations in an A/B fashion and adjusted prompts until they produced coherent diagnostics, explanations, and LaTeX-formatted items. A ten-item diagnostic pretest, mapped to 64 competencies, establishes each learner's baseline. Subsequent practice sets adhere to a mastery-learning rule, comprising 60% weak items, 20% consolidated skills, and 20% stretch material. Cost per session and prompt coherence serve as the primary technical metrics, while pedagogical soundness is assessed qualitatively through a review of generated content.

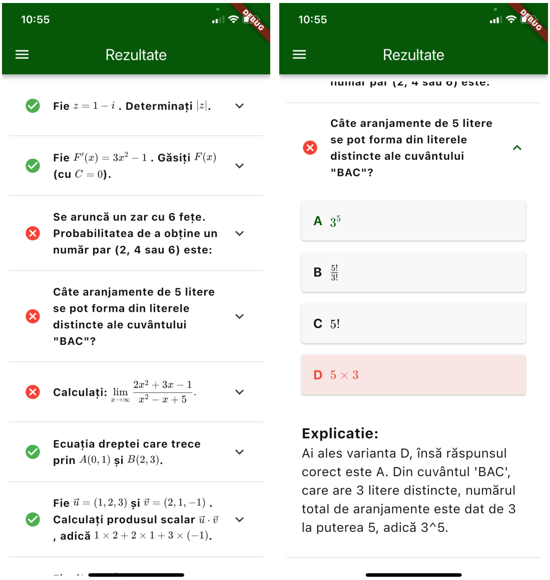


Figure 1. Automated Evaluation View

3 USE CASES

3.1 Automated Evaluation and Feedback

Learners submit answers via a Flutter client; the data is sent through API Gateway [5] to AWS Lambda [6], which builds a prompt including the learner's responses, skill profile, and performance history. GPT-4o-mini [4] returns a JSON with incorrect answers and explanations. Lambda updates the profile, verifies correctness, and publishes results to SNS [10] for database storage, enabling real-time feedback and traceability without teacher grading.

The student profile feeds the adaptive engine described in the next section; every new answer, therefore, refines the learner model and immediately influences the following exercise set. By delegating evaluation to the language model, the platform eliminates routine grading for teachers while preserving traceability: the whole exercise/response pair is stored for later audit, allowing educators to intervene if necessary.

3.2 Adaptive Content Generation

As we depicted from Figure 2, after profile updates, a new problem set is requested from GPT-4o-mini [4] via Lambda [6], using a 60-20-20 rule: 60% weaknesses, 20% mastered skills, 20% new material. This distribution is grounded in mastery-learning research [11], self-efficacy studies [12], [13], and desirable difficulty literature [14]. Lambda vali-

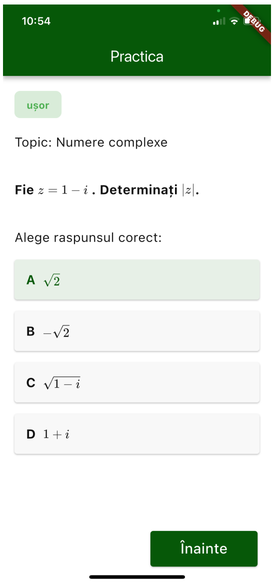


Figure 2. Adaptive Content Generation View

dates the output, assigns UUIDs, stores the content, and enforces syllabus boundaries, Romanian language, LaTeX formatting, and JSON structure.

3.3 Student Performance Dashboard

The dashboard includes:
- A version selector for timestamped snapshots.
- A GPT-generated summary of strengths and weaknesses.
- A list of weak skills shown as colored chips.
- A skill map with color-coded progress bars

$$\text{green} \geq 80\%, \quad \text{yellow} \in [40\%, 79\%], \quad \text{gray} < 40\%.$$

Data streaming is optimized to reduce latency and bandwidth (Figure 3).

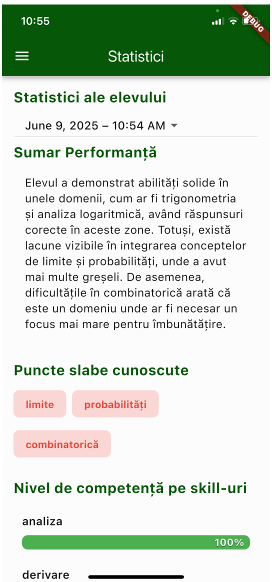


Figure 3. Student Performance Dashboard View

4 System architecture

Traditional homework platforms present identical worksheets to every learner. We aim to deliver adaptive practice that meets each student at their current mastery level. That led to three design goals: instant scalability – handle exam-week spikes without pre-provisioning servers, pay-per-use cost model – when no one studies at 3 a.m., the bill is almost 0 dollars, and LLM integration – use GPT to craft, explain, and grade personalized exercises.

4.1 System overview

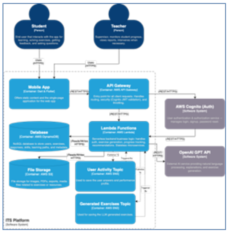


Figure 4. Container-level architecture

The platform is 100% serverless on AWS: Flutter Mobile App – one code base for Android, iOS, and web, API Gateway – public entry point; injects Cognito JWT claims and enforces throttling, Lambda Functions – stateless Python 3.11 functions, each with less than 300 ms cold start, DynamoDB – single-digit-millisecond NoSQL store, S3 – object storage for PDFs, images, and other rich media, Cognito – user auth and token management, and OpenAI GPT API – generates questions and explanations.

4.2 Backend Workflow

One practice session works like this:

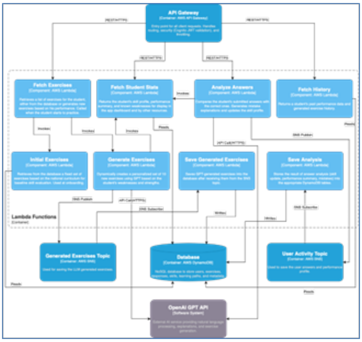


Figure 5. Backend component workflow—solid arrows = synchronous, dashed = async via SNS

- FetchExercises returns either the ten-item diagnostic set (on the first login) or an adaptive batch tailored to the student's profile.
- Student submits answers → AnalyzeAnswers checks local correctness, calls GPT for explanations, and publishes a graded payload to an SNS topic.
- SaveAnalysis (SNS subscriber) writes the graded attempt and updates the learner's skill map.
- GenerateExercises (also SNS-triggered) asks GPT for the next personalized set; SaveGeneratedExercises stores it in DynamoDB.

4.3 Data Model

Each operation fetches items by key or a pre-declared secondary index – no joins, perfect for DynamoDB's latency profile.

Table	Primary Key	Essential Attributes
Users	Id	name, role, email (Cognito UUID)
StudentMemory	studentId	summary, weaknesses [], skillsMap
StudentResponses	id	exerciseId, answer, correct?, timestamp
Exercises	id	topic, difficulty, options {}, correctKey

Figure 6. Core DynamoDB table

4.4 Infrastructure as Code

All cloud resources are expressed declaratively in YAML, including Amazon Cognito pools, API Gateway routes, Lambda functions, DynamoDB tables, Amazon S3 buckets, and Amazon SNS topics. A single deployment script reads the specification to provision the entire AWS stack. The same files drive a LocalStack + Docker environment for offline development, ensuring that the code you test locally behaves exactly the way it will in production.

4.5 Discussion

The architecture demonstrates that LLM-powered, personalized practice can operate at a national scale with zero servers to patch and a cost profile that public schools can afford. The following steps include trajectory-based recommendations and a teacher dashboard, both of which are built using the existing data in the database.

5 Paper limitations

Although the reported results confirm the technical feasibility of a serverless, GPT-powered tutor, three constraints circumscribe the work's generalizability. First, all latency and cost measurements were produced under a synthetic load in a controlled environment. Because no classroom trials have yet been conducted, the platform's actual influence on learning gains, learner engagement, and teacher workload remains unverified. Second, the adaptive engine is currently restricted to grade 12 mathematics in the Romanian Mathematics-Computer Science track. Extending the tutor to other subjects or grade levels will require new competency taxonomies, revised prompt templates, and subject-specific validation of generated content. Third, the entire pipeline relies on the proprietary GPT-4o-mini API, so changes in pricing, availability, or model behavior could impact operating costs and reproducibility. Fallback options, like locally hosted inference or optimized open-source models, may therefore be necessary for sustainable deployment. Addressing these three limitations, through classroom pilots, curriculum expansion, and model diversification, defines the immediate agenda for future work.

6 Conclusions

This paper introduces a cloud-native intelligent tutoring prototype that combines a GPT-based large-language model with a fully serverless AWS backend, aiming to alleviate two systemic constraints of Romanian pre-university education: overcrowded classrooms and a chronic shortage of qualified teaching staff. Through an iterative, design-based research process, we mapped the Grade 12 mathematics syllabus, built an adaptive engine that applies a 60–20–20 mix of weakness, consolidation, and stretch items, and integrated natural-language explanations that update each learner's profile in real-time. The entire workflow – implemented as Infrastructure-as-Code with Lambda, API Gateway, DynamoDB, S3, and SNS – scales elastically and preserves a pay-per-use cost model suitable for public schools. Latency under synthetic load remains below five seconds, demonstrating that personalized content, feedback, and analytics can be delivered without manual intervention or traditional server maintenance.

The current scope is intentionally narrow, focusing on a single subject and addressing the student role exclusively. Future work will expand the curriculum, introduce a dedicated teacher interface with trajectory-based analytics, incorporate more comprehensive learning resources, and conduct controlled classroom pilots to evaluate the educational impact. Despite these limitations, the results suggest that coupling large-language models with serverless cloud services presents a viable and cost-effective approach to personalized feedback in under-resourced school systems.

References

[1] Eurostat, "Primary education statistics," *Statistics Explained*, Available: `https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Primary_education_statistics`.

[2] Eurostat, "Educational expenditure statistics," *Statistics Explained*, Available: `https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Educational_expenditure_statistics`.

[3] OECD, "PISA: Programme for International Student Assessment," Available: `https://www.oecd.org/en/about/programmes/pisa.html`.

[4] OpenAI, "GPT-4o mini: advancing cost-efficient intelligence," Available: `https://openai.com/index/gpt-4o-mini-advancing-cost-efficient-intelligence`.

[5] Amazon Web Services, "What is Amazon API Gateway?" Available: `https://docs.aws.amazon.com/apigateway/latest/developerguide/welcome.html`.

[6] Amazon Web Services, "What is AWS Lambda?" Available: `https://docs.aws.amazon.com/lambda/latest/dg/welcome.html`.

[7] Amazon Web Services, "Amazon Cognito user pools," Available: `https://docs.aws.amazon.com/cognito/latest/developerguide/cognito-user-pools.html`.

[8] Amazon Web Services, "What is Amazon DynamoDB?" Available: https://docs.aws.amazon.com/amazondynamodb/latest/developerguide/Introduction.html.

[9] Amazon Web Services, "What is Amazon S3?" Available: https://docs.aws.amazon.com/AmazonS3/latest/userguide/Welcome.html.

[10] Amazon Web Services, "What is Amazon SNS?" Available: https://docs.aws.amazon.com/sns/latest/dg/welcome.html.

[11] M. Winget and A. M. Persky, "A Practical Review of Mastery Learning," *Am. J. Pharm. Educ.*, vol. 86, no. 10, Dec. 2022, DOI: 10.5688/ajpe8906.

[12] A. Bandura, "Self-efficacy: Toward a unifying theory of behavioral change," *Psychol. Rev.*, vol. 84, no. 2, pp. 191–215, 1977, DOI: 10.1037/0033-295X.84.2.191.

[13] B. Rosenshine, "Principles of Instruction," *American Educator*, 2012.

[14] R. A. Bjork and E. L. Bjork, "Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning," in *Psychology and the Real World: Essays Illustrating Fundamental Contributions to Society*, Worth Publishers, 2011, pp. 56–64.

Florin Olariu[1], Cosmin Pasat[2]

[1]ORCID: https://orcid.org/0009-0003-8260-2020
Faculty of Computer Science, Alexandru Ioan Cuza University,
Strada General Henri Mathias Berthelot Nr. 16,
Iași 700259, România
E–mail: olariu@gmail.com

[2]ORCID: https://orcid.org/0009-0002-0091-8017
Faculty of Computer Science, Alexandru Ioan Cuza University,
Strada General Henri Mathias Berthelot Nr. 16,
Iași 700259, România
E–mail: tcosmin.pasat@gmail.com

Proceedings of Workshop on Intelligent Information Systems
WIIS2025, October 16-18, 2025, Chisinau, Republic of Moldova

A Multi-Dimensional Requirements Analysis for Developing Artificial Intelligence for Healthcare

Maria-Ecaterina Olariu, Adrian Iftene

Abstract

The integration of Artificial Intelligence (AI) in healthcare offers opportunities to improve patient care, diagnostic accuracy, and optimize treatment. However, the development of AI systems in healthcare settings requires consideration of multiple requirements concerning technical, clinical, ethical, and organizational domains. This article discusses essential requirements for developing AI systems in healthcare, found among current research. The analysis shows that data quality and standardization, infrastructure capabilities, interoperability, security and privacy protection, clinical validation, and ethical considerations are the dimensions that should be taken into account in the development phase.

Keywords: Artificial Intelligence, Healthcare, Clinical Decision Support Systems, Data Quality, Privacy, GDPR, AI Act.

1 Introduction

Integrating AI systems into healthcare has many benefits, such as better clinical decision-making, better patient outcomes, and more efficient workflows. However, putting AI technologies into use is challenging due to technical, legal, and ethical concerns. AI offers the promise of improved diagnostic accuracy and personalized treatment plans, although it is hard to solve the aforementioned challenges. The deployment of healthcare AI systems requires a comprehensive understanding of the interconnectedness of different areas and their respective requirements.

Technical considerations include data quality assurance, infrastructure scalability, and robust security frameworks to protect sensitive patient information. Clinical requirements focus on seamless integration with existing workflows, regulatory compliance with medical device standards, and maintaining high levels of diagnostic accuracy and reliability. The ethical implications of AI in healthcare include algorithmic bias, equitable access, and decision accountability. Regulatory frameworks like GDPR and the AI Act introduce complexity by imposing privacy protection and transparency in AI decision-making. This paper explores the dimensions to be followed when developing and deploying AI systems in healthcare environments.

2 Technical Requirements

2.1 Data Quality and Standardization

The quality and uniformity of the data that an AI system uses are among the most crucial components for a successful outcome. For the development and validation of an AI model, it is important to have high-quality, standardized datasets [2]. To avoid systematic bias in models, there is a need for diverse and representative data. The lack of different demographic groups, geographic regions, and clinical presentations might lead to unreliable AI models that can compromise clinical decision-making [6].

2.2 Infrastructure and Interoperability Requirements

To meet the computational demands of modern AI systems in healthcare and ensure reliable performance, infrastructure is a key factor. Cloud computing infrastructure, combined with a microservice-based architecture, provides a scalable solution, which could improve efficiency and scalability [8, 9]. Integrating into existing healthcare systems and Electronic Medical Records (EMRs) or Electronic Health Records (EHRs) without disrupting operations is another important factor to account for [1].

Interoperability represents a critical technical requirement that enables different AI systems and healthcare platforms to communicate and share information effectively. A well-defined knowledge base, modeled using Knowledge Graphs from established sources like medical guidelines and clinical literature, is essential for seamless system integration [15]. Standardized knowledge representation minimizes data gaps and ensures consistent AI performance across diverse healthcare environments and clinical contexts [15].

2.3 Security and Privacy Considerations

Healthcare AI systems require security and privacy considerations due to a lack of consensus on data ownership, confidentiality, and liability. Legal frameworks like GDPR are necessary for compliance and secure data sharing mechanisms to protect patient confidentiality. Secure data sharing and storing mechanisms for healthcare providers must be implemented, while ensuring patient confidentiality and data privacy [10].

3 Clincal Requirements

In the following section, the discussion will revolve around integration with clinical workflows, regulatory compliance, accuracy and reliability, interpretability and explainability.

The successful implementation of AI in healthcare requires seamless integration with existing clinical workflows [8], proper training for physicians, comprehensive documentation [12], and verification of user interface design to ensure systems meet functional and non-functional requirements [13].

Healthcare AI systems must follow strict rules that apply to medical devices and clinical decision support tools [14]. There are many parts to regulatory compliance, such as safety standards, efficacy requirements, and quality management systems.

High diagnostic accuracy and dependability are necessary for AI systems in healthcare, with success rates between 80 and 100 percent

[6]. Maintaining these performance levels across a range of patient populations and clinical settings requires constant validation and observation.

Explainable AI is crucial for transparency in clinical decision-making, requiring healthcare professionals to understand the reasoning behind AI-generated recommendations for informed decisions and professional accountability, especially in high-stakes clinical decisions [5].

4 Ethical Requirements

Ethical requirements for healthcare AI systems demand the implementation of unbiased algorithms across different demographic groups. Equal access to AI-powered healthcare solutions and consideration of diverse patient populations are fundamental ethical imperatives [6].

Clear lines of responsibility for AI-driven decisions must be established, along with liability frameworks for AI-related errors. Regular ethical reviews and assessments ensure ongoing compliance with ethical standards and identification of emerging concerns.

5 Regulatory Compliance and Privacy Considerations

To ensure compliance with European data protection and artificial intelligence regulations, this system, which processes patients' personal information, will adhere to the General Data Protection Regulation (GDPR) and the proposed AI Act.

5.1 General Data Protection Regulation Compliance

Regarding GDPR compliance, two key principles will be prioritized: the right to be forgotten (Chapter III, Article 17) and data minimization (Chapter II, Article 5, 1.c.). To uphold the right to be forgotten, comprehensive data deletion mechanisms will be implemented across all system components. Data minimization will be achieved by collecting only essential medical information. This necessitates the development of a standardized data collection template, automated processes for segregating essential from auxiliary data, and clearly defined, automatically enforced data retention policies (Data protection by design and by default – Article 25 [3]).

5.2 Artificial Intelligence Act Considerations

The AI Act aims to ensure the safety and security of healthcare AI, which is classified as high-risk. A robust risk management system will be implemented, monitoring outputs for errors and biases, conducting regular assessments, and establishing incident reporting protocols [4]. Technical documentation will be maintained, detailing system architecture, training methodologies, data sources, validation procedures, and risk management measures. Compliance with data protection and AI regulations will be achieved through technical solutions in data protection and access control.

5.3 Technical Implementation for Privacy Protection

Techniques such as anonymization, pseudonymization, and differential privacy will be employed to preserve privacy in AI applications. Implementation considerations will focus on patient data and system operations, including consent management systems, data portability mechanisms, comprehensive documentation, Data Protection Impact Assessments, explainable AI (XAI), user-friendly interfaces, and audit trails for system decisions.

6 Conclusions

To successfully use AI systems in healthcare, it is necessary to take into account technical, clinical, ethical, and regulatory issues simultaneously. Even though the possible benefits are substantial, implementation must prioritize data quality, system interoperability, and strong privacy protections to keep patients safe and adhere to the regulations. The analysis shows that AI systems in healthcare need to be very accurate in their diagnoses and also be transparent about how they function through explainable AI. Future advancements should concentrate on formulating standardized protocols for ethical AI implementation, improving cross-system interoperability, and creating extensive regulations that balance innovation with patient safety and regulatory obligations.

References

[1] S. Bozyel, "Artificial Intelligence-Based Clinical Decision Support Systems in Cardiovascular Diseases," *Anatol J Cardiol*, vol. 28, no. 2, pp. 74–86, 2024. DOI: `https://doi.org/10.14744/AnatolJCardiol.2023.3685`.

[2] M. A. Muzammil et al., "Artificial intelligence-enhanced electrocardiography for accurate diagnosis and management of cardiovascular diseases," *Journal of Electrocardiology*, vol. 83, pp. 30–40, Mar. 2024. DOI: `https://doi.org/10.1016/j.jelectrocard.2024.01.006`.

[3] J. Zaki, S. M. R. Islam, N. S. Alghamdi, M. Abdullah-Al-Wadud, and K.-S. Kwak, "Introducing Cloud-Assisted Micro-Service-Based Software Development Framework for Healthcare Systems," *IEEE Access*, vol. 10, pp. 33332–33348, 2022. DOI: `https://doi.org/10.1109/ACCESS.2022.3161455`.

[4] K. I. Shakhgeldyan, B. I. Geltser, and B. V. Potapenko, "Architecture of a Hybrid Clinical Decision Support System," in *Proceedings of the Seventh International Scientific Conference "Intelligent Information Technologies for Industry" (IITI'23)* (Lecture Notes in Networks and Systems, vol. 777), S. Kovalev, I. Kotenko, and A. Sukhanov, Eds. Cham: Springer Nature Switzerland, 2023, pp. 146–156. DOI: `https://doi.org/10.1007/978-3-031-43792-2`.

[5] N. Mehta, K. Born, and B. Fine, "How artificial intelligence can help us 'Choose Wisely'," *Bioelectron Med*, vol. 7, Article

No: 5, Dec. 2021. DOI: https://doi.org/10.1186/s42234-021-00066-8.

[6] P. Z. Chan, M. A. I. B. Ramli, and H. S. J. Chew, "Diagnostic Test Accuracy of artificial intelligence-assisted detection of acute coronary syndrome: A systematic review and meta-analysis," *Comput. Biol. Med.*, vol. 167, Article ID: 107636, Dec. 2023, DOI: https://doi.org/10.1016/j.compbiomed.2023.107636.

[7] D. Lee and S. N. Yoon, "Application of Artificial Intelligence-Based Technologies in the Healthcare Industry: Opportunities and Challenges," *IJERPH*, vol. 18, no. 1, Article No: 271, Jan. 2021. DOI: https://doi.org/10.3390/ijerph18010271.

[8] S. Reza-Soltani et al., "The Role of Artificial Intelligence and Machine Learning in Cardiovascular Imaging and Diagnosis," *Cureus*, vol. 16, no. 9, Article ID: e68472, Sept. 2024. DOI: https://doi.org/10.7759/cureus.68472.

[9] S. Labkoff et al., "Toward a responsible future: recommendations for AI-enabled clinical decision support," *J. Am. Med. Inform. Assoc.*, vol. 31, no. 11, pp. 2730–2739, Nov. 2024. DOI: https://doi.org/10.1093/jamia/ocae209.

[10] H. Moghaddasi, F. Rahimi, A. S. Seddighi, L. Akbarpour, and A. Roshanpoor, "A decision support system to increase the compliance of diagnostic imaging examinations with imaging guidelines: focused on cerebrovascular diseases," *Diagnosis*, 2024. DOI: https://doi.org/10.1515/dx-2024-0072.

[11] B. Solaiman and M. G. Bloom, "AI, Explainability, and Safeguarding Patient Safety in Europe: Toward a Science-Focused Regulatory Model," in *The Future of Medical Device Regulation*, 1st ed., I. G. Cohen, T. Minssen, W. N. Price Ii, C. Robertson, and C. Shachar, Eds. Cambridge University Press, 2022, pp. 91–102. DOI: https://doi.org/10.1017/9781108975452.008.

[12] S. Kim, D. Kim, M. Kim, H. Ko, and O. Jeong, "XAI-based Clinical Decision Support System: A Systematic Review," June 11, 2024, *Computer Science and Mathematics*. DOI: https://doi.org/10.20944/preprints202406.0721.v1.

[13] "REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)," *Official Journal of the European Union*, Apr. 27, 2016. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679#tit_1.

[14] "REGULATION (EU) 2024/1689 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)," *Official Journal of the European Union*, June 13, 2024. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=OJ:L_202401689.

[15] S. Zheng, X. W. Xu, Z. Hu, M. Wang, Y. Yao, and J. Li, "Development and implementation of a guideline-based decision support system for etiological diagnosis of Sustained Monomorphic Ventricular Tachycardia (SMVT)," *Eur. Heart J.*, vol. 45, no. Supplement 1, Article ID: ehae666.3517, Oct. 2024. DOI: 10.1093/eurheartj/ehae666.3517.

Maria-Ecaterina Olariu[1], Adrian Iftene[2]

[1]ORCID: https://orcid.org/0009-0005-0364-6661
Alexandru Ioan Cuza University of Iasi
E–mail: mariaecaterina.olariu@gmail.com

[2]ORCID: https://orcid.org/0000-0003-3564-8440
Alexandru Ioan Cuza University of Iasi
E–mail: adrian.iftene@uaic.ro

Proceedings of Workshop on Intelligent Information Systems
WIIS2025, October 16-18, 2025, Chisinau, Republic of Moldova

Hybrid Assessment Strategies for VR Biomedical Education

Elena-Ecaterina Opait, Adrian Iftene

Abstract

Virtual reality is increasingly used in biomedical education, yet evaluating user interactions in gamified settings, especially for home use, remains an open challenge. This paper examines automated assessment methods, including structured tasks, conversational AI, and behavior-based metrics. Analysis of recent studies suggests that a hybrid strategy is optimal, balancing automated feedback with adaptive learning. Practical customization requirements for integrating this approach into existing virtual reality biomedical applications are outlined, with an emphasis on usability and enhanced student engagement.

Keywords: Virtual reality, Biomedical education, Gamified learning, Learning evaluation.

1 Introduction

1.1 Context

As virtual reality becomes more common in biomedical education [1, 2], there is growing interest in how these systems support learning beyond visual immersion [3]. Many virtual reality (VR) platforms now include interactive tasks, often designed as games, that ask learners to respond to questions. These elements are intended to keep students engaged [4, 5]. However, in practice, it remains unclear how well these interactions are being evaluated, especially outside of formal classroom settings [6, 7].

1.2 Problem

Most VR applications include some form of automated feedback, but these systems tend to rely on predefined rules or structured answers [8]. They rarely adapt to open-ended responses, and few provide meaningful insights into how well a student understands the topic. This becomes a more serious issue in gamified environments, where surface-level engagement can mask poor comprehension. In remote and home-based learning contexts, the lack of instructor oversight makes this even harder to manage the evaluation [6]. There is a need for more flexible and scalable ways to assess how students interact with educational content in these contexts.

1.3 Focus and Scope

This paper reviews existing methods for evaluating interactions in VR biomedical education. The emphasis is on automated approaches, and the goal is to identify gaps, compare strategies, and consider what a hybrid evaluation system might look like. The analysis focuses on studies that combine educational goals with real-time interaction, especially those that explore how learners respond to questions or tasks in a VR setting.

2 Methodology

2.1 Selection Criteria

This review focuses on studies published between 2018 and 2024 that examine automated evaluation strategies in virtual reality environments for biomedical education. Sources were identified using academic databases. Search terms included combinations of "VR", "biomedical education", "automated assessment", "conversational agents", "learning based on tasks", and "behavioral analytics". Only peer-reviewed papers that reported on implemented systems or conducted comparative analysis were considered. Studies that focused purely on hardware, visual fidelity, or unrelated disciplines were excluded.

2.2 Categorization Approach

To support comparison, the selected papers were grouped based on the type of evaluation method they explored. These include structured task assessments, conversational AI, behavior metrics, and hybrid approaches that combine multiple strategies. Within each category, papers were reviewed for their evaluation design, feedback mechanisms, adaptability, and relevance to remote or home-based learning contexts. The goal was to trace common patterns, highlight some approaches, and provide a basis for identifying what works best in biomedical education.

3 Review of Existing Methods

3.1 Structured Task Assessment

Many VR education platforms rely on structured tasks to assess student knowledge. These include multiple-choice questions, drag-and-drop exercises, and step-by-step simulations where each correct action is logged. This method works well for checking factual recall or following clinical procedures, and it integrates easily with existing VR frameworks. However, it often fails to capture deeper reasoning or flexible thinking in a similar manner to written tests. For example, Pottle used VR surgical modules to assess decisions through structured tasks, reporting high usability but limited adaptability to diverse responses [9]. Similarly, Moro *et al.* emphasized the effectiveness of structured evaluations in anatomy learning but noted their inability to reflect reasoning [10].

3.2 Conversational AI and Natural Language Understanding

Recent advances in NLP have made it possible to embed conversational agents in VR education. These systems evaluate student responses through natural language interactions, often simulating clinical interviews or patient consultations. Hamilton *et al.* implemented an AI standardized patient that adapts its feedback based on the learner's dialogue [12]. Haag *et al.* developed a voice AI tutor capable of interpreting student responses in context and offering follow-up questions [13]. While promising, such systems require careful training and are sensitive to speech variability and misunderstanding.

3.3 Behavioral and Performance Metrics

Behavioral evaluation looks beyond verbal or written responses. It tracks metrics such as gaze direction, task completion time, error rates, and physical interactions. Nicholson *et al.* used eye tracking and motion data to identify engagement levels during VR anatomy labs [14]. Al-Ghareeb *et al.* analyzed user interactions in home-based nursing simulations, correlating task fluidity and error frequency with learning outcomes [15]. This method allows continuous assessment but may overlook cognitive reasoning, unless combined with other techniques.

3.4 Hybrid Models

Several studies now suggest that combining different evaluation strategies leads to better outcomes. Maresky *et al.* combined decision tasks with gesture tracking in clinical VR simulations to capture both cognitive and procedural performance [16]. Birbara *et al.* layered gamified feedback with real-time conversational prompts to assess and guide student progress [19]. These systems tend to be more responsive and adaptive, though they require more development effort and tuning.

4 Key Papers and Findings

Pottle [9] evaluated the role of virtual reality in surgical training. His study emphasized task assessment, focusing on skill repetition and procedural accuracy in simulated environments. While outcomes were generally positive in terms of learner satisfaction and technical skill development, the study noted that such evaluation methods did not adequately measure decisions or deeper clinical reasoning [9].

Moro *et al.* [10] performed an analysis of immersive technologies in health education, including VR and augmented reality. The authors found widespread improvements in engagement and retention but also identified a consistent lack of in-depth evaluation frameworks. Most studies reviewed relied on pre/post-tests or self-report surveys, limiting insight into learner behavior or critical thinking processes [10].

Hamilton *et al.* [12] explored the use of conversational agents embedded in VR scenarios to train students in patient diagnosis. Their system allowed learners to engage in realistic medical interviews with AI virtual patients. The study showed promise for real-time, adaptive evaluation, particularly in recognizing diagnostic reasoning and communication skills [12].

Nicholson *et al.* [14] implemented eye tracking and user behavior monitoring within a VR anatomy lab. Their research highlighted how gaze data, focus shifts, and interaction patterns could be correlated with learning outcomes. This type of behavioral tracking provided an alternative to traditional testing, providing insight into how students engage with and navigate complex spatial content [14].

Al-Ghareeb *et al.* [15] assessed home-based nursing education using VR, focusing on behavioral and motion tracking metrics. Their study captured data such as gesture smoothness, error correction, and navigation style to infer engagement and understanding. These metrics enabled scalable assessment while preserving realism and flexibility in the absence of an actual instructor [15].

Haag *et al.* [13] developed a VR training system that used voice AI agents to simulate clinical conversations. The system responded dynamically to spoken input and adjusted feedback based on learner responses. This approach integrated both language understanding and performance monitoring, offering an early model of adaptive, speech-driven evaluation in medical simulations [13].

Maresky *et al.* [16] examined clinical decision making within a gamified VR environment. Their study assessed not only the correctness of student choices but also how those decisions unfolded through gesture input and sequence tracking. This hybrid method revealed both cognitive reasoning and physical task execution, making it suitable for complex training contexts [16].

Birbara *et al.* [19] combined real-time conversational feedback with gamified tasks in a VR anatomy course. Their mixed methods evaluation captured both objective performance data and subjective engagement measures. The study highlighted how layered dialogue and prompts initiated by the system can offer a more complete picture of student learning [19].

5 Patterns and Insights

Hybrid methods consistently outperform single-mode assessments. These approaches not only capture accuracy but also trace process, engagement, and reasoning in real time [16, 19].

Structured tasks ensure broad content coverage and are easy to implement at scale, but they lack nuance. Their binary nature limits insight into how learners approach problems or adapt to complexity [9, 10].

Conversational AI shows strong potential for assessing higher thinking and clinical communication skills [12, 13]. However, current models require substantial training data specific to the domain and are sensitive to variations in speech, language, and context. Integration into real-time systems also poses performance and latency challenges [13].

Behavioral analytics, such as gaze tracking, motion patterns, and task fluency, are promising for passive, continuous assessment, particularly in home-based learning environments [14, 15]. They offer monitoring without disrupting immersion, but often need to be triangulated with cognitive or verbal data to provide meaningful insights.

In all studies, the most effective systems are those that combine these modes. Hybrid frameworks provide both breadth and depth, capturing performance in ways that adapt to individual learners and varied tasks [16, 19].

6 Practical Implications

Scalable hybrid models require a modular design. The systems must be flexible enough to integrate multiple input types, task results, spoken responses, and behavioral data, while maintaining usability across diverse learning contexts [17, 19]. This means designing platforms that can operate on varied hardware, from VR headsets to mobile setups in remote environments [6]. Structured tasks may favor novice learners who benefit from clear goals and feedback, while more experienced users may prefer conversational or exploratory modes [10, 11]. Hybrid systems should support multiple levels of customization and allow transitions between assessment styles based on learner needs and progression [18].

Customization is particularly important in medical and biomedical curricula, where training goals vary by discipline, institution, and learner stage. Modular and hybrid systems can be configured to target specific competencies, from procedural accuracy to diagnostic reasoning, without requiring entirely new infrastructure [4, 16]. This adaptability makes them suitable for integration into existing educational systems, both in classrooms and remote [15, 6].

7 Conclusion

This review highlights the increasing value of hybrid evaluation methods in VR biomedical education. Structured assessments offer reliability and broad coverage but often miss the depth of student reasoning. Conversational AI brings flexibility and a more personalized learning experience, though it depends heavily on design and technical infrastructure [20]. Behavioral analytics contribute by tracking user engagement and performance in real time, yet they can be difficult to interpret without context [21].

The most effective systems combine these approaches. When structured tasks, natural language interactions, and behavioral insights are integrated into a unified framework, educators gain a more complete understanding of student learning. Such hybrid models can support real-time adaptation, tailored feedback, and a range of learning objectives across varied settings [22], where educators are open to using such tools [23].

To move from theory to practice, it is essential to focus on modular, adaptable systems that can be aligned with institutional goals and technical constraints.

Well-designed hybrid tools should be able to adjust to learner needs, scale across platforms, and remain usable for both novice and experienced users [24]. As immersive learning technologies evolve, these systems will play an important role in ensuring educational outcomes are meaningful.

References

[1] M. Venkatesan, H. Mohan, J. R. Ryan, C. M. Schürch, G. P. Nolan, D. H. Frakes, and A. F. Coskun, "Virtual and augmented reality for biomedical applications," *Cell Reports Medicine*, vol. 2, no. 7, Article ID: 100348, 2021.

[2] A. Singh, D. Ferry, A. Ramakrishnan, and S. Balasubramanian, "Using virtual reality in biomedical engineering education," *Journal of Biomechanical Engineering*, vol. 142, no. 11, Article ID: 111013, 2020.

[3] S. Kavanagh et al., "A systematic review of virtual reality in education," *Themes in Science and Technology Education*, vol. 10, no. 2, pp. 85–119, 2017.

[4] J. Radianti, T. A. Majchrzak, B. Fromm, and I. J. Wohlgenannt, "A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda," *Computers & Education*, vol. 147, Article ID: 103778, 2020.

[5] M. A. Minouei, A. Omid, A. Mirzaie, H. Mahdavifard, and A. Rahimi, "Effectiveness of virtual reality on medical students' academic achievement in anatomy: systematic review," *BMC Medical Education*, vol. 24, Article No. 1407, 2024.

[6] A. N. Neher, F. Bühlmann, M. Müller et al., "Virtual reality for assessment in undergraduate nursing and medical education – a systematic review," *BMC Medical Education*, vol. 25, Article No. 292, 2025.

[7] G. Makransky et al., "Equivalence of using a desktop virtual reality science simulation at home and in class," *PLOS ONE*, vol. 14, no. 4, Article ID: e0214944, 2019.

[8] S. Wijewickrema, X. Ma, P. Piromchai, R. Briggs, J. Bailey, G. Kennedy, and S. O'Leary, "Providing automated real-time technical feedback for virtual reality based surgical training: is the simpler the better?" in *Artificial Intelligence in Education: 19th International Conference, AIED 2018, London, UK, June 27–30, 2018, Proceedings, Part I*, 2018, pp. 584–598.

[9] J. Pottle, "Virtual reality and the transformation of medical education," *Future Healthcare Journal*, vol. 6, no. 3, pp. 181–185, 2019.

[10] C. Moro, Z. Phelps, and A. Redmond, "HoloLens and mobile augmented reality in medical and health science education: A systematic review," *British Journal of Educational Technology*, vol. 52, no. 1, pp. 490–510, 2021.

[11] M. Zahabi and A. M. Abdul Razak, "Adaptive virtual reality-based training: A systematic literature review and framework," *Virtual Reality*, vol. 24, no. 4, pp. 725–752, 2020.

[12] J. Hamilton et al., "Virtual standardized patients for clinical training in VR: A randomized trial," *Medical Education*, vol. 54, no. 11, pp. 1032–1043, 2020.

[13] M. Haag, L. Liu, and S. Mehmood, "Conversational agents in virtual reality medical training: A voice-based approach," *International Journal of Medical Informatics*, vol. 141, Article ID: 104234, 2020.

[14] D. T. Nicholson et al., "Can virtual reality improve anatomy education? A randomized controlled study of a computer-generated three-dimensional anatomical ear model," *Medical Education*, vol. 54, no. 11, pp. 1046–1051, 2020.

[15] S. Al-Ghareeb, H. Cooper, and L. Dyson, "Tracking learning behaviors in VR-based nursing education," *Nurse Education Today*, vol. 109, Article ID: 105215, 2022.

[16] H. S. Maresky et al., "Immersive virtual reality in medical education: A new frontier?" *Medical Teacher*, vol. 41, no. 4, pp. 492–497, 2018.

[17] L. Tudor Car, B. M. Kyaw, A. Teo et al., "Outcomes, Measurement Instruments, and Their Validity Evidence in Randomized Controlled Trials on Virtual, Augmented, and Mixed Reality in Undergraduate Medical Education: Systematic Mapping Review," *JMIR Serious Games*, vol. 10, no. 2, Article ID: e29594, 2022.

[18] K. Liu, W. Zhang, W. Li et al., "Effectiveness of virtual reality in nursing education: a systematic review and meta-analysis," *BMC Medical Education*, vol. 23, Article No. 710, 2023.

[19] N. Birbara et al., "Gamified assessment with conversational AI in immersive anatomy education," *Anatomical Sciences Education*, vol. 16, no. 2, pp. 230–239, 2023.

[20] W. Y. W. Lai and J. S. Lee, "A systematic review of conversational AI tools in ELT: Publication trends, tools, research methods, learning outcomes, and antecedents," *Computers and Education: Artificial Intelligence*, vol. 7, Article ID: 100291, 2024.

[21] D. Adams, “Virtual retail in the metaverse: Customer behavior analytics, extended reality technologies, and immersive visualization systems,” *Linguistic and Philosophical Investigations*, no. 21, pp. 73–88, 2022.

[22] A. Kotranza, D. S. Lind, and B. Lok, “Real-time evaluation and visualization of learner performance in a mixed-reality environment for clinical breast examination,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 7, pp. 1101–1114, 2011.

[23] I. S. Khukalenko et al., “Teachers’ perceptions of using virtual reality technology in classrooms: A large-scale survey,” *Education and Information Technologies*, vol. 27, no. 8, pp. 11591–11613, 2022.

[24] J. Saunier, M. Barange, B. Blandin, R. Querrec, and J. Taoum, “Designing adaptable virtual reality learning environments,” in *Proceedings of the 2016 Virtual Reality International Conference*, 2016, pp. 1–4.

Elena-Ecaterina Opait[1], Adrian Iftene[2]

[1]ORCID: https://orcid.org/0009-0007-9227-1324
Faculty of Computer Science, Alexandru Ioan Cuza University, General Henri Mathias Berthelot Street 16, Iasi, 700483, Romania
E–mail: opaitelena@gmail.com

[2]ORCID: https://orcid.org/0000-0003-3564-8440
Faculty of Computer Science, Alexandru Ioan Cuza University, General Henri Mathias Berthelot Street 16, Iasi, 700483, Romania
E–mail: adiftene@info.uaic.ro

Proceedings of Workshop on Intelligent Information Systems
WIIS2025, October 16-18, 2025, Chisinau, Republic of Moldova

AI Guguță: A Culturally Attuned Voice Tutor Powered by Reasoning LLMs

Alexandr Parahonco, Olesea Caftanatov,
Tudor Bumbu, Marius Cerescu, Nichita Degteariov,
Dan Bumbu, Ion Bostan, Iurie Ciala

Abstract

We present AI Guguță, a voice-first tutoring system that integrates dialect-sensitive speech technology with a reasoning-oriented Large Language Model (LLM). The system combines two core components: (i) a persona-based LLM fine-tuned to guide learners through multi-step mathematics problems, and (ii) a Moldavian speech corpus, comprising both child and adult recordings, used to adapt automatic speech recognition (ASR) and text-to-speech (TTS) models. Early evaluation highlights the LLM's effectiveness in supporting learning, the natural expressiveness of the TTS output, and the solid ASR performance on adult speech, with moderate success on children's speech. We also address the system's pedagogical potential, ethical considerations, and plans for classroom validation.

Keywords: intelligent tutoring, reasoning LLMs, ASR, TTS.

1 Introduction

The landscape of education is being reshaped by artificial intelligence, particularly through intelligent tutoring systems (ITS) that adapt to learners. Large Language Models now enable unprecedented levels of personalized learning support [1]. The **AI Guguță Project** (Code: 25.80012.0807.21TC) aims to create the first AI-driven interactive educational platform in the Republic of Moldova, focused on primary school mathematics [2]. A key novelty of this project is the integration of advanced conversational technologies with local linguistic and cultural heritage [3]. In practice, this means uniting two research efforts: (1) developing a reasoning LLM-based virtual math tutor tailored to Moldavian pedagogy, and (2) constructing Moldavian speech resources and models to give the tutor a voice and ears. By merging these, we seek to modernize learning with an AI tutor that children can converse with naturally in their native dialect, while preserving both teaching methods and folklore.

Voice-based tutoring offers distinct advantages for young learners. Unlike text-only tutors, a voice-enabled tutor can hear students explain their reasoning and respond in real time, capturing rich information often lost in written exchanges [4]. Studies indicate that when students articulate math solutions aloud, it engages metacognitive processes and improves understanding [5]. Therefore, equipping our virtual tutor with speech recognition and synthesis is crucial for authentic dialogue-driven learning. At the same time, the tutor's intelligence comes from LLMs capable of reasoning through multi-step problems and delivering instruction in a conversational manner. Combining these technologies yields a voice-first AI tutor that can listen to a child's question, reason through the answer like a skilled teacher, and speak back a guiding response [6]. To our knowledge, AI Guguță is among the first systems to blend an advanced reasoning LLM with a full dialectal ASR/TTS pipeline for education in a low-resource language.

In this paper, we present a short overview of the AI Guguță system. Section 2 discusses related work in intelligent tutors, speech technology for low-resource languages, and cultural adaptation of AI. Section 3 describes our methods, including the LLM tutor development and the speech corpus design and modeling. In Section 4, we detail experimental results of the LLM's tutoring performance and baseline ASR/TTS evaluations. Section 5 provides a discussion on the integration of the components, the pedagogical and cultural implications, and ethical considerations. Finally, Section 6 concludes with the contributions and outlines future work, such as scaling to multimodal inputs and broader dialect support [7].

2 Related work

Intelligent Tutoring Systems (ITS) have progressed from rule-based designs to LLM-powered platforms [8]. Early systems like AutoTutor relied on scripted dialogue, while recent models such as GPT-3 enable adaptive, open-ended interaction. However, challenges persist around factual reliability and curriculum alignment. Our approach builds on reasoning-specialized LLMs like DeepSeek-R1, which demonstrate emergent capabilities such as chain-of-thought reasoning. Unlike standard LLMs that prioritize fluency, we emphasize logical rigor and pedagogical customization.

Another research thread addresses cultural and linguistic personalization of educational AI. Prior work has highlighted that tutoring systems should respect learners' cultural background to improve engagement and outcomes [9]. Pawar et al. [10] survey cultural awareness in language models, and Li et al. [1] introduce methods to infuse cultural context into LLM responses. Our tutor extends this idea by embodying a local folk character and using regional speech patterns. This aligns with efforts like CultureLLM [2, 11], which incorporate cultural differences into model training. Additionally, code-switching and dialect handling have been studied in NLP for education; for example, Alkhamissi et al. [12] examine how multilingual LLMs align with cultural norms. We contribute a concrete case study of adapting an AI tutor to a specific regional culture (Moldavian in our case), demonstrating culturally responsive AI in practice.

In speech technology, we draw on advances in low-resource ASR and child speech processing. Self-supervised models like *wav2vec 2.0* and *XLS-R* [13, 14] offer robust multilingual foundations. However, child speech remains a challenge due to acoustic variability. We extend prior work by curating a Moldavian child corpus and fine-tuning models to improve recognition.

For TTS, systems like VITS [15] and zero-shot models such as VALL-

E enable expressive, speaker-specific synthesis. These tools support our goal of creating a tutor with a warm and culturally resonant voice. Ethical safeguards, including watermarking, are integrated to mitigate risks. Finally, speech-to-speech translation models like Translatotron 2 [16] inspire future multilingual extensions of AI Guguță.

3 System overview

Our system consists of two main components developed in parallel: (i) the LLM-based tutor brain, encompassing the persona design and model fine-tuning that enable pedagogical reasoning, and (ii) the speech interface, including the data pipeline, ASR engine, and TTS voice of the tutor. These components are integrated into a single conversational agent (Fig. 1).

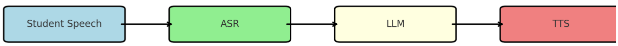


Figure 1. System pipeline (student speech → ASR → LLM → TTS)

The system pipeline illustrated in Figure 1 outlines the overall workflow. When a student poses a spoken query, the Automatic Speech Recognition (ASR) module transcribes the input into text. This textual input is then processed by a Large Language Model (LLM), which generates a pedagogically appropriate response. The response is subsequently synthesized into audio by the Text-to-Speech (TTS) system, producing output in the tutor's designated voice. Figure 1 summarizes this sequence: student → ASR → LLM → TTS → tutor voice output. The following sections describe the design and implementation of each component in detail.

4 Methods

4.1 LLM tutor: persona and adaptation

We designed a persona "Professor Guguță" – a culturally resonant, patient, attentive, and encouraging math tutor – encoded via a persistent system prompt and few-shot examples. This character synthesizes two influences: the teaching style of the late Prof. Boris Ivanovici Cinic, known for rigorous yet engaging math pedagogy, and the charming traits of Guguță, a beloved Moldavian folk hero of children's literature. The tutor is envisioned as a 70-year-old wise educator who carries the soul of Moldavian folklore into every lesson. To establish this persona in the LLM, we wrote a comprehensive system prompt that defines the tutor's identity, voice, and teaching approach.

Key elements of the prompt include: (a) **Identity**: *"You are Professor Guguță, a distinguished elderly mathematician and teacher in Republic of Moldova"* explicitly linking to the folklore figure. (b) **Cultural style**: use of local proverbs and references (e.g., Guguță's cap, traditional foods such as mămăligă – a dense cornmeal porridge, similar to polenta) to explain concepts. (c) **Language usage**: instructing the AI to speak in the Moldavian dialect, reflecting its unique phonetic and lexical features while remaining polite and pedagogical. (d) **Pedagogy**: adopting a Socratic, question-driven tutoring style that encourages critical thinking. (e) **Etiquette**: requiring formal address ("Professor Guguță") and handling multilingual greetings (Romanian, Russian, English) to mirror Moldova's linguistic reality. This carefully engineered prompt ensures that from the first user interaction, the LLM *"stays in character"* as a culturally grounded virtual tutor.

From the technical perspective, we used a hybrid adaptation strategy: prompt engineering (fast iteration) plus LoRA fine-tuning on curriculum-specific dialogues and worked solutions to improve multi-step reasoning and accuracy [6, 7]. Prompt tuning handles tone and persona; LoRA injects domain knowledge efficiently. However, the characteristics of the model also influence the result.

The effectiveness of a virtual tutor heavily depends on the underlying

LLM's ability to reason. Not all LLMs are created equal in this regard. We can distinguish between simple LLMs, which excel at language tasks, and reasoning LLMs, which are designed for complex problem-solving [8].

Simple LLMs, such as earlier versions of Llama, are primarily designed for generating fluent and coherent text. They are excellent for tasks like summarization, translation, and creative writing. However, they may struggle with multi-step reasoning, mathematical problems, and logical puzzles [8].

Reasoning LLMs, on the other hand, are specifically trained to tackle complex cognitive tasks. Models like **DeepSeek-R1, Llama 3.1, Llama 3.2, and Aya** are at the forefront of this new wave of AI [7]. They are trained on vast datasets of code, scientific papers, and mathematical problems, and they often employ techniques like reinforcement learning to enhance their reasoning capabilities.

DeepSeek-R1 demonstrates remarkable reasoning capabilities through large-scale reinforcement learning, naturally emerging with powerful reasoning behaviors including self-verification, reflection, and generating long chains of thought [17]. For our Professor Guguță implementation, we utilize DeepSeek-R1:8b as the base model, as specified in our Modelfile: FROM deepseek-r1:8b.

Llama 3.1 and 3.2 represent significant advances in reasoning capabilities, with Llama 3.1 featuring 128K context length and stronger reasoning abilities, while Llama 3.2 focuses on vision-enabled capabilities for multimodal learning [18].

Aya models, developed by Cohere, focus on multilingual capabilities and cultural awareness, making them particularly relevant for culturally-sensitive educational applications [19].

4.2 Speech dataset: Moldavian Voice

We collected and annotated the Moldavian Voice dataset (approx. 500 hours), containing studio recordings and curated public audio. Key annotation decisions:

- preserve Moldavian diacritics and mark regional lexical variants;
- annotate code-switching occurrences (Russian insertions);
- generate phoneme alignments (forced alignment) and a pronunciation lexicon (∼15k entries).

All minor participants were recorded with parental consent and governed by a data-use policy (restricted raw audio for minors).

4.3 ASR and TTS training

ASR: fine-tuned Whisper and XLS-R backbones with speaker-wise splits, shallow fusion with a domain LM, and augmentations (SpecAugment, speed perturbation). TTS: trained multi-speaker VITS and FastPitch+HiFi-GAN, then adapted an elder-male voice for the tutor. We evaluated both MOS and ASR intelligibility on synthesized audio.

4.4 Evaluation protocol

LLM tutoring: correctness (final answer), scaffold quality (teacher rubric), and persona adherence. ASR/TTS: WER/CER on held-out adult and child test sets; MOS for TTS naturalness; end-to-end pilot with students for usability.

5 Results

5.1 LLM tutoring performance

The AI Guguță Project represents a transformative step in educational technology, advancing the design of culturally responsive virtual tutors built on reasoning-optimized language models. Our research reveals that prompt-level fine-tuning is particularly effective in crafting distinct tutor personas and embedding cultural specificity, while transformer-level fine-tuning offers superior performance in linguistic comprehension and mathematical reasoning – essential for delivering precise, pedagogically sound instruction.

Beyond theoretical modeling, we have successfully developed the foundational tools for an agent-like system (Fig. 2). Professor Guguță interacts seamlessly with a range of file formats – including DOCX, PDF, TXT, and XLSX – enabling flexible integration into diverse educational workflows. The system architecture extends the capabilities of the underlying LLM through persistent chat history and Retrieval-Augmented Generation (RAG), allowing it to draw on external knowledge sources in real time.

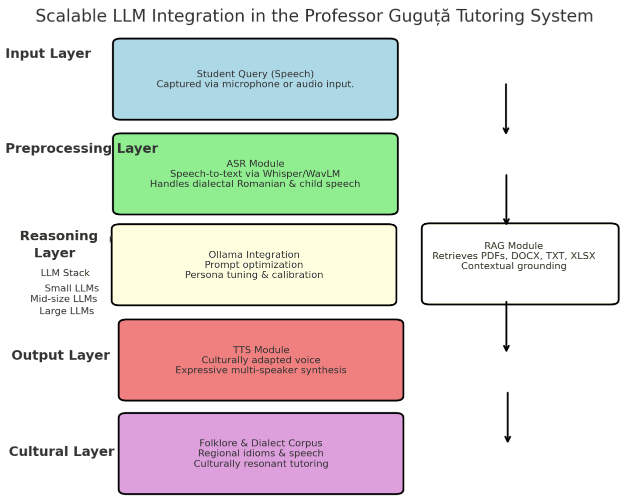


Figure 2. Scalable LLM Integration in the Professor Guguță Tutoring System

This design proves especially powerful for culturally anchored tasks. By invoking regional proverbs, idiomatic expressions, and folkloric references, the tutor can respond not only with intellectual rigor but with cultural resonance. In such moments, the LLM becomes more than a computational engine – it becomes a storyteller, a guide, and a bearer of tradition, bridging modern pedagogy with ancestral wisdom.

A key architectural principle of the Professor Guguță system is its modular support for language models of varying size (Fig. 2), task orientation, and comprehension depth. By designing a flexible interface that accommodates both large-scale reasoning models and compact, task-specific LLMs, our platform lays the groundwork for scalable and adaptive educational agents. Notably, even small and very small models – often overlooked in high-performance settings – can play a vital role in refining user instructions, rewriting prompts, and enhancing interaction fluency.

This layered approach enables dynamic delegation: lightweight models handle prompt restructuring and stylistic adaptation, while more powerful reasoning engines tackle complex pedagogical tasks. The integration of the Ollama system further amplifies this capability, offering a robust framework for prompt optimization across model tiers. Within the Professor Guguță project, Ollama facilitates rapid experimentation with prompt formats, persona tuning, and response calibration – ensuring that each model, regardless of size, contributes meaningfully to the tutor's overall intelligence and cultural sensitivity.

5.2 Performance of the speech recognition system

Fine-tuned Whisper achieved WER $\approx 9.5\%$ (adult) and 14.8% (child); XLS-R baseline obtained $\approx 18.3\%$ WER (child). Code-switching segments remain challenging (WER $\approx$25%). VITS TTS achieved MOS $\approx$4.4/5 (n=15 listeners); intelligibility CER $<2\%$ when re-transcribed by ASR.

Table 1. Selected quantitative results (preliminary)

Metric	Adult speech	Child speech
ASR WER (Whisper finetuned)	9.5%	14.8%
ASR WER (XLS-R finetuned)	11.8%	18.3%
TTS MOS (VITS)	4.4 / 5.0	
LLM tutoring correctness	18 / 20 problems	

6 Discussion

The integration of persona-aware LLMs with dialect-sensitive speech technologies in the Professor Guguță system demonstrates the transformative potential of culturally grounded AI in education. By combining reasoning-specialized models with expressive voice synthesis and regional folklore, the tutor becomes more than a digital assistant – it becomes a pedagogical companion capable of empathy, nuance, and cultural resonance. Prompt engineering played a pivotal role in shaping the tutor's persona, enabling it to respond with warmth, patience, and culturally embedded wisdom. Through iterative refinement and Ollama-based optimization, even small-scale LLMs contributed meaningfully to prompt restructuring and stylistic adaptation. Meanwhile, LoRA fine-tuning at the transformer level enhanced the system's factual precision and procedural reliability, particularly in mathematical reasoning and curriculum-aligned instruction.

Speech adaptation was essential to ensure accessibility across age groups and dialectal variants. The system's ASR pipeline – built on models like Whisper and WavLM – was fine-tuned to handle Moldavian adult and child speech, addressing the acoustic variability and disfluency common in younger learners. Nevertheless, challenges remain: recognition accuracy on highly disfluent child utterances is still limited, occasional code-switching errors persist, and prosodic control in TTS responses requires further refinement to fully capture the tutor's intended tone and emotional cadence.

Ethical safeguards were embedded throughout the system's design. All voice cloning processes were conducted with explicit consent, raw audio access was restricted, and watermarking techniques were applied to prevent misuse of synthetic voices. These measures reflect our commitment to responsible AI development, especially in educational contexts involving minors and culturally sensitive content.

Ultimately, the AI Guguță project affirms that the future of intelligent tutoring lies not only in technical sophistication but in the thoughtful integration of cultural heritage, linguistic diversity, and pedagogical empathy. By building a system that is modular, multilingual, and culturally wise, we lay the foundation for a new generation of educational agents – ones that teach not just with knowledge, but with voice, story, and soul.

7 Conclusion

The AI Guguță project stands as a testament to what is possible when advanced language technologies are guided by cultural insight and educational purpose. By weaving together reasoning-specialized LLMs, dialect-sensitive speech systems, and a richly curated folklore corpus, we have created a tutoring agent that is not only intelligent, but deeply human in its voice, values, and pedagogical style.

Our modular architecture – capable of integrating LLMs of varying size and specialization – lays the foundation for scalable, adaptive learning systems. Whether rewriting prompts with compact models or solving multi-step problems with large reasoning engines, the system demonstrates how layered intelligence can be orchestrated to serve learners with precision and empathy. The Ollama framework further empowers this orchestration, enabling dynamic prompt optimization and persona refinement across model tiers.

Cultural responsiveness is not treated as an aesthetic flourish, but as a core design principle. Through dialectal speech recognition, expressive voice synthesis, and the invocation of regional proverbs and idioms, the tutor becomes a carrier of tradition, as well as a facilitator of learning. This fusion of ancestral wisdom and modern AI opens new pathways for inclusive, meaningful education – especially in underrepresented linguistic and cultural contexts.

As we look ahead, the project invites further exploration: visual mathematics, adaptive learning algorithms, and multilingual extensions are already on the horizon. But the deeper promise lies in the model itself – a tutor who listens, reasons, and responds not just with knowledge, but with cultural grace and pedagogical care.

Acknowledgments. The authors gratefully acknowledge teachers, students, and parents who contributed recordings and feedback, and the funding and administrative support for Project AI Guguță (Code: 25.80012.0807.21TC).

References

[1] C. Li et al., "CultureLLM: Incorporating cultural differences into large language models," in *Advances in Neural Information Processing Systems*, vol. 37, pp. 84799–84838, 2024. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2024/file/9a16935bf54c4af233e25d998b7f4a2c-Paper-Conference.pdf.

[2] O. Ciobanu, "Spiridon Vangheli: paradisul pierdut şi regăsit al copilăriei," *Philologia*, no. 5-6, pp. 59–62, 2013.

[3] J. Shin, C. Tang, T. Mohati, M. Nayebi, S. Wang, and H. Hemmati, "Prompt Engineering or Fine-Tuning: An Empirical Assessment of LLMs for Code," in *Proc. 22nd IEEE/ACM International Conference on Mining Software Repositories (MSR@ICSE)*, (Ottawa, ON, Canada), Apr. 28–29, 2025, pp. 490–502. DOI: 10.1109/MSR66628.2025.00082.

[4] M. A. Ferrag, N. Tihanyi, and M. Debbah, "From LLM Reasoning to Autonomous AI Agents: A Comprehensive Review," arXiv preprint arXiv:2504.19678 [cs.LG, cs.AI], Apr. 28, 2025. DOI: 10.48550/arXiv.2504.19678. [Online]. Available: https://arxiv.org/abs/2504.19678.

[5] P. Nawrot, "nanoT5: A PyTorch Framework for Pre-training and Fine-tuning T5-style Models with Limited Resources," arXiv:2309.02373, v2, Oct. 24, 2023. [Online]. Available: https://arxiv.org/abs/2309.02373.

[6] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, and W. Chen, "LoRA: Low-Rank Adaptation of Large Language Models," in *Proc. International Conference on Learning Representations (ICLR)*, 2022. [Online]. Available: https://arxiv.org/abs/2106.09685.

[7] D. Guo *et al.*, “DeepSeek-R1: Incentivizing reasoning capability in LLMs via reinforcement learning,” arXiv:2501.12948, 2025.

[8] J. Chu-Carroll, A. Beck, G. Burnham, D. O. S. Melville, D. Nachman, A. E. Özcan, and D. Ferrucci, “Beyond LLMs: Advancing the Landscape of Complex Reasoning,” arXiv preprint arXiv:2402.08064, Feb. 12, 2024. DOI: 10.48550/ARXIV.2402.08064. [Online]. Available: https://arxiv.org/abs/2402.08064.

[9] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, “Robust Speech Recognition via Large-Scale Weak Supervision,” arXiv preprint arXiv:2212.04356, Dec. 6, 2022. [Online]. Available: https://arxiv.org/abs/2212.04356 and https://cdn.openai.com/papers/whisper.pdf.

[10] S. Pawar, J. Park, J. Jin, A. Arora, J. Myung, S. Yadav, F. G. Haznitrama, I. Song, A. Oh, and I. Augenstein, “Survey of Cultural Awareness in Language Models: Text and Beyond,” *Computational Linguistics*, vol. 51, no. 3, pp. 907–1004, 2025. DOI: 10.1162/COLI.a.14. (Preprint: arXiv:2411.00860).

[11] A. Baevski, H. Zhou, A. Mohamed, and M. Auli, “wav2vec 2.0: Self-supervised learning of speech representations,” in *NIPS’20: Proceedings of the 34th International Conference on Neural Information Processing Systems*, Article No.: 1044, 2020, pp. 12449–12460.

[12] B. AlKhamissi, M. ElNokrashy, M. AlKhamissi, and M. Diab, “Investigating Cultural Alignment of Large Language Models,” in *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, (Bangkok, Thailand), Aug. 11–16, 2024, pp. 12404–12422. DOI: 10.18653/v1/2024.acl-long.671. [Online]. Available: https://www.aclanthology.org/2024.acl-long.671/.

[13] A. Babu, C. Wang, A. Tjandra, K. Lakhotia, Q. Xu, N. Goyal, K. Singh, P. von Platen, Y. Saraf, J. Pino, A. Baevski, A. Conneau, and M. Auli, “XLS-R: Self-supervised Cross-lingual Speech Representation Learning at Scale,” arXiv preprint arXiv:2111.09296, Nov.

2021. [Online]. Available: https://arxiv.org/abs/2111.09296 (PDF: https://arxiv.org/pdf/2111.09296.pdf).

[14] S. Chen, C. Wang, Z. Chen, Y. Wu, S. Liu, Z. Chen, J. Li, N. Kanda, T. Yoshioka, X. Xiao, J. Wu, L. Zhou, S. Ren, Y. Qian, Y. Qian, J. Wu, M. Zeng, X. Yu, and F. Wei, "WavLM: Large-Scale Self-Supervised Pre-Training for Full-Stack Speech Processing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 16, no. 6, pp. 1505–1518, Oct. 2022. DOI: 10.1109/JSTSP.2022.3188113. (Preprint: arXiv:2110.13900).

[15] Jaehyeon Kim, Jungil Kong, and Juhee Son, "Conditional Variational Autoencoder with Adversarial Learning for End-to-End Text-to-Speech," in *International Conference on Machine Learning (ICML)*, 2021. [Online]. Available: https://scholar.google.com/citations?view_op=view_citation&hl=en&user=-ZJaGikAAAAJ&citation_for_view=-ZJaGikAAAAJ:UeHWp8X0CEIC.

[16] Y. Jia, M. T. Ramanovich, T. Remez, and R. Pomerantz, "Translatotron 2: High-quality direct speech-to-speech translation with voice preservation," in *Proceedings of the 39th International Conference on Machine Learning (PMLR), vol. 162*, (Baltimore, MD, USA), Jul. 17–23, 2022, pp. 10120–10134. [Online]. Available: https://proceedings.mlr.press/v162/jia22b/jia22b.pdf.

[17] S. V. Marjanović, A. Patel, V. Adlakha, M. Aghajohari, P. BehnamGhader, M. Bhatia, A. Khandelwal, A. Kraft, B. Krojer, X. H. Lù, N. Meade, D. Shin, A. Kazemnejad, G. Kamath, M. Mosbach, K. Stańczak, and S. Reddy, "DeepSeek-R1 Thoughtology: Let's think about LLM Reasoning," arXiv preprint arXiv:2504.07128, Apr. 2025. [Online]. Available: https://arxiv.org/abs/2504.07128.

[18] Meta AI, "Introducing Llama 3.1: Our most capable models to date," Meta AI Blog, Jul. 23, 2024. [Online]. Available: https://ai.meta.com/blog/meta-llama-3-1/. (Accessed: Sep. 24, 2025).

[19] Cohere for AI (Cohere Labs), "Aya: Open-access multilingual research LLM family (Aya technical report / Aya 23 releases)," Tech-

nical Report, 2024. Available: https://cohere.com/research/aya (Aya model paper: Feb 13, 2024), and Aya 23 report: https://arxiv.org/abs/2405.15032. [Accessed: Sep. 24, 2025].

Alexandr Parahonco[1], Olesea Caftanatov[2], Tudor Bumbu[3], Marius Cerescu[4], Nichita Degteariov[5], Dan Bumbu[6], Ion Bostan[7], Iurie Ciala[8]

[1,2,3,4,5,6,7,8]Vladimir Andrunachievici Institute of Mathematics and Computer Science, Moldova State University

[1]ORCID: https://orcid.org/0009-0007-3486-5597
Alecu Russo State University from Balti
E–mail: alexandr.parahonco@math.md

[2]ORCID: https://orcid.org/0000-0003-1482-9701
E–mail: olesea.caftanatov@math.md

[3]ORCID: https://orcid.org/0000-0001-5311-4464
E–mail: tudor.bumbu@math.md

[4]ORCID: https://orcid.org/0009-0000-5704-6577
E–mail: marius.cerescu@gmail.com

[5]ORCID: https://orcid.org/0009-0008-2352-8971
E–mail: nichita.degteariov@math.md

[6]ORCID: https://orcid.org/0009-0001-2730-3744
E–mail: dan.bumbu@math.md

[7]ORCID: https://orcid.org/0009-0006-8690-5135
E–mail: ion.bostan@math.md

[8]ORCID: https://orcid.org/0009-0002-4048-8138
E–mail: iurie.ciala@math.md

Digitizing of Bessarabian Folklore: From Textual Heritage to Digital Treebank

Mircea Petic, Olesea Caftanatov, Svetlana Cojocaru, Ludmila Malahov, Atul Kr. Ojha

Abstract

The paper investigates aspects of digitization and transliteration of Bessarabian folklore from the 19th to the 20th century using the HeDy platform. In our research, we present the FolkAI Treebank, part of diachronic parallel corpus, aligned with Universal Dependencies, and adapted to the Bessarabian dialect. The Treebank preserves the linguistic and cultural specificity of folklore through digitization, normalization, tokenization, and annotation in CoNLL-U format. The initial results include two digitized works with progress in transliteration, tokenization, and annotation. Beyond academic value, FolkAI supports the learning of the Romanian language, digital education, and cultural tourism by transforming folklore into an accessible and interactive resource.

Keywords: Folklore, Treebank, parallel diachronic corpus, digitization.

1 Introduction

Folklore is an essential part of a people's cultural identity and collective memory, serving as a way to reflect on history and the social and economic traditions of the past. Encyclopedia Britannica [1] defines folklore as follows: "*In modern usage, an academic discipline the subject matter of which (also called folklore) comprises the sum total of* *traditionally derived and orally or imitatively transmitted literature, material culture, and custom of subcultures within predominantly literate and technologically advanced societies; comparable study among wholly or mainly non-literate societies belongs to the disciplines of ethnology and anthropology*".

In other words, folklore is a way of transmitting knowledge. We are more accustomed to the fact that through oral or written creation, knowledge about the history and customs of the region, as well as ethical and behavioral norms and ritual procedures, is propagated. Perhaps this mode of communication is not so visible for reasons related to the exact sciences. This aspect is humorously described in Nikita Razgovorov's science-fiction novel "Four Furry Feet" (1963).

The hero of this story, the physicist Prof. Yr, finds that certain idiomatic expressions, proverbs, and sayings are, in fact, less common formulations of physical laws, for example:

– *Strike iron while it is hot* – express the results of observations of changes in the aggregate state of iron with increasing temperature;

– *All that glitters is not gold* – this is an absolute truth and the beginning of spectral analysis of metals;

– *No one can embrace the infinite* – a concise formulation of the theory of relativity.

Of course, these examples are humorous, but they deserve the consideration that folklore is a tool for promoting knowledge in various fields. This tool can also spark interest in certain phenomena, laws of nature, etc., in an accessible, attractive, and memorable way. Therefore, folklore, in its apparent simplicity, contains a veritable treasure trove of knowledge, which we aim to accumulate, classify, and contribute to its preservation and propagation.

The development and application of IT tools could facilitate both academic studies aimed at restoring the authenticity of folklore and make a significant contribution to the promotion of our cultural heritage. Although folklore is understood to encompass all literary, musical, artistic, and intangible creations, in our study, we will limit ourselves to folkloric works expressed through language, thereby placing our research within the area of natural language processing.

Any natural language application typically relies on two basic components: (i) a collection of linguistic resources and (ii) the operating tools (processing/interaction with them) for achieving certain goals. In our case, the object of study is the digitization of folklore works, the first component being a parallel diachronic corpus (with Cyrillic and Latin script), which will include texts from different historical periods (19th-20th centuries, with possible incursions into earlier periods), allowing for the analysis of the evolution of the vocabulary characteristic of Bessarabian folklore. The lower limit of the corpus volume is estimated to be a minimum of 100,000 tokens.

Unlike traditional corpora, some of the entries in this corpus will be accompanied by images created using innovative AI technologies that will illustrate the meanings of various folkloric elements. At the same time, based on the collected materials (including illustrative ones), a folklore book (in traditional and interactive electronic versions) could be published, which could be used for educational purposes. These materials, digitized, transliterated, and improved with AI, can support the learning of the Romanian language, including for those who learn it as a foreign language. They can also fit into digital education tools. Visually rich folktales and proverbs can be effectively incorporated into interactive lessons, language apps, and cultural programs. Economically, the platform can boost local and international cultural tourism by providing digital museum guides, heritage trails, and AR storytelling experiences. This makes folklore accessible and engaging to tourists, educators, and learners around the world.

In this article, our aim is to describe the process of developing the diachronic parallel FolkAI Treebank. Since the very notion of a diachronic parallel corpus in general and an FolkAI Treebank in particular presupposes a technical instrument, we introduce HeDy, the platform used for digitization and transliteration. A short review of related works is presented in the next *Section 2.* The functionalities of the HeDy platform are briefly outlined in *Section 3.* The specific features of the FolkAI Treebank, aligned with *Universal Dependencies* (UD) standards but also adapted to the Bessarabian dialect, are discussed in *Section 4.*

2 Related works

We should mention that the digitization of folklore texts has its specific characteristics, which have been highlighted in various publications. In many cases, related works emphasize that the description of folklore subjects was carried out by different collectors, who left their personal mark on everything from writing (in the case of handwriting) to different styles of presentation and even spelling. In such cases, it is necessary to create a separate digitization model for each collector [2]. On the other hand, the material collected by a single collector may be insufficient to train the model, so it will be necessary to supplement it with material collected by other people or to create a model for a group of collectors.

The authors of the works in this field consider broadening the notion of digitizing folklore by preserving elements of the speaker's voice tone and using various multimedia means (including virtual or augmented reality) [3]. According to [4], folkloric texts differ fundamentally from philological ones. A key characteristic of folklore is its tendency to generate multiple versions. The authors emphasize that "*the peculiarity of folkloric texts is that versions are created not only during collection and documentation (through transcription, editing, and publication), where a change of environment takes place, but even the primary text itself can only be transcribed in variants*". Therefore, the aspect of variability must be taken into account when digitizing folkloric works. A decade earlier [5], it was observed that traditional corpora lack emotional tone or narrative trajectories, and digitizing these aspects would require advanced annotation elements that distinguish them from general-purpose corpora.

3 Workflow of the HeDy platform illustrated by processing a fragment of folkloric text

The operating tool will be based on the HeDy platform [6]. The HeDy platform is an open software tool designed for digitizing and transliterating printed documents or historical Romanian manuscripts written in Cyrillic and transitional scripts, covering the 16th to 20th centuries. HeDy covers the entire workflow, from image preprocessing and text recognition (OCR) to proofreading and exporting to standardized formats. Among the many applications of this platform, we would like to mention the production of resources based on the processing of folkloric texts.

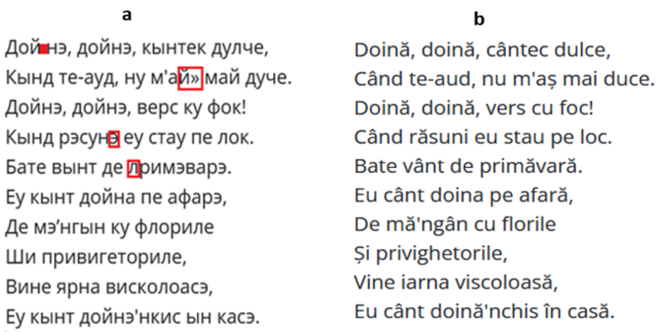
To illustrate the capabilities of the HeDy platform, we will select a small fragment from the well-known "Doina" (Figure 1), a representative of the folk lyric literature genre [7].

Дойнэ, дойнэ, кынтек дулче,
Кынд те-ауд, ну м-аш май дуче.
Дойнэ, дойнэ, верс ку фок!
Кынд рэсунь, еу стау пе лок.
Бате вынт де примэварэ.
Еу кынт дойна пе афарэ,
Де мэ-нгын ку флориле
Ши привигеториле,
Вине ярна висколоасэ,
Еу кынт дойнэ'нкис ын касэ.

Figure 1. Fragment of the folk lyric “Doina”

The scanned image file is taken over by HeDy and preprocessed to improve image quality. Typical operations include deskew, contrast/brightness adjustment, background noise removal, etc. We note that the quality of this source image is not excellent, but there is hope that the preprocessing stage will improve the situation. The next step is optical character recognition.

HeDy offers several recognition models, the number of which is constantly growing. In our case, the model for the 20th century, based on the Soviet Cyrillic alphabet, is suitable. The raw OCR output is presented in Fig. 2a. HeDy allows for error correction, providing virtual keyboards tailored to the relevant historical periods. Once the errors are corrected, the text is transliterated, producing the final version shown in Figure 2b.



a	b
Дойнэ, дойнэ, кынтек дулче,	Doină, doină, cântec dulce,
Кынд те-ауд, ну м'ай» май дуче.	Când te-aud, nu m'aș mai duce.
Дойнэ, дойнэ, верс ку фок!	Doină, doină, vers cu foc!
Кынд рэсунэ еу стау пе лок.	Când răsuni eu stau pe loc.
Бате вынт де примэварэ.	Bate vânt de primăvară.
Еу кынт дойна пе афарэ,	Eu cânt doina pe afară,
Де мэ'нгын ку флориле	De mă'ngân cu florile
Ши привигеториле,	Și privighetorile,
Вине ярна висколоасэ,	Vine iarna viscoloasă,
Еу кынт дойнэ'нкис ын касэ.	Eu cânt doină'nchis în casă.

Figure 2. Figure 2. a) Recognized text, with recognition errors marked b) Final result – recognized and transliterated text

Thus, we can conclude that the HeDy platform provides us with an excellent component for creating the FolkAI corpus, taking on the functions of preprocessing, digitization, and transliteration, as well as offering possibilities for postprocessing the document.

4 FolkAI Treebank

In this section, we describe the process of constructing the FolkAI Treebank, a diachronic resource aligned with the principles of UD [8], while also capturing the specific features of the Bessarabian dialect. The creation of such a Treebank involves several steps, beginning with the collection and digitization of primary texts, followed by data cleaning, normalization, and metadata annotation. One important motivation for building this Treebank is the absence of any comparable resource in the Bessarabian linguistic space, which makes our work pioneering in this direction. Inspired by this gap, we began by digitizing two key works of Moldovan folklore: "O creație populară: proverbe și zicători (A folk creation: proverbs and sayings)" by G. Botezatu et al. (1981, 309 pages) [9] and "Folclor din părțile Codrului (Folklore from the Codru region)" by G. Botezatu et al. (1973, 362 pages) [10]. Both books were carefully digitized, validated through recognition processes, and subsequently transliterated from Cyrillic to Latin script. More details on the digitization process are described in Section 3. Once the textual foundation was established, the next crucial stage was tokenization and word segmentation, which determine the basic linguistic units on which morphological and syntactic annotation will be built.

4.1 Tokenization and Word Segmentation in the FolkAI Treebank

Tokenization in the FolkAI Treebank follows the general principles of UD while also addressing the specific orthographic and morphological characteristics of the Bessarabian (Moldovan) dialect. In general, words are delimited by whitespace or punctuation, but typographical conventions, dialectal usage, and the presence of hyphens or apostrophes introduce additional complexity that requires careful documentation.

In line with UD guidelines, multi-word tokens are generally not required, but specific exceptions are systematically enumerated. These include:

– Intervals: numeric ranges are split into separate tokens.
Example: 1721–1726 → 1721, –, 1726
(20/1X-1978) → (, 20, /,1X, -, 1978,)

– Abbreviations: treated as a single token.
Example: **СССР → СССР, АН → АН, ДФМР →ДФМР**

– Inflection markers: split from the root.
Example: **мош-улуй → мош, -улуй ("the old man's")**
урс-улуй → урс, -улуй (the bear's)

– Hyphenated words: generally split into components unless otherwise specified.
Examples: **Хан-тэтар → Хан, -, тэтар**
ложико-семиотик → ложико, -, семиотик
Фрате-меу→ Фрате, -,меу

– Lexicalized compounds: words recorded in the DOOM dictionary [11] are treated as single tokens:
Examples: **прим-министру,**
and some folklore mwe-units *Example*: **Пенеш-ымпэрат**.

– Names of places: treated as single tokens.
Example: **Болбочий-Бекь (Bolbocii-Vechi).**

– Ad-hoc compounds: treated as multi-word tokens.
Examples: **еуро-асиатикэ (Euro-Asiatic), молдо-русо-украинянэ (Moldavian Russian Ukrainian).**

– Contractions and clitics: split into separate tokens, generally keeping the hyphen with one of the components.
Short pronominal forms: **Вре-о → вре, -о, Че-ай → Че-, ай**

– Hyphens usually indicate a missing sound, which stays with the root word:
гура-ць → гура, -ць ("your mouth")
а-шь → а, -шь ("I would")

– Apostrophes: treated similarly to hyphens in contraction contexts.
Examples:
н'аш → н', аш ("I would not")
н'аре → н', аре ("Don't have")

– Special prepositional cases: some prepositions carry a hyphen after them when combined with the following words.
Example: **де-а луй → де-, а, луй ("of the" or "of a(n)")**

Proverbs and idiomatic expressions are tokenized as separate words even when they contain hyphenated forms or apostrophes, such as ну-й ("isn't") or де-ар ("if... would"). By documenting these rules, the FolkAI Treebank provides a reproducible tokenization framework suitable for morphological and syntactic annotation, ensuring that both standard Romanian forms and dialectal variations are correctly represented for computational processing. For the automation of the tokenization process, we employed the UDPipe tool (version 1.2.0) together with the Romanian model romanian-rrt-ud-2.5-191206. For the Romanian model to work, we transliterated our dataset from Cyrillic to Latin Script. The model was tested on a sample page from the Proverbs collection, as illustrated in Fig. 3. Afterwords, we transliterated it back from Latin to Cyrillic Script.

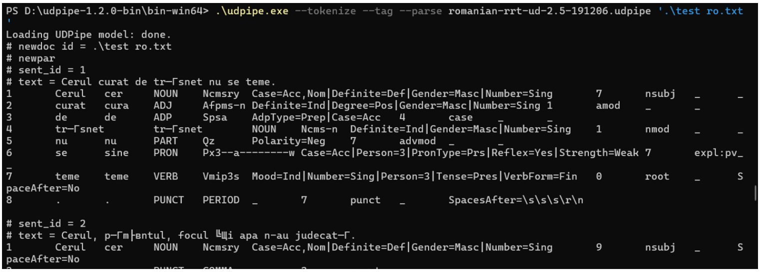


Figure 3. Sample tokenization output on a page from proverbs collection using UDPipe and Romanian UD model

4.2 Annotation Process

The second stage in the construction of the FolkAI Treebank is the **annotation process**, which follows the **CoNLL-U format**. The CoNLL-U format used in the annotation process is structured into ten tab-separated columns, with each column capturing a distinct layer of linguistic information. The **ID** column represents the word index, starting from 1, while the **FORM** column contains the surface form of the word or punctuation symbol. The **LEMMA** column provides the canonical form or stem of the word, and **UPOS** specifies its universal part-of-speech tag. Language-specific information is recorded in the **XPOS** column, or replaced with an underscore if unavailable. Morphological details are included in the **FEATS** column, again replaced with an underscore when no features are present. Syntactic structure is encoded through the **HEAD** column, which indicates the head of the current word (either another word's ID or zero if the word is the root), and the **DEPREL** column, which defines the dependency relation to the head. Additional structural information may be provided in the **DEPS** column in the form of enhanced dependency graphs, expressed as head–relation pairs. Finally, the **MISC** column stores any supplementary annotation not covered in the previous categories. An illustration of this annotation scheme is given in Fig. 4, which shows the encoding for the sentence id_1:

"Черул курат де трэснет ну се теме", translated as *"The clear sky doesn't fear the thunderbolt"*.

# newdoc id = .\test ro.txt									
# newpar									
# sent_id = 1									
# text = Черул курат де трэснет ну се теме.									
1	Черул	Чер	NOUN	Ncmsry	Case=Acc,Nom|	7	nsubj	_	_
2	курат	курат	ADJ	Afpms-n	Definite=Ind|Deg	1	amod	_	_
3	де	де	ADP	Spsa	AdpType=Prep|C	4	case	_	_
4	трэснет	трэснет	NOUN	Ncms-n	Definite=Ind|Ger	1	nmod	_	_
5	ну	ну	PART	Qz	Polarity=Neg	7	advmod	_	_
6	се	сине	PRON	Px3--a--------w	Case=Acc|Persc	7	expl:pv	_	_
7	теме	теме	VERB	Vmip3s	Mood=Ind|Numb	0	root	_	SpaceAfter=No
8	.	.	PUNCT	PERIOD	_	7	punct	_	SpaceAfter=No

Figure 4. Sample of annotation scheme based on CoNLL-U format

For morphological annotation, the FolkAI Treebank employs all 17 Universal POS categories[12], ensuring compatibility with the UD framework. In addition, for syntactic relations, we adopt the full set of 37 Universal Dependencies relation types [13], as illustrated in Figure 3 on sentence id_4.:

"Ши тина лучеште, кынд соареле о избеште", which means *"Even the mud shines when the sun strikes it"*.

The annotation process is carried out using INCEpTION [14], a text-annotation environment designed to support a wide range of linguistic and machine learning tasks on written text. INCEpTION is implemented as a web application that allows multiple users to collaborate on the same annotation project while also supporting the management of multiple projects simultaneously. One of its key features is an integrated recommender system, which assists annotators in producing

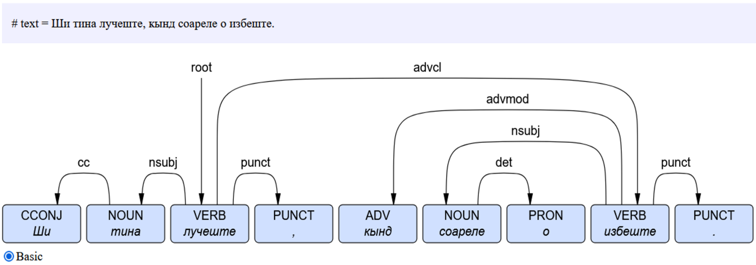


Figure 5. Example of syntactic relation visualisation

annotations more efficiently and consistently. In addition to annotation, INCEpTION enables the creation of corpora by searching external document repositories and importing relevant texts. It also provides access to knowledge bases, which can be leveraged for advanced tasks such as entity linking. An example of the lemmatization process within INCEpTION is illustrated in Fig. 5.

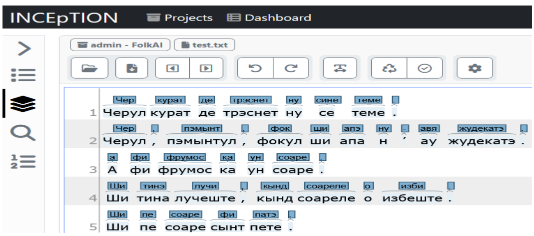


Figure 6. Lemmatization of the first five sentences in the INCEpTION annotation environment

In sum, the FolkAI Treebank integrates standardized UD principles with dialect-specific rules of tokenization and annotation, while employing tools such as UDPipe and INCEpTION to ensure both automation and manual refinement.

5 Conclusion

The creation of the FolkAI Treebank is an important step in preserving and valuing the cultural heritage of Bessarabian folklore. By combining digitization, transliteration, and detailed annotation that follows Universal Dependencies standards, the project provides a clear method and an innovative linguistic resource. The HeDy platform has been crucial in dealing with the challenges of historical scripts and dialectal variation, while also ensuring precision and reproducibility.

In terms of concrete progress, two books have been digitized, one and a half books have been transliterated, and the book of proverbs has been fully tokenized automatically and is currently under validation. At the annotation level, 30 sentences have been marked. The ultimate goal is to expand the corpus to a minimum of 100,000 tokens.

Beyond its academic value, the corpus offers new opportunities for education, language learning, and cultural tourism, making folklore an accessible and engaging digital resource. This approach helps to protect intangible heritage and brings it into today's digital culture.

Acknowledgments. Part of the research for this paper was supported by the project 25.80012.0807.50SE, "Digitization and creation of the FolkAI corpus with texts of Bessarabian folklore from the 19th–20th centuries", as well as by the COST Action CA21167, "Universality, Diversity and Idiosyncrasy in Language Technology (UniDive)".

References

[1] "Folklore. Academic discipline," Britannica. [Online]. Available: `https://www.britannica.com/topic/folklore-academic-discipline`. Accessed August 24, 2025.

[2] B. Ó. Raghallaigh, A. Palandri, and C. M. Cárthaigh, "Handwritten Text Recognition (HTR) for Irish-Language Folklore," in *Proceedings of the 4th Celtic Language Technology Workshop within*

LREC2022, (Marseille, France), European Language Resources Association, 2022, pp. 121–126.

[3] A. Koichumanova, “Folk Tales in the Digital Age: New Perspectives on Storytelling and the Oral Tradition,” Master’s Degree in Communication Strategies, Universita degli studi di Padova. School of Human and Social Sciences and Cultural Heritage, 2024, 104 p.

[4] M. Domokos, “Textual Concepts and Textological Practices in Hungarian Folkloristics,” *Acta Ethnographica Hungarica*, vol. 67, no. 2, pp. 437–465, 2023. DOI: 10.1556/022.2023.00010.

[5] J. Abello, P. Broadwell, and T. R. Tangherlini, “Computational Folkloristics,” *Communications of the ACM*, vol. 55, no. 7, pp. 60-70, 2012. DOI: `https://doi.org/10.1145/2209249.2209267`.

[6] T. Bumbu, L. Burtseva, S. Cojocaru, A. Colesnicov, and L. Malahov, *Digitization of Romanian Historical Printings*, Chișinău: Vladimir Andrunachievici Institute of Mathematics and Computer Science, 2023, 176 p. ISBN 978-9975-68-497-2.

[7] E. Junghietu, *Who said doinu-doina*, (In Romanian: Cine-a zis doinu-doina, in Cyrillic alphabet). Chişinău: Editura Literatura Artistică, 1981, p. 27.

[8] Universal Dependencies. [Online]. Available: `https://universaldependencies.org/`. Accessed August 24, 2025.

[9] G. Botezatu et al., *Folk Creation. Proverbs and Sayings.* (In Romanian: Creație populară. Proverbe și zicători.) 1981, p. 309.

[10] G. Botezatu et al., *Folklore from the Codru Region.* (In Romanian: Folclor din părțile Codrului.) 1973, p. 362.

[11] Dicționarul ortografic, ortoepic și morfologic al limbii române. [Online]. Available: `https://doom.lingv.ro/`. Accessed August 24, 2025.

[12] Universal POS tags. [Online]. Available: https://universaldependencies.org/u/pos/index.html. Accessed August 24, 2025.

[13] Universal Dependency Relations. [Online]. Available: https://universaldependencies.org/u/dep/index.html. Accessed August 24, 2025.

[14] A semantic annotation platform offering intelligent assistance and knowledge management. [Online]. Available: https://inception-project.github.io/. Accessed August 24, 2025.

Mircea Petic[1], Olesea Caftanatov[2], Svetlana Cojocaru[3], Ludmila Malahov[4], Atul Kr. Ojha[5]

[1,2,3,4] "V. Andrunachievici" Institute of Mathematics and Computer Science, Chisinau, Republic of Moldova

[1]ORCID: https://orcid.org/0000-0001-6044-7646
E–mail: mircea.petic@math.md

[2]ORCID: https://orcid.org/0000-0003-1482-9701
E–mail: olesea.caftanatov@math.md

[3]ORCID: https://orcid.org/0009-0003-1025-5306
E–mail: svetlana.cojocaru@math.md

[4]ORCID: https://orcid.org/0000-0001-9846-0299
E–mail: ludmila.malahov@math.md

[5]ORCID: https://orcid.org/0000-0002-9800-9833
Insight SFI Research Centre for Data Analytics, Data Science Institute, Galway University, Galway, Ireland.
E–mail: shashwatup9k@gmail.com

Maintaining the value of data/knowledge assets in medical diagnostics: an exploratory study

Olga Popcova, Elena Guţuleac, Iulian Secrieu, Carolina Ţâmbală

Abstract

We conducted a survey to explore the experiences of medical professionals regarding maintaining the relevance over time of data/knowledge assets – such as scoring systems, data and knowledge warehouses, and AI models. It aimed to obtain insights from medical practitioners in order to concretize the subject of the longevity of existing information tools in the domain of medical diagnostics. This exploratory study provided us with a clearer understanding of the use of information tools, including the importance of a second opinion, trust in historical data, regular updates of the data/knowledge assets, and the need for additional explanations of the generated conclusions.

Keywords: information tools, scoring systems, life cycle, longevity, decision making, medical diagnostics.

1 Introduction

Every information tool (knowledge-based system, decision support system, scoring system, etc.) typically goes through several life cycle stages [1] as follows: emergence, adoption, growth, maturity, decline, replacement/renewal.

We try to track the evolving use of information tools over time, identifying when older methods become less effective, when new technologies are applied, and how these changes impact efficiency and decision quality. The domain of medical diagnostics was selected as the primary area of study for the problem.

The use of information tools to support decision-making in medical diagnostics has evolved from manually calculated scores to rule-based expert systems, knowledge-based systems, integrated electronic health records support, decision support systems, and predictive tools based on AI; each phase being driven by technological progress, clinical needs, and regulatory acceptance.

We seek to gain, through a survey, a better understanding of the life cycle of information tools, specifically how their use, value, and relevance evolve over time and in different contexts within the field of medical diagnostics. It will allow us, in the first stage, to identify the constraints and limitations of the information tools used in medical diagnostics. In the second stage, we will be able to propose solutions to minimize the negative effects of the most important ones, identified in the first stage.

In this way, enhancing the longevity of existing information tools used in medical diagnostics implies enhancing the longevity of the utilised data/knowledge assets.

There exist different understandings [2] of "increasing the longevity of a knowledge base" as follows:

- Prolonging the life cycle of the knowledge base by maintaining it for a longer period without necessarily implying significant changes or enhancements. It suggests extending its longevity by keeping it relevant over time.

- Extending the life cycle of the knowledge base by increasing the period during which it remains useful and relevant. It entails a more proactive approach, potentially incorporating updates and expansions to keep its ongoing value.

- Sustaining the knowledge base. It is a broader approach aimed at maintaining the relevance and usefulness of the knowledge base without necessarily focusing on time or extending its life cycle.

In our understanding, enhancing the longevity of the knowledge base means focusing on improving its long-term usefulness, potentially through updates, optimizations, or reorganization.

This understanding is based on the authors' long-term experiences in developing medical decision support tools, studying the current state of the art of medical diagnostics as the selected domain, and reviewing our existing data/knowledge assets with an emphasis on longevity perspectives [3].

2 The evolution of formalization of decision-making in medical diagnostics

In the medical diagnostics domain, there is no single approach sufficient to structure and formalize the decision-making process. All approaches aim to reduce subjectivity, ensure consistency, and improve patient treatment procedures; however, each has its own specific advantages and limitations. The widespread use of knowledge-based, data-driven, and mixed approaches is determined by how they support decisions under uncertainty and complexity.

The evolution of formalizing decision-making in medical diagnostics progressed from scoring tools to rule-based approaches, statistical methods and probabilistic methods, knowledge-based approaches, data-driven and machine learning approaches, and presently – hybrid and patient-centered ones. Each of these approaches balances the integration of evidence, interpretability, and adaptability in different ways.

The longevity and life cycle of information tools depend significantly on the approach used in their development. They are particularly dependent on the limitations inherent in the chosen approach, since the negative impacts of any limitations usually increase over time.

Therefore, it is crucial to consider these factors from the very beginning of the design of information systems, namely, to recognize all (or at least the main) limitations.

Let's outline the main advantages and limitations of the basic approaches [4]-[7] used to formalize decision-making (see Table 1).

Table 1. Basic approaches used to formalize decision-making

Approaches	General advantages and limitations
Rule- and knowledge-based approaches	+ Transparent and interpretable, widely adopted. – Knowledge bases are hard to maintain and may lag behind new evidence.
Scoring systems and threshold models	+ Easy to apply, evidence-based, widely adopted. – Oversimplifies complex, multifactorial cases.
Data-driven and machine learning approaches	+ High accuracy with large datasets, can detect complex patterns, widely adopted. – Require interpretability methods.
Knowledge and data fusion	+ Balances interpretability and adaptability, widely adopted. Combine guidelines (knowledge-driven) with machine learning (data-driven). – Can create reduced sensitivity to alerts because of overexposure and over-trusting the tool at the expense of human expertise.
Statistical and probabilistic models	+ Quantitative and adaptable to uncertainty. – Requires quality data and statistical literacy.
Intuitive/heuristic decision-making	+ Very fast, efficient in familiar scenarios, leverages tacit knowledge from experience. – Prone to cognitive biases, can miss atypical cases.

The knowledge-based approach relies on expertise, rules, heuristics, and past experiences stored in knowledge bases. It provides explainable decisions, which helps ensure consistency and standardization in repetitive decisions. This approach is crucial where expert reasoning and explainability are essential.

The data-driven approach focuses on identifying patterns, correlations, and trends. It enables real-time, adaptive decision-making using big data (often through statistical analysis or machine learning techniques), and can uncover hidden patterns that may not be apparent to human intuition. The data-driven approach is effective when real-time data or extensive datasets are available.

A mixed (knowledge-based and data-driven) approach is emerging as the gold standard for high-stakes decisions requiring both adaptability and transparency. Rapidly growing, it addresses the need for both performance and explainability.

Scoring systems provide a structured method for assigning numerical values to various criteria/parameters, allowing for an objective evaluation and comparison of different options. Traditional scoring systems play a significant role as decision-making tools, facilitating more data- and evidence-based decisions. Scoring systems play an important role, particularly in diagnostics, prognosis, and triage.

Scoring systems are essential in each of the highlighted approaches. However, there are some peculiarities regarding the explicit and implicit use of scores:

- In rule- and knowledge-based approaches, scores serve as a bridge between heuristics and formalization;

- In scoring systems and threshold models, scores help define cut-off thresholds;

- In data-driven and machine learning approaches, scores generated by machine learning may lack interpretability, unlike classic (traditional) scores;

- In knowledge and data fusion, implicit scores become automated and sometimes invisible to the clinician;
- In statistical and probabilistic models, scoring provides a user-friendly surrogate for complex regression equations;
- In intuitive/heuristic decision-making, scores act as a bias-corrector without replacing intuition.

We can conclude that scoring systems are deeply embedded in almost every formal approach to medical decision-making. They may be used explicitly, such as in risk calculators and cut-off values, or implicitly, as seen in the embedded weights within models or AI.

3 The revision of our existing data/knowledge assets

Subsequently, we decided to analyze how the use of our existing medical diagnostics information tools, as well as their value and relevance, has evolved over time and across different contexts. In our long-term experience of developing systems based on professional knowledge in the domain of medical diagnostics, we utilized scoring systems in explicit and implicit modes, embedded in knowledge-based, data-driven, and mixed approaches.

Four assets were analyzed:

1. Expert knowledge bases in the field of ultrasound diagnostics of the hepato-pancreato-biliary area (129 nodes for liver, 149 – for pancreas, 115 – for kidney, 114 – for spleen) [9];
2. The core of knowledge bases restricted by some principles: limited by the examination time, examined area, purpose and conditions of investigation, etc. (30 nodes for hemorrhagic shock, 63 – for the Extended Focused Assessment with Sonography for Trauma) [9];

3. Scoring systems developed for the differentiated diagnosis of liver cirrhosis [10];

4. KB for supporting triage based on vital signs for patients in multiple casualty disasters (7 vital signs as basic characteristics, 4 triage categories) [11].

The revision of our data/knowledge assets confirms: i) knowledge bases are hard to maintain and may lag behind new evidence; ii) scoring systems can oversimplify complex, multifactorial cases, particularly when used implicitly; however, when used explicitly, they necessitate interpretability methods to ensure understanding.

Besides, two major problems were highlighted [3]:

1. The trust in historical data obviously decreases over time and the lack of a clear, transparent procedure for its regular validation.

2. Sometimes thresholds (cut-off points) can be perceived by a practitioner as not grounded enough, especially when small differences in scores can lead to very different conclusions.

It is essential to validate our assumptions with the end-users of the data/knowledge assets in medical diagnostics. To achieve this, we conducted a survey to investigate the experiences of medical professionals regarding maintaining the relevance of data/knowledge assets over time.

4 Survey results

The 12-question survey covered various aspects, such as: use of information tools in medical practice (3 questions), their relevance over time (2 questions), validation and update procedures (2 questions), interpretation and explanation issues (3 questions), and cut-off points in scoring systems (2 questions).

The survey was distributed as a hard-printed copy to experienced medical practitioners and resident doctors from public and private medical centers. We received completed responses from 8 respondents specialised in medical imagistics.

Based on the provided answers, we can identify the following main findings:

- All respondents agreed on the importance of having a second opinion in their medical decision-making practice, 87.5% indicating it as "important" and 12.5% as "very important".
- All respondents reported using knowledge bases in their medical practice. Additionally, 62.5% also use scoring systems, and 50% utilize decision support systems.
- The vast majority (87.5%) noted more frequent use of the mentioned tools in the last 3-5 years, answering "yes" to the corresponding question.
- 37.5% of respondents consider that trust in historical data (used in the data/knowledge assets such as decision support systems, knowledge bases, and scoring systems) decreases over time, while 62.5% answered "no" on the corresponding question.
- All respondents agreed on the importance of keeping these data/knowledge assets relevant over time, with 75% indicating it as "important" and 25% as "very important".
- All respondents believe that a transparent validation procedure is necessary in medical practice to address the lack of clarity in the creation of data/knowledge assets.
- Most (62.5%) respondents support updates of the data/knowledge assets "on the appearance of important studies/findings related to your subspecialty", while the rest (37.5%) consider that it should be done "periodically in 5/10/15 years". No one selected "on demand (only at the personal request of the medical practitioner)".

- Dealing with both quantitative and qualitative values, all respondents think that a duplication (with providing an interpretation, explanation, or evidence base) of qualitative values by quantitative ones would be beneficial.
- All respondents indicated that they seek additional information to fully understand the parameters/values of a scoring system.
- The vast majority (87.5%) emphasized the need for an additional grounding/explanation in scoring systems, especially in cases when there are small differences between cut-off points.
- Most (62.5%) of respondents reported encountering cases in their medical practice where small differences between cut-off points implied different conclusions, answering "yes" to the corresponding question.

Our exploratory study has certain limitations, as the sample size was small. However, it provides valuable insights for future research and can serve as a basis for further surveys on a larger scale, possibly focusing on different medical subdomains and specific problems.

5 Conclusions

Our findings emphasize the importance of information tools in the domain of medical diagnostics as a support in decision-making processes as a “second opinion”.

We identified and highlighted that "trust in historical data decreases over time" during the revision of our existing data and knowledge assets. This is a significant problem that affects the longevity of information tools and was confirmed by over a third of the respondents, making it particularly important in the medical field. Therefore, it is recommended to seek and propose solutions to address this concern, especially considering that all the respondents believe it is important (and very important) to keep data/knowledge assets relevant over time. They also agree that a transparent validation process is necessary.

Our survey revealed that the vast majority of respondents emphasized the need for an additional grounding/explanation in scoring systems, especially in cases when there are small differences between cut-off points. This finding aligns fully with our expectations.

Acknowledgments. The project in the framework of stimulating excellence in research 24.80012.5007.24SE ”BOOSTing decision making: applying an Integrated Decision Intelligence approach to overcome the limitations and challenges of traditional scoring systems” and the institutional project 011301 ”Information systems based on Artificial Intelligence” (Republic of Moldova) have supported this research.

The authors thank all the medical professionals from the Medical Center 'AnaMaria-Med' (Chisinau) who participated in the study for their cooperation.

References

[1] E. Ammenwerth, J. Brender, P. Nykänen, H. U. Prokosch, M. Rigby, and J. Talmon, “Visions and strategies to improve evaluation of health information systems: Reflections and lessons based on the HIS-EVAL workshop in Innsbruck,” *International Journal of Medical Informatics*, vol. 73, no. 6, pp. 479–491, 2004. DOI: https://doi.org/10.1016/J.IJMEDINF.2004.04.004.

[2] R. Capilla, E. Y. Nakagawa, U. Zdun, and C. Carrillo, “Toward Architecture Knowledge Sustainability: Extending System Longevity,” *IEEE Software*, vol. 34, no. 2, pp. 108–111, Mar-Apr 2017. DOI: https://doi.org/10.1109/MS.2017.54.

[3] Iu. Secrieru, C. Țâmbală, E. Guțuleac, O. Popcova, and A. Spătaru, “IDI to Enhance the Longevity of the Knowledge Bases Established on Scoring Systems,” in *2024 E-Health and Bioengineering Conference (EHB)*, (Iasi, Romania), November 14-15, 2024. DOI: https://doi.org/10.1109/EHB64556.2024.10805691.

[4] D. J. Power and R. Sharda, "Model-driven decision support systems: Concepts and research directions," *Decision Support Systems*, vol. 43, no. 3, pp. 1044–1061, 2007. DOI: https://doi.org/10.1016/j.dss.2005.05.030.

[5] N. Desai and J. Gross, "Scoring systems in the critically ill: uses, cautions, and future directions," *BJA Education*, vol. 19, no. 7, pp. 212–218, 2019. DOI: https://doi.org/10.1016/j.bjae.2019.03.002.

[6] A. Scarffe, A. Coates, K. Brand, and W. Michalowski, "Decision threshold models in medical decision making: a scoping literature review," *BMC Medical Informatics and Decision Making*, vol. 24, Article No: 273, 2024. DOI: https://doi.org/10.1186/s12911-024-02681-2.

[7] C. W. Mangus and P. Mahajan, "Decision Making: Healthy Heuristics and Betraying Biases," *Critical Care Clinics*, vol. 38, no. 1, pp. 37–49, 2022. DOI: https://doi.org/10.1016/j.ccc.2021.07.002.

[8] B. S. Patel, E. Steinberg, S. R. Pfohl, and N. H. Shah, "Learning decision thresholds for risk stratification models from aggregate clinician behavior," *Journal of the American Medical Informatics Association*, vol. 28, no. 10, pp. 2258–2264, 2021. DOI: https://doi.org/10.1093/jamia/ocab159.

[9] C. Gaindric, O. Popcova, S. Puiu, Iu. Secrieru, E. Guțuleac, and S. Cojocaru, "An approach to structure information regarding patient diagnostics in the form of taxonomy in management of mass casualty disasters," in *Proceedings of the Workshop on Intelligent Information Systems WIIS2020*, (Chisinau, Republic of Moldova), December 4-5, 2020, pp. 75–82. ISBN: 978-9975-68-415-6.

[10] C. Tambala and Iu. Secrieru, "Portal hemodynamics disorders severity in liver cirrhosis assessment by duplex ultrasound,"

Curierul medical, vol. 59, no. 1, pp. 37–40, 2016, Chisinau, Republic of Moldova. ISSN: 1857-0666.

[11] C. Gaindric, S. Șandru, S. Puiu, O. Popcova, Iu. Secrieru, and E. Guțuleac, "Advanced pre-hospital triage based on vital signs in mass casualty situations," in *Proceedings of the Workshop on Intelligent Information Systems WIIS2021*, (Chisinau, Republic of Moldova), October 14-15, 2021, pp. 130–134. ISBN: 978-9975-68-438-5.

Olga Popcova[1], Elena Guţuleac[2], Iulian Secrieru[3], Carolina Ţâmbală[4]

[123] Affiliation: Vladimir Andrunachievici Institute of Mathematics and Computer Science, Moldova State University
[4] Affiliation: Nicolae Testemiţanu State University of Medicine and Pharmacy
E–mails:
[1]olga.popcova@math.usm.md,
[2]elena.gutuleac@math.usm.md,
[3]iulian.secrieru@math.usm.md,
[4]tambala.carolina@gmail.com
ORCIDs:
[1]https://orcid.org/0009-0005-2186-2835,
[2]https://orcid.org/0009-0009-4925-9295,
[3]https://orcid.org/0009-0007-7429-0417

Proceedings of Workshop on Intelligent Information Systems
WIIS2025, October 16-18, 2025, Chisinau, Republic of Moldova

Concept of a scoring system as a solution for minimization of possible conflicts in university course timetabling

Andrei Rusu, Elena Rusu, Iulian Secrieru,
Olga Popcova, Elena Guţuleac

Abstract

A concept of a scoring system for university timetables is proposed in order to facilitate establishing a university timetable that respects its hard and soft restrictions. The idea is to ease the creation of the initial timetable that is closer to the optimal one. The concept can be used further to automate parts of a computer-aided design system for university timetables.

Keywords: scoring system, CAD system, university timetable, mathematical models.

1 Introduction

The process of creating a timetable for university courses is a complex and demanding task. In many faculties in Moldova, this activity is still carried out in a traditional way, using paper sheets to track schedules for student groups, professors, and classrooms. Any single change made to the timetable of a student group, a professor, or a classroom requires corresponding updates in all related timetables, which makes the process highly error-prone and time-consuming.

Although there are many software systems available on the market designed to automate the scheduling process and generate high-quality timetables more efficiently, these tools are often not used in practice.

The main reasons include their complexity, high cost, lack of compliance with the specific requirements of the faculty, or simply that the generated results do not meet expectations.

Comprehensive surveys on university timetable systems can be found in [1, 2, 3].

An alternative approach was proposed in [4], where a relatively simple computer-aided design (CAD) system was introduced. This solution relied on widely familiar tools, such as Microsoft Excel and Google Spreadsheets, enhanced with their built-in scripting capabilities (VBA for Excel and JavaScript for Google Docs). The advantage of this approach is that it builds on widely known software, thus lowering the learning curve and reducing resistance to adoption.

One of the main challenges in constructing an initial timetable is establishing appropriate priorities when assigning time slots – for example, for administrative staff, visiting professors, or other specific categories – while simultaneously respecting the hard constraints of the schedule (e.g., room availability, course overlaps, or legal workload limits). A common practice is to allocate priority to these special cases first and then refine the timetable in collaboration with all involved stakeholders (professors, students, and administrative staff). The closer the initial timetable is to an optimal solution, the less effort is required for subsequent adjustments.

In this paper, we propose the concept of a scoring system as a means of facilitating the generation of the initial timetable. The scoring system assigns numerical values to timetable configurations by aggregating contributions from both hard and soft constraints. Hard constraints (such as non-overlapping sessions or room capacity) are strictly enforced and penalized heavily in the score, while soft constraints (such as preferences of professors or distribution of classes across the week) are assigned moderate weights. In this way, the scoring function not only quantifies the quality of a timetable but also provides a structured mechanism for guiding the search process toward feasible and near-optimal solutions.

Such an approach naturally enables the integration of optimization techniques, such as heuristic search, metaheuristics, or reinforcement learning methods, where the scoring function serves as the reward signal. It creates the basis for developing adaptive and intelligent scheduling systems capable of improving over time as new constraints or preferences are introduced.

2 Problem formulation

2.1 Informal Problem Formulation

The *timetabling problem* is generally defined as the optimal allocation of resources subject to a set of constraints, and it admits several variations depending on the application domain. In this paper, we consider a standard case representative for many faculties in Moldova and Romania. For simplicity, we assume that all course activities are organized in the same manner within a weekly schedule.

Let us denote the main entities involved:

- A set of rooms
$$S = \{S_1, S_2, \ldots, S_{N_S}\},$$
where each room is characterized by specific properties such as seating capacity, availability of a projector, number of computers, and other facilities.

- A set of study groups of students
$$G = \{G_1, G_2, \ldots, G_{N_G}\}.$$

- A set of professors
$$P = \{P_1, P_2, \ldots, P_{N_P}\}.$$

- A set of week-days
$$Z = \{\text{Mo}, \text{Tu}, \text{We}, \text{Th}, \text{Fr}, \text{Sa}, \text{Su}\},$$

and a set of standard teaching intervals

$$T = \{8{-}10,\ 10{-}12,\ 12{-}14,\ 14{-}16,\ 16{-}18,\ 18{-}20\}.$$

A *timeslot* is defined as the pair consisting of a given day and a given time interval.

- A set of courses

$$D = \{D_1, D_2, \ldots, D_{N_D}\},$$

each course being associated with one or more activity types: *Course*, *Seminar*, *Laboratory*. These activities are allocated to specific student groups.

Each professor is assigned to teach a subset of these activities to certain student groups. Every such tuple

$$(P_i, G_j, D_k)$$

is called a *timetable event* (or teaching activity).

The problem is to allocate to each timetable event a *day*, a *time interval*, and a *room*, in such a way that the following **hard constraints** are satisfied:

1. Each professor can participate in at most one event in a given timeslot.

2. Each student group can participate in at most one event in a given timeslot.

3. At most one event can take place in the same room during a given timeslot.

In addition to hard constraints, there exist **soft constraints**, which reflect the preferences of students, professors, or the administration.

While some soft constraints represent desirable conditions (e.g., avoiding late evening classes), others may effectively behave as hard constraints (e.g., specific availability restrictions of professors or administrative requirements, deviations from professor preferences).

The objective is therefore to construct a timetable that satisfies all hard constraints and maximizes the satisfaction of soft constraints.

2.2 Compact Mathematical Formulation

Sets.

- $E = \{1, \ldots, N_E\}$: set of teaching events (each event = one course–group–type taught by a professor).
- $R = \{1, \ldots, N_R\}$: set of rooms.
- $H = \{1, \ldots, N_H\}$: set of timeslots (day $\times$ interval).

Parameters.

- $a_{e,p} \in \{0, 1\}$: event e is taught by professor p.
- $b_{e,g} \in \{0, 1\}$: event e involves student group g.
- $s_{e,r} \in \{0, 1\}$: room r is suitable for event e (equipment/type).
- $\mathrm{cap}_r \in \mathbb{N}$: capacity of room r.
- $\mathrm{size}_e \in \mathbb{N}$: required seats for event e (e.g., group size).
- $\mathcal{A}^P_{p,h} \in \{0, 1\}$: professor p is available at timeslot h.
- $\mathcal{A}^R_{r,h} \in \{0, 1\}$: room r is available at timeslot h.
- $\mathcal{A}^G_{g,h} \in \{0, 1\}$: group g is available at timeslot h.

Decision variables: $x_{e,r,h} \in \{0, 1\}$, where $x_{e,r,h} = 1$, if event e is assigned to room r at timeslot h.

Hard constraints.

- each event exactly once

$$\sum_{r \in R} \sum_{h \in H} x_{e,r,h} = 1, \quad \forall e \in E \tag{1}$$

- no professor overlap

$$\sum_{e \in E} \sum_{r \in R} a_{e,p}\, x_{e,r,h} \leq 1, \quad \forall p,\ \forall h \in H \tag{2}$$

- no group overlap

$$\sum_{e \in E} \sum_{r \in R} b_{e,g}\, x_{e,r,h} \leq 1, \quad \forall g,\ \forall h \in H \tag{3}$$

- at most one event per room/timeslot

$$\sum_{e \in E} x_{e,r,h} \leq 1, \quad \forall r \in R,\ \forall h \in H \tag{4}$$

- room suitability

$$x_{e,r,h} \leq s_{e,r}, \quad \forall e, r, h \tag{5}$$

- capacity constraint

$$\text{size}_e \leq \text{cap}_r + M(1 - x_{e,r,h}), \quad \forall e, r, h \tag{6}$$

- room availability

$$x_{e,r,h} \leq \mathcal{A}^R_{r,h}, \quad \forall e, r, h \tag{7}$$

Soft constraints and scoring. We define a weighted penalty function:

$$\min f(x) \;=\; \sum_{m=1}^{M} w_m\, \Phi_m(x), \tag{8}$$

where each term $\Phi_m(x)$ measures the violation of a soft constraint, and $w_m > 0$ is its associated weight.

Typical examples include:

- $\Phi_{\text{late}}(x)$: penalizes late evening classes (e.g., 18–20).
- $\Phi_{\text{gaps}}(x)$: penalizes gaps in the schedule for professors or groups.
- $\Phi_{\text{prefP}}(x)$: penalizes deviations from professor preferences.
- $\Phi_{\text{balance}}(x)$: penalizes unbalanced distributions of classes across days.
- $\Phi_{\text{moves}}(x)$: penalizes consecutive changes of building/room.

For example, the penalty for late classes can be defined as:

$$\Phi_{\text{late}}(x) = \sum_{e \in E} \sum_{r \in R} \sum_{h \in H_{\text{late}}} x_{e,r,h}.$$

The overall objective is therefore to find a feasible assignment x that satisfies all hard constraints (1)–(7) and minimizes $f(x)$.

3 Proposed solution

The concept of the scoring system proposed below is intended to facilitate the creation of an initial university timetable that is sufficiently close to an optimal solution. The underlying idea is inspired by well-known methods used in operations research for generating initial feasible solutions in transportation problems, such as the north-west corner method and the minimum-cost method, where homogeneous goods are distributed from a fixed number of suppliers to a fixed number of customers (see, for example, [5]).

We use the scoring function to *construct* an initial timetable that satisfies all hard constraints and accommodates as many soft constraints as possible. The process is incremental:

1. **Initialization.** The timetable is empty. At each step we select one teaching event and one feasible assignment (room, timeslot).
2. **Score-guided move.** For every feasible candidate move, we evaluate its *marginal score* and choose the best one.

3. **Iterate.** We commit the best move, update the partial timetable and repeat until all events are placed.

State, actions, feasibility. Let x denote the current (partial) assignment and E_{rem} – the set of unassigned events. For an unassigned event $e \in E_{\text{rem}}$, a candidate action is a triplet (e, r, h). Feasibility is enforced by the hard constraints:

$$\begin{aligned}\mathcal{F} \;=\; \{(e,r,h) \mid \;\; & \text{no professor/group/room conflicts at } h,\\ & \text{room suitable and available}\}.\end{aligned}$$

We *only* consider $(e, r, h) \in \mathcal{F}$.

Marginal (incremental) scoring. Let $f(x) = \sum_{m=1}^{M} w_m \, \Phi_m(x)$ be the weighted penalty over soft constraints. For a feasible candidate (e, r, h), define the marginal cost

$$\Delta f\big((e,r,h) \mid x\big) \;=\; f\big(x \cup \{(e,r,h)\}\big) \;-\; f(x).$$

We select the move minimizing Δf (ties broken by deterministic or randomized rules).

Local, decomposed updates. To keep the procedure fast, each soft term Φ_m should be *locally updatable*:

$$\Phi_m(x \cup \{(e,r,h)\}) \;=\; \Phi_m(x) \;+\; \Delta\Phi_m^{(e,r,h)}(x),$$

where $\Delta\Phi_m^{(e,r,h)}$ depends only on entities affected by (e, r, h) (e.g., the same professor/group/day, adjacent timeslots for gap penalties, etc.).

Pseudocode (generic).

Algorithm 1: Greedy constructive scheduling guided by marginal score

Input : Events E, rooms R, timeslots H; weights w_m; scoring terms Φ_m; feasibility filters.
Output: Feasible timetable x.

$x \leftarrow \emptyset$; $E_{\text{rem}} \leftarrow E$.
while $E_{rem} \neq \emptyset$ **do**
  choose next event $e \in E_{\text{rem}}$ (e.g., by a priority heuristic);
  $\mathcal{C} \leftarrow \{(e, r, h) \in R \times H \mid (e, r, h) \in \mathcal{F}(x)\}$;
  if $\mathcal{C} = \emptyset$ **then**
    repair or backtrack (or relax minor soft terms);
    continue
  end
  compute $\Delta f((e, r, h) \mid x)$ for all $(e, r, h) \in \mathcal{C}$ using local updates;
  pick $(e^\star, r^\star, h^\star) = \arg\min_{(e,r,h)\in\mathcal{C}} \Delta f((e, r, h) \mid x)$;
  $x \leftarrow x \cup \{(e^\star, r^\star, h^\star)\}$;
  update local caches for Φ_m around entities affected by $(e^\star, r^\star, h^\star)$;
  $E_{\text{rem}} \leftarrow E_{\text{rem}} \setminus \{e^\star\}$.
end

Event ordering (prioritization). To reduce conflicts early, we prioritize events by a static or dynamic key, e.g.:

$$\text{key}(e) = (\text{fewest feasible slots}) \oplus (\text{largest group}) \oplus (\text{strict availabilities}),$$

or any lexicographic combination thereof.

Tie-breaking and diversification. When multiple candidates share the same Δf, we can use (i) deterministic tie-breakers (e.g., prefer earlier hours, fewer room moves), or (ii) randomized selection (e.g., ε-greedy) to escape myopic choices. A light *beam search* can keep the K best partial states to increase robustness.

Correctness and objective. Hard constraints are maintained by construction (candidate set $\mathcal{C}$). The process terminates when all events are assigned. The resulting timetable x is a feasible solution minimizing accumulated marginal penalties in a greedy sense; it forms a strong seed for later local improvements (e.g., swap/2-opt) or for RL/Metaheuristics, where $-\Delta f$ serves as a stepwise reward.

Notes on implementation.

- Pre-filter infeasible (e, r) by suitability/capacity; pre-filter (r, h) by availability.
- Maintain fast indices by professor, group, room, day, and timeslot to compute $\Delta\Phi_m$.
- Typical soft terms with local updates: late-slot penalty, gaps for a professor/group (depend on neighbors of h), day balance, room-change penalty along a day, preference (professor/day/time) penalties.

4 Conclusion

This paper introduced a scoring-based approach for creating university timetables. The method facilitates the creation of an initial feasible schedule that satisfies all hard constraints while incorporating soft constraints as much as possible, thus minimizing later adjustments. Inspired by classical operations research heuristics, the scoring function guides the incremental process toward solutions close to optimal.

The concept supports both automatic and semi-manual design, offering flexibility for diverse institutional contexts. Future work includes integrating AI tools to capture and formalize soft constraints from natural language input, as well as exploring reinforcement learning and metaheuristic methods where the scoring function serves as a reward signal. Another promising research direction is the use of non-classical logics (e.g., diagonalizable, paraconsistent, and modal logics) to enhance the representation of constraints and decision rules, building on results such as those in [6], [7], and [8]. Overall, the approach provides a solid basis for developing adaptive and user-friendly timetable systems.

Acknowledgments. The National Agency for Research and Development (Moldova) has supported part of the research for this paper through the research project 011301 ”Information Systems based on Artificial Intelligence”.

References

[1] Hamed Babaei, Jaber Karimpour, and Amin Hadidi, “A survey of approaches for university course timetabling problem,” *Computers & Industrial Engineering*, vol. 86, 2015, pp. 43–59, DOI: https://doi.org/10.1016/j.cie.2014.11.010

[2] M. C. Chen, S. N. Sze, S. L. Goh, N. R. Sabar, and G. Kendall, “A Survey of University Course Timetabling Problem: Perspectives, Trends and Opportunities,” *IEEE Access*, vol. 9, pp. 106515–106529, 2021, DOI: https://doi.org/10.1109/ACCESS.2021.3100613.

[3] H. Alghamdi, T. Alsubait, H. Alhakami, A. Baz. “A review of optimization algorithms for university timetable scheduling,” *Engineering, Technology & Applied Science Research*, vol.10, 2020, pp. 6410-6417, DOI: https://doi.org/10.48084/etasr.3832

[4] A. Rusu and E. Rusu, “A Simple CAD System for Faculty Timetabling Design,” in *International Conference dedicated to the 60th anniversary of the foundation of Vladimir Andrunachievici Institute of Mathematics and Computer Science, MSU: Proceedings IMCS-60, 10-13 octombrie 2024, Chişinău*, (Chisinau, Republic of Moldova), 2024, pp. 361–364. ISBN: 978-9975-68-515-3, https://ibn.idsi.md/ro/vizualizare_articol/223048

[5] P. Moloşnic and E. Rusu, *Matematici speciale: Note de curs*, Chişinău: Tehnica-UTM, 2023, 288 p., ISBN: 978-9975-45-983-9, https://repository.utm.md/handle/5014/25113

[6] A. Rusu and E. Rusu, “On Algorithms Related to Expressibility of Functions of Diagonalizable Algebras,” in *Studies in Fuzziness and Soft Computing*, Berlin:

Springer Science and Business Media Deutschland GmbH, 2021, pp. 257–277. ISSN: 1434-9922, DOI: https://doi.org/10.1007/978-3-030-61334-1_13

[7] A. Rusu and E. Rusu, "On weak-functionally complete systems of formulas containing paraconsistent negation in a paraconsistent logic," in *International Conference on Intelligent Information Systems,* (Chisinau, Republic of Moldova), 2020, pp. 170–175. ISBN: 978-9975-68-415-6, https://ibn.idsi.md/ro/vizualizare_articol/116217

[8] A. Rusu and E. Rusu, "On some classes of formulas in S5 which are precomplete relative to existential expressibility," *Computer Science Journal of Moldova*, vol. 31, nr. 3(93), pp. 395–408, 2023. ISSN: 1561-4042. DOI: https://doi.org/10.56415/csjm.v31.21.

Andrei Rusu[1,6,8], Elena Rusu[2,7], Iulian Secrieru[3,6], Olga Popcova[4,6], Elena Guţuleac[5,6]

[1]ORCID: https://orcid.org/0000-0002-0259-3060
E–mail: andrei.rusu@math.usm.md
E–mail: andrei.rusu@365.univ-ovidius.ro

[2]ORCID:
https://orcid.org/0000-0002-2473-0353
E–mail: elena.rusu@mate.utm.md

[3]ORCID: https://orcid.org/0009-0007-7429-0417
E–mail: iulian.secrieru@math.usm.md

[4]ORCID: https://orcid.org/0009-0005-2186-2835
E–mail: olga.popcova@math.usm.md

[5]ORCID: https://orcid.org/0009-0009-4925-9295
E–mail: elena.gutuleac@math.usm.md

[6]Moldova State University, Vladimir Andrunachievici Institute of Mathematics and Computer Science

[7]Technical University of Moldova, Dep. of Mathematics

[8]Ovidius University of Constanţa, Dep. of Mathematics and Informatics

Cultivating Awareness: Using Augmented Reality to Combat Plant Blindness in Home Environments

Inga Ţiţchiev, Olesea Caftanatov,
Dan Dancuţa, Arcadie Fron, Dimitrii Bumbac

Abstract

Plant care can be a frustrating experience for many, especially when lacking basic knowledge about the plant's identity and needs. This challenge, often caused by "plant blindness", a tendency to overlook or misidentify plant species, affects both novice and seasoned plant enthusiasts. Traditional identification methods, such as books or static images, offer limited engagement and utility. To address this, we propose an innovative mobile augmented reality (AR) application designed to identify home and decorative plants in real time and provide personalized, comprehensive care guidance.

Keywords: augmented reality, plant blindness, 3D modeling, plant care, ecological literacy

1 Introduction

What if the houseplant beside your window was silently signaling distress – and you didn't even know its name? Each day, countless houseplants suffer not from overt neglect, but from a more insidious phenomenon known as *plant blindness* – the cognitive bias that causes individuals to overlook, undervalue, or misidentify the plants in their immediate environment [1]. This perceptual gap results in diminished awareness of plant care needs, ecological functions, and the intrinsic value of botanical life. Rooted in educational, cultural, and psychological factors, plant blindness undermines both environmental literacy and sustainable behaviors within domestic spaces.

Scholars [6] have emphasized that this issue extends beyond individual perception, pointing to systemic shortcomings in education and science communication. Plants are often relegated to the background in biology curricula, with animals occupying the central focus. As a result, students at all levels – from early childhood to university – develop weaker associations with plants, contributing to disinterest and poor recall. To address this imbalance, researchers argue that plants must be given a more central role in biological education, while plant scientists themselves should actively engage the public by showcasing the fascinating features of plant life and its vital role in human affairs.

Therefore, here is where our team wants to contribute to filling this gap. Given the increasing integration of technology into daily life, this research argues that Augmented Reality (AR) offers a promising solution for cultivating awareness. By making plant features, identities, and needs visible through interactive, real-time digital overlays, AR has the potential to counteract plant blindness, fostering a more informed and empathetic relationship between humans and their home-based flora.

Section 2 discusses briefly the state-of-the-art research on the impact of plant blindness, outlining how it affects human perception and interaction with plants. Moreover, we will discuss some traditional approaches that have been proposed to address this issue, and conclude with modern technology used for cultivating greater awareness of this phenomenon.

To fully realize this potential, however, accurate and visually compelling representations of plants are essential; thus, exploring the development of three-dimensional (3D) models is being conducted (see Section 3). As a result, fifteen common houseplants were modelled using professional software tools, such as Blender, Maya, ZBrush, and Substance Painter. The main purpose of these models is the optimization for AR environments and educational platforms, which motivates learners to explore botanical structures and care requirements with scientific accuracy and visual realism.

Beyond recognition and interaction, plant education must also address a critical challenge: health and disease management; thus, Section 5 we dedicated to this subject.

2 Plant Blindness: Concept and Traditional Approaches

The term *plant blindness* was first introduced in 1999 by Wandersee and Schussler [2] to describe the widespread human tendency to ignore, overlook, or undervalue plants in everyday life. At its core, this concept captures more than simple inattentiveness: it encompasses the inability to notice plants in one's surroundings, the failure to recognize their fundamental roles within ecosystems, and an ingrained anthropocentric bias that places greater importance on animals and humans than on botanical life. In this sense, plant blindness is not merely a perceptual oversight but a deeply rooted cognitive and cultural phenomenon that shapes how societies conceptualize and interact with the natural world.

People who are "plant blind" fail to attend to plants as individual biological units, but rather group them together into a large green backdrop. One explanation for this phenomenon is that humans pay the most attention to items that are within 15 degrees above or below the midline of their vision [7], leaving out objects low to the ground (grasses and herbs) or high above our heads (trees). Another explanation is that our brain searches for visual cues to distinguish individual objects, and a mass of green plants is not as distinct as the shape of a moving animal [8]. This evidence collectively points to how important attention is to the idea of "plant blindness", but this differential attention has cascading impacts on attitude, knowledge, and relative interest in plants versus animals.

For many people, plants are simply less engaging than animals. Students often show a preference for studying animals over plants and tend to recall plant-related information less accurately, while plants are frequently underrepresented in educational materials. This aspect was explored by Balas et. al., [3]; they employed an established paradigm in visual cognition, the "attentional blink," to examine how images of plants and animals compete for attentional resources. The results indicate that participants detect animals more reliably than plants in rapid sequences of images, and that the refractory period of visual attention differs when a plant is detected. These findings suggest that the human visual system processes plants differently from animals, providing a perceptual basis for the plant blindness phenomenon.

Despite the crucial roles plants play in sustaining ecosystems and human life, it is considered a significant challenge in biological education. Aragón et al. advocate for the use of gardens as teaching-learning contexts in Early Childhood Education (ECE) to promote an approach to the plant world. His [4] study examines preschool children's mental representations of a garden before and after an educational project. It analyzes pre- and post-project drawings from 39 four-year-olds and 43 five-year-olds using two methods: (i) a quantitative approach, counting the presence of curriculum-related elements (sun, water, earth, animals, plants, trees), and (ii) a qualitative approach, assessing the richness and specificity of these elements. Findings highlight children's interest in and emotional connection to gardens, as well as their science learning. The study emphasizes that focusing on specific curricular scientific content in early education can shape children's mental models and support later scientific understanding.

Traditional strategies to address plant blindness – such as curriculum changes and teacher-led activities – remain important, but the computing era has introduced powerful new tools. One example is iNaturalist [9], a community science platform where users record and identify species. Within it, Projects allow observations to be grouped around specific themes and include features like dedicated pages and journals for member communication. While projects can be useful, the most meaningful contribution comes from making and identifying observations.

Another study, [5], introduces AImPLANT, a prototype artificial intelligence-based mobile application, tested for the first time with prospective science teachers during outdoor activities. The app supports botanical education by offering plant identification, an AI-based chatbot, and virtual herbarium modules, thereby linking experiential learning with cutting-edge technology. Findings showed that AImPLANT significantly enhanced participants' academic knowledge of plants and successfully mitigated plant blindness, as evidenced by the Plant Awareness Test, which revealed a shift toward prioritizing plants over animals in spontaneous organism listings. However, as a prototype, the application also displayed limitations, including errors in plant recognition, slow image upload times, and occasional screen freezes, highlighting the need for further development to ensure stability, accuracy, and a smoother user experience.

3 3D Plant Modeling

Three-dimensional (3D) models of plants are essential tools for realistic visualization, interactive learning, and virtual exploration of plant structures. They support a variety of applications, including botanical education, disease detection, agricultural planning, and gamified experiences in AR and VR environments. By providing detailed anatomical representations, 3D models allow users to examine plant morphology, simulate growth and disease progression, and integrate AI-based tools for informed plant care and analysis.

In our project, we have created fifteen common houseplant models using professional software such as Blender, Maya, ZBrush, and Substance Painter, see Figure 1. These models are optimized for augmented reality environments, combining scientific accuracy with visual realism. Each model includes a dedicated interface with key botanical information – species and family, origin, flowering period, morphological traits, medicinal properties, and cultivation guidelines – along with modules addressing common plant diseases. This integration of visual and textual information allows users to interactively explore and learn about plants in a way that is both immersive and educational.

The modeling framework also enables dynamic simulations, including plant growth and life-cycle visualization. When combined with AR and AI technologies, these models can provide interactive guidance for plant care, monitor environmental conditions, and support gamified educational experiences. By bridging scientific data with interactive digital experiences, our 3D models contribute to increasing botanical literacy, fostering engagement with plant life, and supporting a new generation of plant-focused learning tools.

Figure 1. Results of 3D modeling of houseplants

Figure 1 illustrates some examples of 3D plant models used for visualization and interactive applications. The collection includes eight species from 15, with diverse morphological features: (1) Acacia – a woody species often studied for its branching patterns; (2) Strelitzia – known as the "bird of paradise" flower, with striking inflorescences; (3) Spathiphyllum – a popular indoor ornamental plant with white spathes; (4) Miltoniopsis – an orchid species characterized by its delicate flowers; (5) Aloe vera – a succulent valued for its medicinal properties and fleshy leaves; (6) Hippeastrum – a bulbous plant producing large, showy flowers; (7) Crassula – a succulent with compact growth and thick leaves; and (8) Ardisia – an evergreen plant recognizable by its red berries. Together, these 3D models illustrate the diversity of plant forms and their potential for use in education, AR/VR environments, and AI-based applications.

4 Encyclopedia of Plant Life

The encyclopedia is an ongoing initiative designed to provide an interactive and scientifically accurate environment for exploring plant anatomy, identifying species, simulating life cycles, and learning care practices. The possibility to integrate detailed botanical information with immersive 3D visualizations will allow the users to examine plants from multiple perspectives and engage with their growth and development over time.

We have developed a marker-based mobile augmented reality (AR) application that detects and tracks predefined images, called markers, using the camera of a smartphone or tablet. These markers act as triggers: when the camera recognizes one of them in the real environment, our AR system overlays virtual content – such as 3D plant models, care instructions, or animations – on top of the live camera feed, making it appear as though the digital object exists in the physical world.

By carefully selecting and optimizing our markers, we ensure accurate recognition and smooth interaction with virtual plant content. The rating of the trigger images directly affects the accuracy, speed, and stability of the AR experience. Therefore, in our application, we have developed images (see Fig. 2) with the highest possible rating in order to provide a fluid and reliable augmented reality interaction for plant identification and education.

At the moment, each plant entry includes comprehensive data such as taxonomy, origin, flowering period, morphological characteristics, and practical care recommendations, see Figure 3. For example, the Aloe genus is represented as follows:

- **Genus**: Aloe
- **Family**: Liliaceae

Figure 2. Markers used for triggering augmented plants interaction

- **Homeland**: South Africa
- **Flowering**: Irregular

Short Description: *Among indoor floriculture, the tree aloe (Aloe arborescens) is particularly popular. This succulent has a woody stem with lateral shoots and narrow fleshy leaves edged with small thorns. Its juice is valued for its medicinal properties. Another common species is the variegated aloe (Aloe variegata), which grows in a basal rosette of lanceolate fleshy leaves decorated with white dots. Both species thrive in dry environments, with an optimal winter temperature around 10 ° C.*

By combining scientific accuracy, visual realism, and interactive elements, the Encyclopedia seeks to foster botanical literacy and deepen engagement with plant life across educational and recreational contexts.

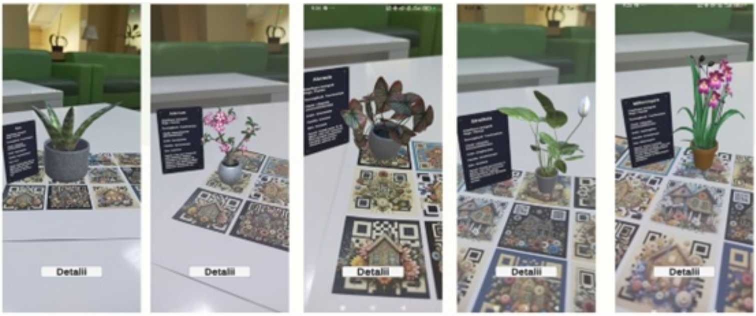


Figure 3. Some examples of plants classification.

5 Plant Diseases

Houseplants, much like outdoor vegetation, are vulnerable to fungal infections, root rot, nutrient deficiencies, and insect infestations. By embedding diagnostic cues and prevention strategies into AR visualizations, learners can be trained to recognize early symptoms before irreversible harm occurs. This integration of plant identification, 3D visualization, and health awareness offers a holistic pathway to improve both plant care at home and ecological literacy.

For example, Aloe vera is a common home plant, but, like all plants, it can suffer from diseases, especially when grown indoors under suboptimal conditions. Here are some common diseases and problems related to aloe vera that you might encounter: leaf spot, root rot, powdery mildew, etc. (see Fig. 4).

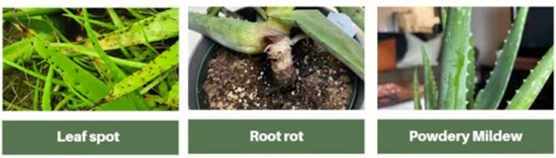


Figure 4. Some of diseases for Aloe Vera

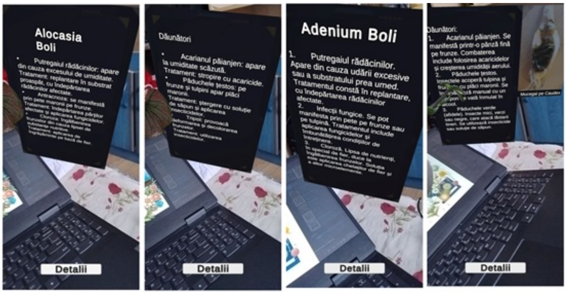


Figure 5. Screenshots with Alocasia and Adenium disease information

6 Future Work

Building on our experience with augmented reality technologies, we see AR to have tremendous potential to contribute to the fight against plant blindness and to improve plant-human interaction. Beyond serving as a visualization tool, it could form the foundation of a digital encyclopedia of plant life, allowing users to explore plants in immersive 3D, even if they have never encountered them in real life. It can also support digital twinning of houseplants, simulating plant growth to assess whether a specific pot or environment provides enough space, and assisting in disease management through interactive diagnostic overlays. Perhaps the most transformative application lies in enabling plants to "communicate" with people. By integrating soil sensors that track temperature, nutrients, pH, light availability, etc., AR systems could translate this data into intuitive messages – alerting users when plants need care or celebrating when they thrive. Through gamification, these experiences can be made engaging for children, turning plants into interactive companions and fostering empathy, awareness, and long-term botanical literacy.

7 Conclusion

Our engagement in this field is driven by strong enthusiasm, yet we also recognize that our work is at an early stage of development. As an initial step, we are constructing an Encyclopedia of Plant Life, which currently comprises 3D models of fifteen species. For each species, we have designed interfaces that provide basic botanical information – including taxonomy (species and family), morphological characteristics, typical blooming periods, medicinal properties, and guidelines for cultivation and care. In addition, the system incorporates interactive 3D visualizations of the plants as well as dedicated modules addressing their most common diseases. This foundation establishes the groundwork for progressively expanding toward more advanced and comprehensive applications in future tools.

Acknowledgments. This research was supported by SIBIA – 011301, "Information systems based on Artificial Intelligence" and CA23135 - Bringing Digital Data and Reality Together - Augmented Reality in Forestry (ARiF) projects.

References

[1] J. Sarah, W. Chih-Hang, and K. Sophien, "Overcoming plant blindness in science, education, and society," *Plants, People, Planet*, vol. 1, 2019. DOI: 10.1002/ppp3.51.

[2] J. H. Wandersee and E. E. Schussler, "Preventing plant blindness," *The American Biology Teacher*, vol. 61, no. 2, pp. 82+84+86 (3 pages), 1999.

[3] Benjamin Balas and Jennifer L. Momsen, "Attention "Blinks" Differently for Plants and Animals," *CBE—Life Sciences Education*, vol. 13, no. 3, pp. 437–443, 2014. DOI: 10.1187/cbe.14-05-0080.

[4] Lourdes Aragón, Inés Ortega-Cubero, Marcia Eugenio-Gozalbo, "Are gardens useful for promoting early-age science education? Evidence from analysis of children's drawings," *Interdisciplinary Journal of Environmental and Science Education*,

vol. 19, no. 4, Article ID: e2320, 2023. [Online]. Available: https://www.ijese.com/download/are-gardens-useful-for-promoting-early-age-science-education-evidence-from-analysis-of-childrens-13727.pdf

[5] B. Ceylan and M. A. Karakus, "Development of an Artificial Intelligence-Based Mobile Application Platform: Evaluation of Prospective Science Teachers' Project on Creating Virtual Plant Collections in terms of Plant Blindness and Knowledge," *Int. J. of Technology in Education and Science*, vol. 8, no. 4, pp. 668–688, 2024. DOI: 10.46328/ijtes.595.

[6] Jose S. B., Chien-Hsun Wu, and Sophien Kamoun, "Overcoming plant blindness in science, education, and society," *Plants, People, Planet*, vol. 1, pp. 169–172, 2019. DOI: 10.1002/ppp3.51.

[7] J. H. Wandersee and E. E. Schussler, "Toward a theory of plant blindness," *Plant Science Bulletin*, vol. 47, no. 1, pp. 2–9, 2001.

[8] J. Feldman, "What is a visual object?," *Trends in Cognitive Sciences*, vol. 7, no. 6, pp. 252–256, 2003. DOI: 10.1016/S1364-6613(03)00111-6.

[9] iNaturalist. https://www.inaturalist.org/

Inga Țîțchiev[1], Olesea Caftanatov[2], Dan Dancuța[3], Arcadie Fron[4], Dimitrii Bumbac[5]

[1,2,3,4,5] Vladimir Andrunachievici Institute of Mathematics and Computer Science, Moldova State University

[1]ORCID: https://orcid.org/0000-0002-0819-0414
E–mail: inga.titchiev@math.usm.md

[2]ORCID: https://orcid.org/0000-0003-1482-9701
E–mail: olesea.caftanatov@math.usm.md

[3]ORCID: https://orcid.org/0009-0004-0710-6198
E–mail: dan.dancuta@math.md

[4]ORCID: https://orcid.org/0000-0001-9709-483X
E–mail: arcadie.fron@gmail.com

[5]ORCID: https://orcid.org/0009-0005-9157-2171
E–mail: dimitrii.bumbac@math.md

Proceedings of Workshop on Intelligent Information Systems
WIIS2025, October 16-18, 2025, Chisinau, Republic of Moldova

Navigating GenAI Societal Transformation: The 3×2A Strategy for Human–AI Synergy

Valeriu Ungureanu

Abstract

This paper introduces the *3×2A Strategy*, a conceptual and operational framework for societal adaptation to Generative AI (GenAI). It structures six strategic dimensions – Automation, Augmentation, Alliance, Alignment, Adaptation, and Accountability – essential for fostering human-AI synergy. The strategy supports both diagnostic reflection and policy planning. Applications in education, healthcare, and scientific research illustrate its role in guiding responsible integration. By aligning GenAI advances with societal goals, institutional norms, and long-term resilience, the framework enables multi-level governance and innovation.

Keywords: Generative AI, Large Language Models, Human-AI Symbiosis, Societal Adaptation, Governance, 3×2A Strategy.

1 Introduction

Recent advances in GenAI, especially large language models (LLMs) [4], are reshaping decision-making, knowledge production, and institutional workflows. This paper proposes the 3×2A Strategy as a flexible yet structured response to these transformations. The strategy was initially introduced and its core conceptual structure outlined in [10], where its foundational premises were formulated. Here, we develop the framework more systematically, articulating its six dimensions and exploring their application across key sectors. The strategy addresses the interplay between technological innovation and societal adaptation by outlining six dimensions that together support effective governance, alignment, and collaboration between humans and AI systems [7].

2 The 3×2A Strategy Framework

The 3×2A Strategy articulates six interrelated dimensions:

- **Automation** – delegating repetitive or procedural tasks to GenAI systems.
- **Augmentation** – enhancing human cognition, creativity, and productivity.
- **Alliance** – fostering collaboration and trust in human-AI teams.
- **Alignment** – ensuring AI behavior accords with human intentions, norms, and institutional policies.
- **Adaptation** – evolving practices, tools, and mindsets in response to GenAI.
- **Accountability** – ensuring transparency, traceability, and governance of AI outcomes.

These dimensions form three conceptual axes: (i) Automation & Augmentation, (ii) Alliance & Alignment, and (iii) Adaptation & Accountability.

Each pair of dimensions forms a conceptual axis – structured like a continuum between two strategic orientations. This axial design enables a multi-dimensional mapping of societal responses, where strategies can lean toward one pole, balance both, or evolve dynamically along the spectrum.

These axes can also be interpreted as three systemic levers or *"controllers"* through which societies steer their adaptation to GenAI [3]. By adjusting the balance along each axis – automation versus augmentation, alliance versus alignment, and adaptation versus accountability

– stakeholders can influence how progress unfolds and ensure it remains both innovative and responsible.

Importantly, the 3×2A Strategy also acts as a *diagnostic and planning lens*, enabling institutions to view their GenAI practices, spot imbalances or risks, and chart context-sensitive implementation paths.

3 Applications in Key Sectors

To illustrate the operational relevance of the 3×2A Strategy, we examine its application across three societal domains where GenAI integration is already underway: education, healthcare, and scientific research. These sectors exemplify diverse challenges – ranging from epistemic and ethical to institutional and infrastructural – that the six strategic dimensions help address. Each subsection below highlights how the 3×2A framework can guide human-AI synergy, foster institutional resilience, and inform responsible innovation.

3.1 Education

GenAI can enhance pedagogical design, personalize learning, and support teacher productivity. The 3×2A Strategy emphasizes the co-evolution of curricula, digital literacy, and alignment with humanistic values. Augmentation and accountability help ensure pedagogical coherence [9].

3.2 Healthcare

GenAI tools contribute to triage, diagnosis support, medical documentation, and ambient intelligence [5]. While automation accelerates workflows, alliance and alignment build trust [8]. Accountability ensures clinical safety and regulatory compliance.

3.3 Scientific Research

GenAI enables knowledge discovery, literature synthesis, and co-authoring. The 3×2A lens supports augmentation of inquiry and creativity, while emphasizing the adaptation of peer review and reproducibility norms.

4 Related Work and Conceptual Positioning

While initiatives, such as the EU AI Act and the AI4People framework, emphasize risk management and ethical principles [2], the 3×2A Strategy offers a more integrative, sociotechnical approach. It complements interpretability tools and policy models with a structural lens focused on human-AI interaction, organizational change, and normative alignment.

5 Conclusion and Outlook

The 3×2A Strategy offers a flexible framework for aligning GenAI technologies with human values, institutional priorities, and systemic resilience. By linking six strategic dimensions across three conceptual axes – each functioning as a societal lever – the framework enables multi-dimensional steering of innovation trajectories. These levers allow stakeholders to balance competing priorities such as automation and augmentation or adaptation and accountability, promoting responsible and context-sensitive innovation.

This work provides a more detailed exposition and operationalization of the core ideas introduced in [10], serving both as a consolidation of the conceptual framework and as a foundation for more advanced future developments. Future work includes empirical validation, policy co-design, the development of domain-specific toolkits for implementation, and a more detailed articulation of the strategy itself to deepen its theoretical and practical foundations [1, 6].

Acknowledgments. This work was partially supported by the research project *SENSE (Sensory Engagement Network for Shared Experiences)*, grant no. 31PCBROMD, code PN-IV-PCB-RO-MD-2024-0392.

The author affirms that the research design, analysis, and conclusions are entirely their own. Generative AI tools (ChatGPT, Gemini, DeepSeek, and Grok) were employed solely for language refinement, stylistic improvements, and comparative reflections consistent with the paper's topic, with all outputs carefully reviewed and verified.

References

[1] D. Autor, "Work of the Past, Work of the Future," NBER Working Paper No. 25588, 2019. `https://www.nber.org/system/files/working_papers/w25588/w25588.pdf`.

[2] L. Floridi et al., "AI4People – An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations," *Minds and Machines*, vol. 28, no. 4, pp. 689–707, 2018. DOI: `https://doi.org/10.1007/s11023-018-9482-5`.

[3] I. Gabriel, "Artificial Intelligence, Values and Alignment," *Minds and Machines*, vol. 30, pp. 411–437, 2020. DOI: `https://doi.org/10.1007/s11023-020-09539-2`.

[4] M. Bommasani et al., "On the Opportunities and Risks of Foundation Models," Stanford Institute for Human-Centered AI, Tech. Rep., 2023. `https://crfm.stanford.edu/report.html`.

[5] M. Matheny et al., *Artificial Intelligence in Health Care: The Hope, the Hype, the Promise, the Peril* (The Learning Health System Series), NAM Special Publication, Washington DC: National Academy of Medicine, 2020.

[6] B. Mittelstadt et al., "The Ethics of Algorithms: Mapping the Debate," *Big Data & Society*, vol. 3, no. 2, pp. 1–21, 2016. DOI: 10.1177/2053951716679679.

[7] I. Rahwan, "Society-in-the-Loop: Programming the Algorithmic Social Contract," *Ethics and Information Technology*,

vol. 20, pp. 5–14, 2018. DOI: https://doi.org/10.1007/s10676-017-9430-8.

[8] E. Topol, *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*, Basic Books, 2019, 400 p.

[9] Fengchun Miao and Wayne Holmes, "Guidance for Generative AI in Education and Research," UNESCO, 2023, 44 p. ISBN: 978-92-3-100612-8. DOI: https://doi.org/10.54675/EWZM9535.

[10] V. Ungureanu, "Generative AI and Workforce: A Synergistic Future with LLMs," in *Mathematics & Information Technologies: Research and Education (MITRE-2025): International Conference,* June 26-29, 2025, Chişinău: CEP USM, 2025, pp. 105–106.

Valeriu Ungureanu

ORCID: https://orcid.org/0000-0001-5184-0040
Faculty of Mathematics and Computer Science, Moldova State University,
Mateevici str., 60, Chisinau, MD-2009, Republic of Moldova
E–mail: valeriu.ungureanu@usm.md

Table of contents

Firma poligrafică „VALINEX" SRL, Chişinău, str. Florilor, 30/1A, 26B, tel./fax 43-03-91,
e-mail: info@valinex.md, http:\\www.valinex.md

Coli editoriale 20,41. Coli de tipar conv. 21,97. Format 60x84 1/16. Garnitură „Times".
Hirtie ofset. Tirajul 100.