HUMAN TECHNOLOGY

ISSN: 1795-6889

https://ht.csr-pub.eu Volume 20(1), May 2024, 6-24

CONCEPTUAL FRAMEWORK FOR ETHICAL ARTIFICIAL INTELLIGENCE DEVELOPMENT IN SOCIAL SERVICES SECTOR

Miroslavas Seniutis
Institute of Sociology and Social Work
Vilnius University
Lithuania
miroslavas.seniutis@fsf.vu.lt
ORCID 0000-0002-8089-3341

Valentas Gružauskas
Institute of Computer Science
Vilnius University
Lithuania
valentas.gruzauskas@mif.vu.lt
ORCID 0000-0002-6997-9275

Angele Lileikiene
Lithuania Business College, Lithuania
angele.lileikiene@ltvk.lt
ORCID 0000-0002-8414-5906

Valentinas Navickas
Lithuania Business College, Lithuania
valentinas.navickas@ltvk.lt
ORCID 0000-0002-7210-4410

Abstract: *This research explores the domain of Artificial Intelligence (AI) for social good, with a particular emphasis on its application in social welfare and service delivery. The study seeks to establish a universal conceptual framework for ethically integrating AI into the social services sector, recognizing the sector's significant yet underexplored potential for AI utilization. The objective is to develop a comprehensive framework applicable to the ethical deployment of AI in social services, using Lithuania as a case study to illustrate its practicality. This involves analysing the political discourse on AI, examining its applications in social welfare, identifying ethical challenges, evaluating the digitalization progress in Lithuania's public services, and formulating guidelines for AI integration at various stages of delivering social services. Our methodology is rooted in document analysis, encompassing a thorough review of both normative and scientific literature pertinent to the ethical application of AI in social welfare. Key findings reveal that AI's anticipated positive impacts on diverse social and economic areas, as highlighted in political declarations, are being partially realized, as corroborated by scientific studies. Although the global application of AI in social welfare is expanding, Lithuania presents a unique case with its strategic planning gaps in this sector. The developed conceptual framework offers vital criteria for the ethical implementation of AI systems designed to be universally applicable to various stages of social services, accommodating different AI applications, client groups, and institutional environments.*

Keywords: *artificial intelligence, social service, ethics, innovations*

DOI: https://doi.org/10.14254/1795-6889.2024.20-1.1

INTRODUCTION

Background and Problem Statement

There is a field of research known as Artificial Intelligence1 (AI) for Social Good2, which focuses on using AI to benefit society as a whole. It often addresses issues that have a broader impact and are not limited to specific individuals or groups. This includes exploring large-scale social and environmental challenges (Akula & Garibay, 2021; Holzmeyer, 2021; Tomašev et al., 2020), such as how to manage natural (United Nations, 2023; Butler, 2017) and war (Cornebise et al., 2018) disasters more effectively, minimize the effects of global warming, and develop alternative energy sources (Chui et al., 2018).

Alongside this stream of research, there are AI studies that concentrate more on the well-being of individuals and communities, particularly those who are vulnerable or marginalized. Research on AI in social welfare is mostly dedicated to exploring possibilities for automation and enhancement of social welfare systems (Oravec, 2019) comprised of diverse services, policies, and programs aimed at ensuring that people's basic needs – such as health, education, housing, employment, and income security – are met.

Another area of research that can be distinguished is focused on the application of AI in providing specific social services in both the private and public sectors. This type of research is concerned with how organizational structures implement specific interventions, programs, practices, and work methods in assisting the functioning and well-being of individuals, families, or communities. It is considered that AI could be applied in assessing eligibility and needs, making enrolment decisions, providing benefits, and monitoring and managing the delivery of benefits to clients of social services (Ohlenburg, 2020). AI is also being used to conduct risk assessments, assist people in crisis, strengthen prevention efforts, identify systemic biases in the delivery of social services, provide social work education, and predict social worker burnout and service outcomes (Reamer, 2023). These benefits aligned with the efforts of steady digital society development in the EU (Kersan-Škabić & Vukašina, 2023).

However, research on the application of AI in the field of social services is not widely developed. In Lithuania, this is, on one hand, related to the lack of a strategically organized plan for financing AI scientific research, and on the other hand, both in Lithuania and worldwide, the development of AI systems in the social welfare does not proceed very rapidly. AI development is privileged in other sectors that bring more financial benefit (Ministry of Economy and Innovations, 2019). In this regard, proofs of positive impact of AI on business development are obtained by Kolková & Ključnikov (2022); Letkovsky et al. (2023); Roshchyk et al. (2022). Therefore, to promote future research on AI development and implementation in the social service sector, critically evaluated and even competing conceptual frameworks are needed, which will help in creating reliable knowledge.

[1] „Artificial Intelligence (AI) refers to systems that display intelligent behavior by analyzing their environment and taking actions – with some degree of autonomy – to achieve specific goals. AI-based systems can be purely software-based, acting in the virtual world (e.g., voice assistants, image analysis software, search engines, speech and face recognition systems), or AI can be embedded in hardware devices (e.g., advanced robots, autonomous cars, drones, or Internet of Things applications)“ (Ministry of Economy and Innovations, 2019).

[2] The similar term 'AI for Good' refers to a United Nations project. It is the leading action-oriented, global, and inclusive platform on AI. Its goal is to identify practical applications of AI that can advance the United Nations Sustainable Development Goals and scale those solutions for global impact. (United Nations, 2022).

Current Research

Research on the development and application of AI for social welfare purposes has been particularly expanding in recent years. For instance, computer science representatives like Floridi et al. (2021) and Tomašev et al. (2020) discuss fundamental ethical principles and practices in designing ethical AI or applying AI in the realm of social welfare. Application of AI in law enforcement is analysed by Carton et al. (2016), Fang et al. (2016); in transportation by Bojarski et al. (2016); in education by Lakkaraju et al. (2015); in knowledge management by Bencsik (2021) and Bilan et al. (2023); in healthcare by Ross & Swetlitz (2017); Yu et al. (2018) and Strickland (2019). While there are fewer studies focusing on the ethical aspects of AI application in the social service sector or addressing specific social problems. In the field of social sciences, management, and economics, Kim et al. (2022) provide guidelines for creating AI implementation strategies in social innovation projects. Well-known cases involve AI being used for predicting poverty risk (Jean et al., 2016), mandatory education for homeless youth (Yadav et al., 2016), revealing factors contributing to criminal victimization among homeless adults (Shah et al., 2021), automation of decisions made by child welfare specialists (Gillingham, 2021).

In Lithuania, research is conducted on AI application in medical diagnosis (Janušonytė, 2021; Venclovaitė, D., Stramkauskaitė, A., Kuzmienė, 2022); the possible impact of AI on different aspects of human consciousness is discussed from a philosophical perspective (Vidauskytė, 2021); challenges raised by AI are analysed in constitutional law, the influence of AI on human rights in law enforcement activities, etc. (Juškevičiūtė-Vilienė, 2020; Skardžiūtė, 2018; Zakaras, 2022).

Previous research has primarily focused on general ethical aspects of AI rather than the specific ethical challenges that may arise within the processes of social service delivery. This involves the development of tailored AI solutions for socially vulnerable groups, ensuring their unique needs are addressed while safeguarding their interests.

Research Objectives

The main aim of this study to devise a universally applicable conceptual framework for the ethical implementation of artificial intelligence (AI) in social services, grounded in an unstructured review of international scientific and normative literature on AI's application in social welfare and its ethical governance. This framework, while globally relevant, will include the backdrop of the Lithuanian context, offering insights into its practical application in a specific national setting.

In order to achieve this aim, the study addresses the following objectives, which reflect the main structure of this article: a) To analyse the growing expansion of AI as articulated in political declarations and its implications for society; b) To investigate and document a range of AI applications within the context of social welfare; c) To identify and scrutinize ethical concerns associated with AI; d) To evaluate the current status of digitization in the Lithuanian public sector; e) To delineate the requirements and directions for future research in the field of ethical AI-based engineering of social services.

Research Relevance

It is expected that conceptual framework applied in future empirical research will provide key insights and practical benefits for various stakeholders. Policymakers and service organizers will gain guidance on key aspects like human resource investment, essential skills for AI integration, impact measurement, risks management, and implementation of appropriate AI applications. Service providers and practitioners will benefit from guidelines on maintaining unbiased AI assessments, preserving professional autonomy, and maximizing the benefits of AI. Additionally, service clients will be equipped with information and advice on handling potential harm, ensuring transparent decision-making, and protecting their privacy in the context of AI utilization.

METHODOLOGY

This study presents an unstructured internet sources literature review of 14 normative documents on AI implementation, with a particular focus on international ethical regulations regarding AI, and 40 scientific sources on Ethical AI applications for Social Good, with particular attention to the field of social welfare (see references). This review can be typologically aligned with an expository literature review, wherein social researchers engage in scholarly discourse by offering a detailed explanation of the subject matter. They utilize pilot evidence to foster a comprehensive understanding of the topic under examination (Tayo et al., 2023). This research method was chosen because it grants access to existing knowledge, offering solutions or suggestions aligned with the main research aim, and it also guides the research towards achieving its findings (Yekeen, 2006). The main stages of review such as formulation of selection criteria, selection of sources, data extraction and analysis were performed according to guidelines of Kitchenham and Charters (2007).

The selection criteria for normative literature sources were established to identify documents that regulate AI implementation across multiple levels: nationally in Lithuania, internationally within the European Union and the United States, and globally under United Nations guidelines. Additionally, for scientific sources, the focus was on selecting articles that explore AI implementation in areas such as social good initiatives, social welfare, and public The selection criteria for normative literature sources services, with particular attention to the ethical challenges presented in these areas. As for each qualitative study, it was important that selected sources would be representative in relation to the research problem and not to the volume of available literature or the prevalence of certain viewpoints (Židžiūnaitė & Sabaliauskas, 2017).

To search for relevant normative literature, we utilized Google Search Engine. For scientific literature, our approach included accessing databases with open access, such as Google Scholar, Scopus, and Web of Science using different combination of these keywords: AI, Ethics, Regulation, Social Good, Social Welfare, Social and/or Public Service.

Data analysis in our study was performed through qualitative thematic analysis using MAXQDA 2022 software. This approach enabled us to identify and organize key themes that are crucial to achieving the study's objectives systematically and inductively. Data analysis was conducted according to the following procedure: familiarization with research data, data

coding, searching for themes, reviewing them, description, and report preparation (Broun & Clarke, 2014). These main themes are integral not only to the study's structure, reflected in its objectives and section titles, but they also form the core categories of our conceptual framework. These themes include: 1) Applications of AI; 2) AI's impact on various social service processes; 3) the process of delivering social services; 4) Ethical issues in AI usage; and 5) Strategies for risk prevention and resolution in the context of AI implementation in social services.

ARTIFICIAL INTELLIGENCE IN SOCIAL WELFARE: A LITERATURE REVIEW

Rising Expansion and Expectations Toward AI for Society in Political Declarations

There is a growing number of political declarations that not only encourage the development of artificial intelligence (AI) globally but also express the expectation that AI will contribute to social good in a broad sense and ensure social welfare specifically.

Documents from the European Commission and various international bodies over the last five years highlight substantial aspirations regarding the advancement of AI. Firstly, there is a strong ambition for the European Union (EU) to become a world leader in both the development and application of AI. It is expected that each member of the EU will contribute towards establishing the EU as a champion of an AI approach that benefits both individuals and society (European Commission, 2018), there is an optimistic view that AI will significantly contribute to creating a more sustainable society and drive economic growth. This includes improving the efficiency of healthcare and agriculture, aiding in the reduction of climate change, enhancing the productivity of manufacturing systems, and strengthening security measures (European Commission, 2020). The goal set for 2030 is for 75% of European companies to integrate cloud computing, big data, and artificial intelligence technologies. The expectation is that the adoption of AI will lead to improvements in job quality, workplace safety, efficiency, and employee well-being (European Commission, 2021a). Thirdly, one of the EU digital policy objectives for the coming decade is the development of Ethical Artificial Intelligence. This aims to foster responsible and reliable AI that benefits humanity and upholds human rights globally (European Commission, 2021c). Similarly, the Global Goals established by the United Nations in 2015 emphasize that AI should play a crucial role in advancing social welfare, protecting human rights, and promoting environmental sustainability, among other objectives (United Nations, 2015).

Thus, the ambition to become a leader in the AI industry, coupled with elevated expectations for AI's positive impact across various social and economic areas of life, and the simultaneous focus on the ethical challenges of AI implementation, may be considered key elements of the EU's political agenda related to AI development.

Exploring the Range of AI Applications in Social Welfare

Scientific studies confirm that expectations towards AI are being, to some extent, realized. Various AI applications have demonstrated success in solving a range of social problems and

in delivering diverse social services within specific institutional frameworks, thereby contributing to the enhancement of the social welfare system (Shi et al., 2020).

Illustrative of this are the following applications: a) Natural Language Processing (NLP) – algorithms capable of reading and comprehending human language have been adapted as virtual assistants for social service users, practitioners, or administrators (Chui et al., 2023); b) Computer Vision (CV) – algorithms that process visual information can enhance public safety by detecting instances of violence, locating missing persons (Trilupaitytė, 2022), aiding the visually impaired (Patel & Parmar, 2022), and assisting seniors with dementia (Bucholc et al., 2023; Chignell et al., 2020); c) Robotics – mobile systems capable of autonomous movement and environmental interaction (Burgard, 2023) can significantly aid in the care and support of the elderly (Wellman & Rajan, 2017); d) Machine Learning (ML) – this method uses historical or real-time data to predict future trends (Engin & Treleaven, 2019), assisting in identifying vulnerable tenant issues (Yeung, 2018), modelling migration patterns (Robinson & Dilkina, 2018), supporting the homeless (Abelson et al., 2014), developing victimization models (Shah et al., 2021), and optimizing food distribution and usage (Shi et al., 2020).

The presented classification of various AI applications, or in other words, the breadth of AI industries, is neither definitive nor exhaustive regarding the fields of its application. However, this overview offers essential insights into the already achieved development of AI in the field of social services in different countries around the world.

Identifying Ethical AI issues

However, numerous ethical challenges arise in relation to the implementation of mentioned AI applications in various socio-economic contexts (United Nations, 2021). These potential ethical challenges are well-summarized in the White House's (2022) Blueprint for AI-related legislation and ethical considerations.

The ethical challenges linked to AI application encompass a spectrum of issues. Firstly, there are security risks and effectiveness concerns, where AI applications might endanger users' safety and amplify deception, affecting trust and reliability. Secondly, algorithmic bias presents a problem, as automated algorithms could inadvertently favor certain demographics, leading to inequality based on age, gender, ethnicity, and more. Thirdly, data privacy issues arise from excessive data collection and use without explicit consent or for unspecified purposes. Additionally, the opacity of AI systems makes their results and processes often difficult to comprehend and verify. Lastly, there's the issue of autonomy loss, where reliance on AI leaves no alternative options for people, limiting their freedom of choice.

The ethical guidelines promoted by the (European Commission, 2019) highlight seven key requirements for AI systems to be deemed trustworthy. Additionally, the document “Artificial Intelligence Act” (European Commission, 2021b) presents a more developed and comprehensive treatment of AI requirements. Many of these requirements are in line with the previously presented White House AI regulations, for example: a) Technical robustness and safety, including resilience to attack and security, a fall-back plan and general safety, accuracy, reliability, and reproducibility. b) Diversity, non-discrimination, and fairness, encompassing the avoidance of unfair bias, accessibility and universal design, and stakeholder participation. c) Privacy and data governance, which covers respect for privacy, quality and integrity of data, and access to data. d) Transparency, including traceability, explainability, and communication.

e) Human agency and oversight, ensuring fundamental rights, human agency, and human oversight.

The additional requirements that are less developed in United States regulation include Societal and Environmental Well-being, which mandates that the entire supply chain of AI systems must be environmentally friendly and beneficial to all human beings, including future generations. It involves a critical examination of resource usage and energy consumption, particularly during the AI systems' training phase. Additionally, AI systems must ensure a positive impact on the social relationships, physical, and mental well-being of users. It is also essential to assess their influence on political processes and structures, including political decision-making and electoral contexts.

One more less developed requirement is accountability, which primarily entails the establishment of mechanisms that enable the auditing of AI systems. This means evaluating and assessing the processes and outcomes of AI systems' algorithms. AI systems should be subject to independent and continuous auditing throughout their entire lifecycle to ensure their safety. AI developers must implement risk and impact assessment procedures, and AI users should have the opportunity to report any negative impacts they might encounter.

The competition for leadership in the AI industry inevitably leads to joint efforts to define ethical principles for AI application and to formulate corresponding AI requirements. Both the European Commission and the White House have promoted a human-centric approach to AI, aiming to enhance individual and collective human well-being. This leads to the fostering of elaboration and international consensus on AI ethics.

Status Quo of the Digitization of the Lithuanian Public Sector

The digitalization of public services (health, education, legal services, etc.) in Lithuania is progressing rapidly. According to the Digital Economy and Society Index (DESI) (European Commission, 2022) Lithuania ranks tenth – 83,85 (score 0 to 10) among European Union countries in this regard. The EU average is – 77, 03 (score 0 to 100) (see figure 1).

However, the pace of AI development in the public sector is not particularly rapid, which could be attributed to the others strategic planning priorities and to the relatively modest financial resources dedicated to AI development in Lithuania. From 2015 to 2018, the public sector invested 26.5 million euros in AI development. For comparison, France's AI development plan is valued at 1.5 billion euros, and China's plan is worth 150 billion euros (Ministry of Economy and Innovations, 2019).

There are known cases where AI systems are applied in law enforcement institutions: police, border guard services, prison department, and other state institutions (Zakaras, 2022), as well as in the health sector, for example in diagnosing diseases (Janušonytė, 2021; Venclovaitė, D., Stramkauskaitė, A., Kuzmienė, 2022) and the business sector (European Commission, 2022), but there are no cases of AI application in the social services sector. It seems that AI development in Lithuania is prioritized in the financial sector. Only 4.45 % of all enterprises (excluding the financial sector) have applied artificial intelligence, whereas the European Union average is 7.91 % (see figure 2).

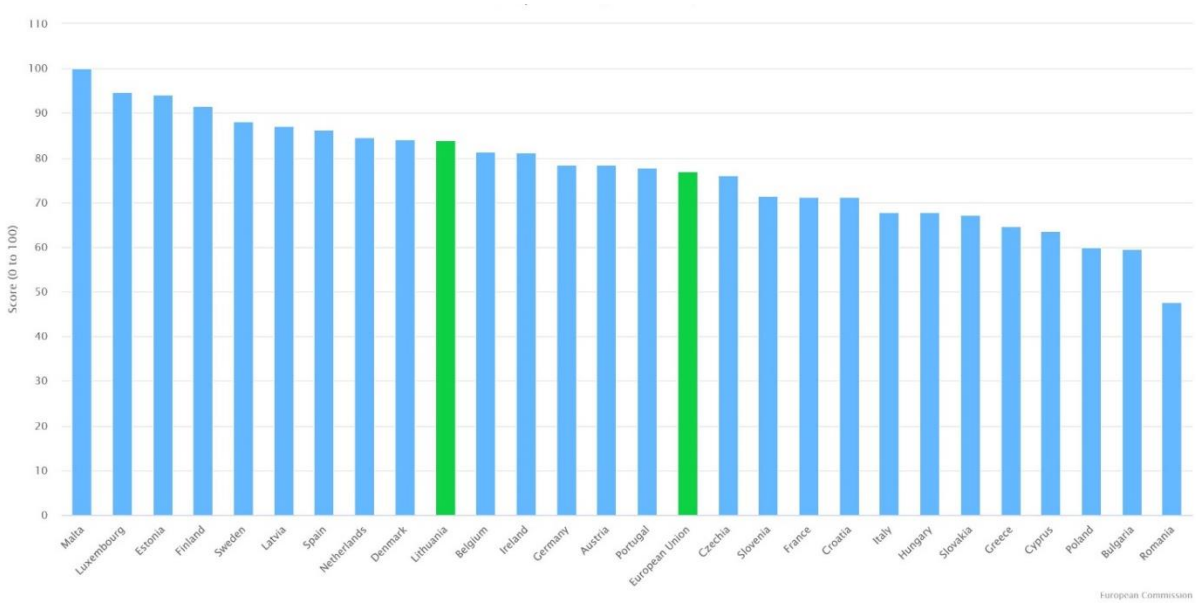

Figure 1. Digital public service for citizens (score 0 to 100), DESI 2022 (data from 2021)

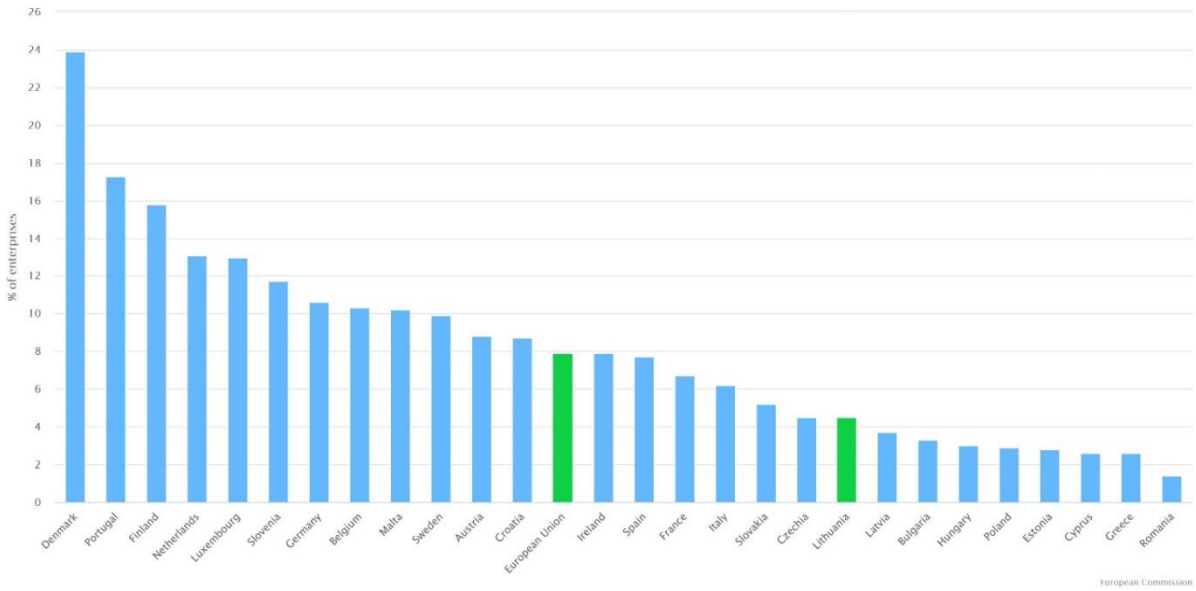

Figure 2. AI implementation across all enterprises (employing 10 or more people and excluding the financial sector) DESI 2022 (data from 2021)

In the Lithuanian Artificial Intelligence Strategy, initiated by the Ministry of Economy and Innovation in 2018-2019, multiple sectors (see figure 3) were pinpointed for prioritized AI development, chosen based on their economic benefit and sectoral impact. The strategy targets a) Manufacturing, aiming to boost productivity through automation; b) Agriculture, employing AI for tasks like robotic harvest collection and intelligent soil analysis; c) Transport, leveraging AI for traffic management and autonomous vehicles; d) Energy, utilizing AI for more efficient

energy supply methods, reducing foreign energy dependence; and e) Healthcare, integrating AI to manage increasing patient loads and paperwork, thereby optimizing healthcare delivery.

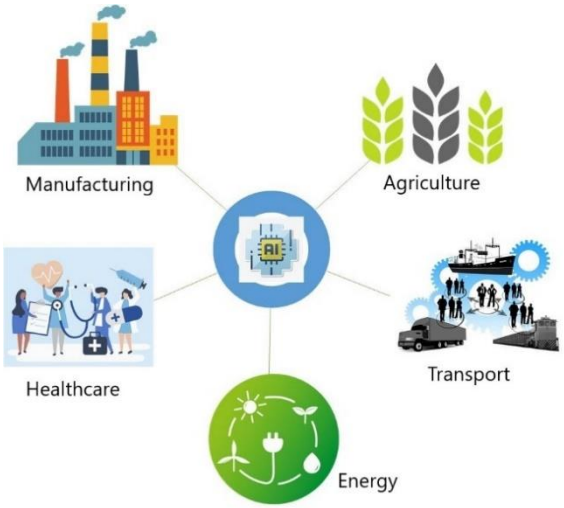

Figure 3. Priority sectors for AI development in Lithuania
Source: own elaboration.

Except the health sector, others are not directly related to social welfare. Therefore, it can be stated that the goals of social welfare are not sufficiently expressed in Lithuania's AI strategy. In a broader sense, the implementation of AI in all these sectors could eventually in long term contribute to a better quality of life of entire population. However, the urgent needs of individuals and groups in need, for whom the social welfare system is designed to provide assistance in various forms, including financial support, healthcare, education, housing, and other services, seem to be insufficiently relevant and prospective to be considered in planning AI development in the country.

In the main strategic documents concerning the state's social welfare, such as the Lithuanian Artificial Intelligence Strategy by the Ministry of Economy and Innovations (2019) and the Strategic Plan by the Ministry of Social Security and Labour (2023), there is a lack of mention of AI implementation in the field of social services. Overall, such plans will have to emerge, along with legal regulation, funding, and specific attempts to create and apply AI in the field of social services. It remains to wish that this process at the political decision-making level would go as smoothly as possible, and at the practical level, opportunities would be created to prepare for the upcoming innovations, such as professional training programs on artificial intelligence, publicly accessible resources about AI opportunities and risks, scientific research, and the like.

RESULTS

Conceptual Framework for Integrating Ethical AI into the Social Service Process

There are well-known national initiatives that involve collaborations between the government, industry (AI companies), researchers (including computer engineers and social scientists), and practitioners (such as social service organizers and providers) (Fox et al., 2022; Bartosz Gajderowicz, 2019). These initiatives are aimed at developing social service solutions based on AI systems. Interdisciplinary and cross-sectoral expert groups are working on social service engineering, which includes how each stage of the social services chain can benefit from engineering design, planning, and delivery. The inclusion of AI systems in this engineering process has the potential to enhance the effectiveness and efficiency of social services. This is achieved by delivering the right services to the right people at the right time, and by preventing and/or addressing potential ethical challenges. For example, a total of $4.9 million in funding was allocated for the implementation of the Compass project in Canada. The goal is to develop a national AI-based platform that aligns the needs of individuals, families, and communities with the right combination of social service options (Darling, 2022).

In this study, a proposed conceptual framework integrates, details, and illustrates potential relationships among four main categories: social service process, the impact of AI on various stages in social service process, AI applications, ethical issues, and risk prevention and solution (see figure 4).

Firstly, the social service process is viewed as a value chain, identifying universal stages such as: design of social services, access to services; identification of client needs; planning assistance; delivery of social services; evaluating service effectiveness (see Saunders, 2016).

Secondly, AI systems can be integrated into various stages of the social service process, potentially having a significant impact on these stages. a) AI can analyse large-scale data to simulate clients' interactions in social services identifying trends and gaps. This helps the design of services focused on real clients' needs and proven practices (Gajderowicz et al., 2014). b) It can provide personalized guidance for users, helping them navigate and access appropriate services more efficiently (Vendeville, 2022). c) AI's data analysis capabilities enhance the precise identification of individual needs by emulating and analysing client behaviour. This understanding of specific client behaviours then guides the customization of suitable assistance methods (Gajderowicz & Fox, 2017). d) AI may enhance the process of searching for and aligning appropriate services with client needs, as well as predicting the potential success rates of these services (Rosu et al., 2017). e) AI can automate routine tasks and decision-making process in service delivery (Jankovic & Bernier, 2023). f) AI can enhance the evaluation process through advanced data analytics, thereby assisting managers and policy makers in improving social services and policies (Vendeville, 2022).

Thirdly, a variety of AI applications, including natural language processing, computer vision, robotics, and machine learning, are examined for their alignment with each stage of the service process (see section 2).

Subsequently, ethical challenges are identified, encompassing issues such as security risks and effectiveness, algorithmic bias, data privacy issues, opacity, autonomy loss, societal and environmental well-being, accountability (see section 3).

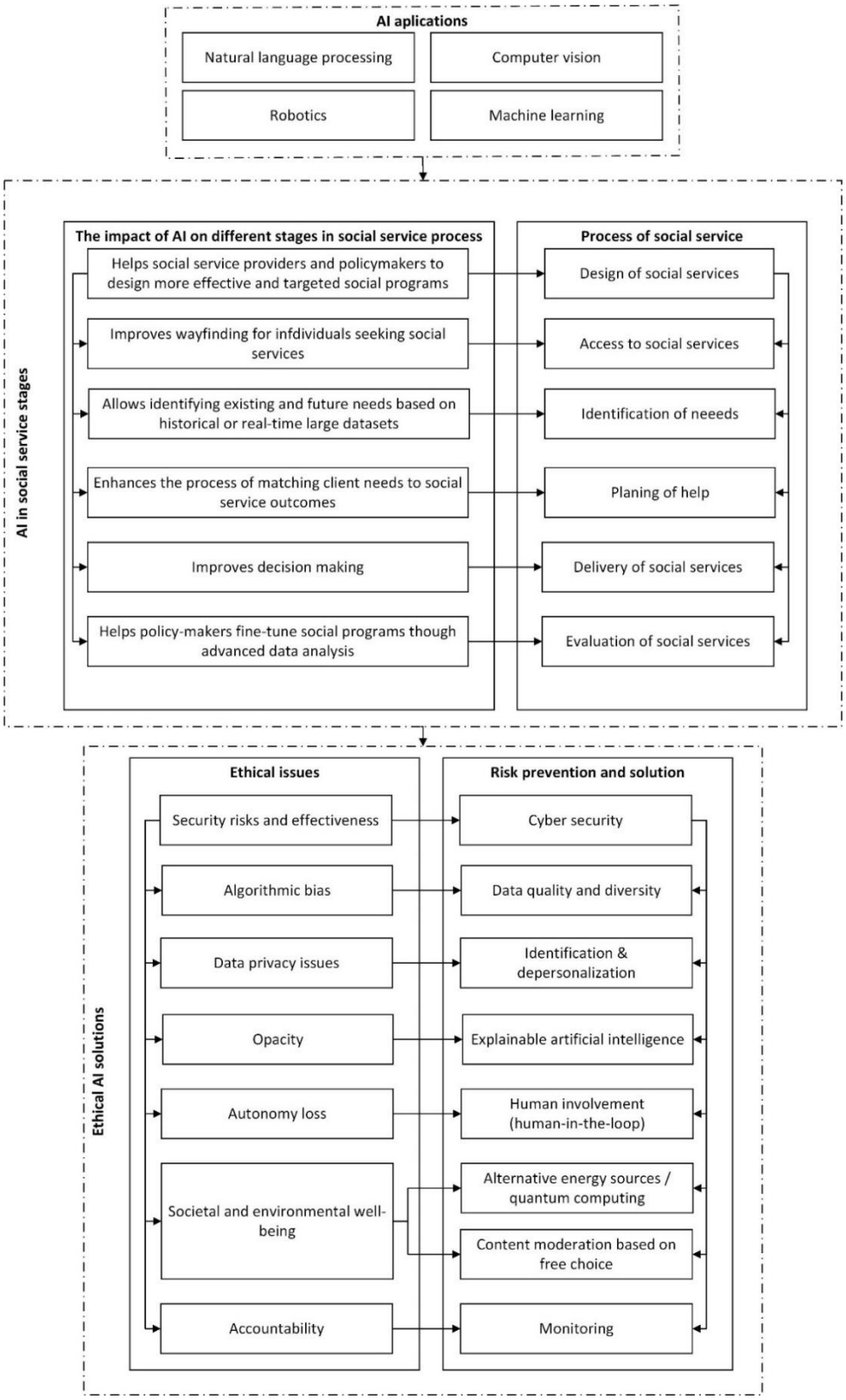

Figure 4. Conceptual Framework for Ethical AI-based social service engineering
Source: own elaboration.

Finally, opportunities for preventing or addressing these challenges are highlighted, including strategies like a) cyber security – the practice of protecting computer systems, networks, and data from digital attacks, theft, or damage. It encompasses technologies, processes, and controls designed to defend against cyber threats (Hendrycks et al., 2022; Zhou et al., 2023); b) data quality and diversity – refers to the accuracy, completeness, reliability, and relevance of data in AI systems. Diversity in data ensures that AI algorithms are trained on a wide range of datasets to avoid bias and to perform effectively across different scenarios (Suresh & Guttag, 2021; Torralba & Efros, 2011); c) identification & depersonalization – identification in AI involves recognizing and distinguishing individual entities, whereas depersonalization is the process of removing personally identifiable information from data sets, ensuring privacy and anonymity (Lison et al., 2021; Patsakis & Lykousas, 2023); d) explainable artificial intelligence (XAI) – the field of AI that focuses on the creation of AI systems whose actions can be easily understood by humans. XAI aims to make AI decisions transparent, understandable, and interpretable (Barredo Arrieta et al., 2020; Molnar, 2019); e) human involvement (human-in-the-loop) – a system design paradigm that incorporates human judgment into AI systems, allowing humans to provide feedback, make decisions, or adjust outputs in real-time, ensuring that the AI remains aligned with human values and goals (Mullainathan & Obermeyer, 2017; Shevlane et al., 2023); f) alternative energy sources / quantum computing – this refers to the exploration and use of renewable energy sources, like solar or wind power, to run AI computations, and the application of quantum computing to dramatically increase computational power for certain types of problems, potentially improving AI efficiency and capabilities (Ajagekar & You, 2019; Jaschke & Montangero, 2023; Li et al., 2022; McDonald et al., 2022); g) content moderation based on free choice – n AI allows users to customize content filters to their ethical preferences and cultural norms, offering a personalized approach to blocking or allowing content instead of a universal moderation policy. This model promotes a balance between preventing harm and upholding diverse expressions (Bubeck et al., 2023; Open AI, 2023); h) monitoring – the continuous observation, checking, and tracking of AI systems' performance and activities. Monitoring aims to ensure that AI systems function as intended, remain secure, and any issues are promptly identified and addressed (Sculley et al., 2015).

This Conceptual Framework presents a general model for evaluating the impact of diverse AI applications across the social service process, pinpointing key ethical concerns that may emerge. It underscores the necessity of incorporating ethical considerations at the initial stages of AI development. While the framework is universally applicable, it is especially pertinent in examining the prospective integration of AI within Lithuania's social service sector. This case study approach offers valuable perspectives on the effective and ethical implementation of AI. The framework's alignment with standard social service stages ensures its adaptability for investigating the potential deployment of various AI systems in delivering a range of social services to different client groups in varied institutional settings.

CONCLUSION

Human-centered ethical AI development, guided by international efforts, is a priority yet remains an unachieved goal. The leadership of the EU and USA in AI regulation and the AI industry raises expectations for AI's positive social and economic impact. Research shows that AI's technological breakthrough has successfully applied many AI systems in the field of social welfare. However, in Lithuania, strategic social welfare documents lack specific AI integration plans, highlighting the need for legal regulation, funding, and AI strategy development in social services. The Conceptual Framework proposed in this research provides critical insights for the ethical and effective integration of AI into social service sectors, exemplified through the Lithuanian context but designed for universal applicability. Moreover, the universality of social service stages enhances the framework's utility in applying it to diverse AI applications across various client groups and institutional settings.

The proposed framework delineates a methodical integration of AI into social service delivery, highlighting the transformative potential of technologies such as natural language processing, computer vision, robotics, and machine learning. These technologies, when judiciously applied, can streamline the processing of information and enhance client interactions through automation, offering personalized and efficient service provision. Delving into the social service stages, the framework elucidates how AI can critically reformulate the lifecycle of services—from the design of predictive and adaptive social programs, to improving accessibility with intelligent guidance systems, from harnessing big data for need identification and planning, to the alignment of services with client requirements through smart matching algorithms. At the ethical forefront, the framework insists on the rectification of algorithmic biases by curating diverse datasets for training AI systems. It underlines the imperative of safeguarding data privacy, asserting transparency in the workings of AI, and upholding accountability for AI-induced decisions, with an insistent recommendation for human oversight in key decision-making junctures. Risk prevention is addressed through a set of practical guidelines that encompass stringent cybersecurity protocols, the enhancement of data quality, the promotion of AI explainability, and the integration of humans in the loop, particularly in sensitive decision-making processes. For pragmatic enactment, the framework advocates regular auditing, stakeholder education, and community engagement, reinforcing the role of AI in social services as a collaborative nexus of technology, ethics, and human welfare.

In the conclusive synthesis of our research, it is imperative to translate the theoretical underpinnings of our framework into actionable recommendations. These practical suggestions are intended to guide key stakeholders—policymakers, service providers, clients, and the community at large—on the responsible integration of AI into the social services sector. For policymakers, the framework recommends the development of robust AI governance models that not only foster innovation but also prioritize ethical considerations. This includes creating policies that encourage the ethical collection and use of data, implementing standards for transparency, and promoting equitable access to AI-enhanced services. Service providers are encouraged to adopt AI solutions that are congruent with the core values of social work, ensuring that such technologies are used to augment, rather than replace, human judgment and empathy. Service providers should also focus on training their staff to work effectively with AI tools and to understand their capabilities and limitations. Clients, as end-users of social services, should be educated on how AI may impact their access to and the quality of services

provided. They should be empowered with the knowledge to navigate AI-driven services safely and with awareness of their rights, particularly regarding data privacy. For other involved parties, including technologists, ethicists, and social welfare experts, the framework advises a collaborative approach to ensure that AI technologies are developed and implemented in a socially responsible manner. This includes continuous dialogue on the ethical implications of AI, shared responsibility for risk mitigation, and a commitment to the ongoing evaluation of AI's impact on social welfare.

REFERENCES

Normative documents

European Commission. (2018, April 20). Artificial Intelligence for Europe. https://eur-lex.europa.eu/legal-conhttps://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52018DC0237

European Commission. (2019, April 8). Ethics Guidelines for Trustworthy AI. High-Level Expert Group on Artificial Intelligence. https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html#:~:text=The AI HLEG presented, and published in April 2019

European Commission. (2020, February 19). White Paper on Artificial Intelligence a European approach to excellence and trust. https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-next-commission_en.pdf.

European Commission. (2021a, March 9). 2030 Digital Compass: the European way for the Digital Decade. https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A52021DC0118

European Commission. (2021b, April 21). Artificial Intelligence Act. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206

European Commission. (2021c, April 21). Fostering a European approach to Artificial Intelligence. https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=COM%3A2021%3A205%3AFIN

European Commission. (2022). The Digital Economy and Society Index. https://digital-strategy.ec.europa.eu/en/policies/desi

United Nations. (2015). Transforming our world: the 2030 Agenda for Sustainable Development. https://www.refworld.org/docid/57b6e3e44.html

United Nations. (2021). Recommendation on the Ethics of Artificial Intelligence (Issue November). https://doi.org/10.7551/mitpress/14102.003.0010

United Nations. (2022). United Nations Activities on Artificial Intelligence (AI). https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52018DC0237

United Nations. (2023). Building fertile ground for data science in Uganda. United Nations Global Pulse. https://medium.com/un-global-pulse/building-fertile-ground-for-data-science-in-uganda-a950dfd3ca0b

White House. (2022). Blueprint for an AI Bill of Rights. Making automated systems work for the American peopleTEMS WORK FOR THE AMERICAN PEOPLE. https://www.whitehouse.gov/ostp/ai-bill-of-rights

Ministry of Economy and Innovations. (2019). Lithuanian artificial intelligence strategy: a vision for the future.

Ministry of Social Security and Labour. (2023). Strategic Plan.

Scientific sources

Abelson, B., Varshney, K. R., & Sun, J. (2014). Targeting direct cash transfers to the extremely poor. Proceedings of the ACM SIGKDD *International Conference on Knowledge Discovery and Data Mining*, 1563–1572. https://doi.org/10.1145/2623330.2623335

Akula, R., & Garibay, I. (2021). Ethical AI for Social Good. Lecture Notes in Computer Science (*Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics*), 13095 LNCS, 369–380. https://doi.org/10.1007/978-3-030-90963-5_28/COVER

Bucholc, M., James, C., Al Khleifat, A., Badhwar, A., Clarke, N., Dehsarvi, A., Madan, C. R., Marzi, S. J., Shand, C., Schilder, B. M., Tamburin, S., Tantiangco, H. M., Llewellyn, D. J., & Ranson, J. M. (2023). Artificial intelligence for dementia research methods optimization *The Deep Dementia Phenotyping* (DEMON) Network 1 Ilianna Lourida 14. https://doi.org/10.1002/alz.13441

Burgard, W. (2023). Artificial Intelligence: Key Technologies and Opportunities. In K. Frankish & W. M. Ramsey (Eds.), The Cambridge handbook of artificial intelligence (pp. 11–19). Cambridge University Press.

Chignell, M., Matulis, H., & Nejati, B. (2020). Motivating Physical Exercise in the Elderly with Mixed Reality Experiences. In N. Streitz & S. Konomi (Eds.), 8th International Conference Distributed, Ambient and Pervasive Interactions (DAPI) 2020: Vol. 12203 LNCS (pp. 505–519). Springer. https://doi.org/10.1007/978-3-030-50344-4_36/TABLES/3

Chui, M., Harryson, M., Manyika, J., Roberts, R., Chung, R., van Heteren, A., & Nel, P. (2018). Notes from the AI frontier. Applying AI for Social Good.

Chui, M., Issler, M., Roberts, R., & Yee, L. (2023). McKinsey Technology Trends Outlook 2023. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-top-trends-in-tech#tech-trends-2023

Cornebise, J., Worrall, D., Farfour, M., & Marin, M. (2018). Witnessing atrocities: quantifying villages destruction in Darfur with crowdsourcing and transfer learning. AI for Social Good NeurIPS2018 Workshop, NeurIPS '18,

Darling, E. (2022). Digital Technology Supercluster Announces Investment to Improve the Accessibility of Social Services. https://www.digitalsupercluster.ca/digital-technology-supercluster-announces-investment-to-improve-the-accessibility-of-social-services/

Engin, Z., & Treleaven, P. (2019). Algorithmic Government: Automating Public Services and Supporting Civil Servants in using Data Science Technologies. Advance Access Publication On, 11. https://doi.org/10.1093/comjnl/bxy082

Floridi, L., Cowls, J., King, T. C., & Taddeo, M. (2021). How to Design AI for Social Good: Seven Essential Factors. Philosophical Studies Series, 144, 125–151. https://doi.org/10.1007/978-3-030-81907-1_9/COVER

Fox, M. S., Gajderowicz, B., Rosu, D., Turner, A., & Lyu, D. (2022). An Ontological Approach to Analysing Social Service Provisioning. ISC2 2022 - 8th IEEE International Smart Cities Conference. https://doi.org/10.1109/ISC255366.2022.9922132

Gajderowicz, Bart, & Fox, M. S. (2017). Requirements for Emulating Homeless Client Behaviour. Conference: AAAI Workshop on Operations Research and Artificial Intelligence for Social Good. https://www.researchgate.net/publication/313556478

Gajderowicz, Bart, Fox, M. S., & Grüninger, M. (2014). Requirements for an Ontological Foundation for Modelling Social Service Chains. Proceedings of the 2014 Industrial and Systems Engineering Research Conference Y. Guan and H. Liao, Eds.

Gajderowicz, Bartosz. (2019). Artificial Intelligence Planning Techniques for Emulating Agents with Application in Social Services. University of Toronto (Canada).

Gillingham, P. (2021). Algorithmically Based Decision Support Tools: Skeptical Thinking about the Inclusion of Previous Involvement. *Practice*, 33(1), 37–50. https://doi.org/10.1080/09503153.2020.1749584

Holzmeyer, C. (2021). Beyond 'AI for Social Good' (AI4SG): social transformations—not tech-fixes—for health equity. *Interdisciplinary Science Reviews*, 46(1–2), 94–125. https://doi.org/10.1080/03080188.2020.1840221

Janušonytė, E. (2021). Odos vėžio diagnostika ir dirbtinis intelektas. Sveikatos Mokslai, 31(3), 175–180. https://doi.org/10.35988/sm-hs.2021.103

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, 353(6301), 790–794. https://doi.org/10.1126/SCIENCE.AAF7894/SUPPL_FILE/JEAN.SM.PDF

Juškevičiūtė-Vilienė, A. (2020). Dirbtinis intelektas ir konstitucinė teisė į teisingumą. *Acta Universitatis Lodziensis. Folia Iuridica*, 93, 117–136.

Kim, E., Jang, G. Y., & Kim, S. H. (2022). How to Apply Artificial Intelligence for Social Innovations. *Applied Artificial Intelligence*, 36(1). https://doi.org/10.1080/08839514.2022.2031819

Oravec, J. A. (2019). Artificial Intelligence, Automation, and Social Welfare: Some Ethical and Historical Perspectives on Technological Overstatement and Hyperbole. *Ethics and Social Welfare*, 13(1), 18–32. https://doi.org/10.1080/17496535.2018.1512142

Patel, K., & Parmar, B. (2022). Assistive device using computer vision and image processing for visually impaired; review and current status. *Disability and Rehabilitation: Assistive Technology*, 17(3), 290–297. https://doi.org/10.1080/17483107.2020.1786731

Reamer, F. G. (2023). Artificial Intelligence in Social Work: Emerging Ethical Issues. *International Journal of Social Work Values and Ethics*, 20(2), 52–71. https://doi.org/10.55521/10-020-205

Robinson, C., & Dilkina, B. (2018). A Machine Learning Approach to Modeling Human Migration. COMPASS '18: Proceedings of the 1st *ACM SIGCAS Conference on Computing and Sustainable Societies*, 1–9. https://doi.org/10.1145/3209811.3209868

Rosu, D., Aleman, D. M., Beck, J. C., Chignell, M., Consens, M., Fox, M. S., Grüninger, M., Grüninger, G., Liu, C., Ru, Y., & Sanner, S. (2017). A Virtual Marketplace for Goods and Services for People with Social Needs.

Saunders, R. P. (2016). Implementation Monitoring and Process Evaluation.

Shah, O. R., Willoughby, L., & Bowersox, N. (2021). Issues in Information Systems Tackling homelessness through AI powered social innovations: A novel and ground-breaking assessment of criminal victimization of homeless populations in los angeles employing predictive analytics and machine learning models such as ARIMA and LSTM. 22(3), 264–277. https://doi.org/10.48009/3_iis_2021_283-297

Shi, Z. R., Wang, C., & Fang, F. (2020). Artiicial Intelligence for Social Good: A Survey. Preprint: ArXiv:2001.01818. https://aaai.org/Symposia/Spring/sss17symposia.php

Skardžiūtė, J. (2018). Dirbtinio intelekto įtaka konkurencijos teisei: konkurencijos teisės normų aiškinimo ir taikymo problemos [Vilniaus universitetas]. https://epublications.vu.lt/object/elaba:29809699/

Tomašev, N., Cornebise, J., Hutter, F., Mohamed, S., Picciariello, A., Connelly, B., Belgrave, D. C. M., Ezer, D., Haert, F. C. van der, Mugisha, F., Abila, G., Arai, H., Almiraat, H., Proskurnia, J., Snyder, K., Otake-Matsuura, M., Othman, M., Glasmachers, T., Wever, W. de, … Clopath, C. (2020). AI for social good: unlocking the opportunity for positive impact. *Nature Communications* 2020 11:1, 11(1), 1–6. https://doi.org/10.1038/s41467-020-15871-z

Trilupaitytė, S. (2022). Vizualioji kontrolė šiandienos visuomenėse: veidų ir emocijų (ne)atpažinimas. *Politologija*, 2(106), 131–164.

Venclovaitė, D., Stramkauskaitė, A., Kuzmienė, L. (2022). Dirbtinis intelektas oftalmologijoje. *Artificial intelligence in ophthalmology*. https://doi.org/10.37499/LBPG.942

Vendeville, G. (2022). Using AI to optimize social services: Professor Mark Fox among U of T researchers to team up with industry and government. https://www.mie.utoronto.ca/using-ai-to-optimize-social-services-professor-mark-fox-among-u-of-t-researchers-to-team-up-with-industry-and-government/

Vidauskytė, L. (2021). Dirbtinis intelektas: visuomenės infantilizacija ir bejėgiškumas. *Logos* (Vilnius), 109, 71–77. https://doi.org/10.24101/LOGOS.2021.77

Wellman, M. P., & Rajan, U. (2017). Ethical Issues for Autonomous Trading Agents. *Minds and Machines*, 27. https://doi.org/10.1007/s11023-017-9419-4

Yadav, A., Chan, H., Jiang, A., Rice, E., Kamar, E., Grosz, B., & Tambe, M. (2016). POMDPs for assisting homeless shelters – Computational and deployment challenges. *International Conference on Autonomous Agents and Multiagent Systems*, 10003 LNAI, 67–87. https://doi.org/10.1007/978-3-319-46840-2_5/FIGURES/14

Yeung, K. (2018). Algorithmic regulation: A critical interrogation. *Regulation & Governance*, 10, 505–523. https://doi.org/10.1111/rego.12158

Yu, K. H., Beam, A. L., & Kohane, I. S. (2018). Artificial intelligence in healthcare. *Nature Biomedical Engineering* 2018 2:10, 2(10), 719–731. https://doi.org/10.1038/s41551-018-0305-z

Zakaras, K. (2022). Dirbtinio intelekto naudojimo teisėsaugos institucijų veikloje iššūkiai žmogaus teisių požiūriu. Mykolo Romerio universitetas.

Secondary sources

Ajagekar, A., & You, F. (2019). Quantum computing for energy systems optimization: Challenges and opportunities. *Energy*, 179, 76–89. https://doi.org/10.1016/j.energy.2019.04.186

Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82–115. https://doi.org/10.1016/J.INFFUS.2019.12.012

Bencsik, A. (2021). The sixth generation of knowledge management – the headway of artificial intelligence. *Journal of International Studies, 14*(2), 84-101. doi:10.14254/2071-8330.2021/14-2/6

Bilan, Y., Oliinyk, O., Mishchuk, H., & Skare, M. (2023). Impact of information and communications technology on the development and use of knowledge. *Technological Forecasting and Social Change*, *191*, 122519. DOI: 10.1016/j.techfore.2023.122519

Bojarski, M., Testa, D. Del, Dworakowski, D., Firner, B., Flepp, B., Goyal, P., Jackel, L. D., Monfort, M., Muller, U., Zhang, J., Zhang, X., Zhao, J., & Zieba, K. (2016). End to End Learning for Self-Driving Cars.

Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Lee, Y. T., Li, Y., Lundberg, S., Nori, H., Palangi, H., Ribeiro, M. T., & Zhang, Y. (2023). Sparks of Artificial General Intelligence: Early experiments with GPT-4. Microsoft Research. https://arxiv.org/abs/2303.12712v5

Carton, S., Helsby, J., Joseph, K., Mahmud, A., Park, Y., Walsh, J., Cody, C., Patterson, C. P. T. E., Haynes, L., & Ghani, R. (2016). Identifying police officers at risk of adverse events. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 13-17-August-2016, 67–76. https://doi.org/10.1145/2939672.2939698

Fang, F., Nguyen, T. H., Pickles, R., Lam, W. Y., Clements, G. R., An, B., Singh, A., Tambe, M., & Lemieux, A. (2016). Deploying PAWS: Field Optimization of the Protection Assistant for Wildlife Security. Proceedings of the AAAI Conference on Artificial Intelligence, 30(2), 3966–3973. https://doi.org/10.1609/AAAI.V30I2.19070

Hendrycks, D., Google, N. C., Schulman, J., & Steinhardt, J. (2022). Unsolved Problems in ML Safety. Preprint. https://arxiv.org/abs/2109.13916

Jankovic, J., & Bernier, A. (2023). Research shows decision-making AI could be made more accurate when judging humans. https://www.utoronto.ca/news/research-shows-decision-making-ai-could-be-made-more-accurate-when-judging-humans

Jaschke, D., & Montangero, S. (2023). Is quantum computing green? An estimate for an energy-efficiency quantum advantage. *Quantum Sci. Technol*, 8, 25001. https://doi.org/10.1088/2058-9565/acae3e

Kersan-Škabić, I., & Vukašina, M. (2023). Contribution of ESIFs to the digital society development in the EU. *Journal of International Studies, 16*(2), 195-210. doi:10.14254/2071-8330.2023/16-2/13

Kolková, A., & Ključnikov, A. (2022). Demand forecasting: AI-based, statistical and hybrid models vs practice-based models - the case of SMEs and large enterprises. *Economics and Sociology, 15*(4), 39-62. doi:10.14254/2071- 789X.2022/15-4/2

Lakkaraju, H., Aguiar, E., Shan, C., Miller, D., Bhanpuri, N., Ghani, R., & Addison, K. L. (2015). A machine learning framework to identify students at risk of adverse academic outcomes. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2015-August, 1909–1918. https://doi.org/10.1145/2783258.2788620

Letkovsky, S., Jencova, S., Vasanicova, P., Gavura, S., & Bacik, R. (2023). Predicting bankruptcy using artificial intelligence: The case of the engineering industry. *Economics and Sociology, 16*(4), 178-190. doi:10.14254/2071- 789X.2023/16-4/8

Li, B., Samsi, S., Gadepally, V., & Tiwari, D. (2022). Clover: Toward Sustainable AI with Carbon-Aware Machine Learning Inference Service. Findings of the Association for Computational Linguistics: NAACL , 1962–1970. https://doi.org/10.1145/3581784.3607034

Lison, P., Pilán, I., Sánchez, D., Batet, M., & Øvrelid, L. (2021). Anonymisation Models for Text Data: State of the art, Challenges and Future Directions. ACL-IJCNLP 2021 - 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, Proceedings of the Conference, 4188–4203. https://doi.org/10.18653/V1/2021.ACL-LONG.323

McDonald, J., Li, B., Frey, N., Tiwari, D., Gadepally, V., & Samsi, S. (2022). Great Power, Great Responsibility: Recommendations for Reducing Energy for Training Language Models. Findings of the Association for Computational Linguistics: NAACL 2022 - Findings, 1962–1970. https://doi.org/10.18653/v1/2022.findings-naacl.151

Molnar, C. (2019). Interpretable Machine Learning A Guide for Making Black Box Models Explainable. Leanpub. http://leanpub.com/interpretable-machine-learning

Mullainathan, S., & Obermeyer, Z. (2017). Does Machine Learning Automate Moral Hazard and Error? American Economic Review, 107(5), 476–480. https://doi.org/10.1257/AER.P20171084

Ohlenburg, T. (2020). AI in social protection -exploring opportunities and mitigaiting risks. Deutsche Gesellschaft für Internationale Zusammenarbeit. https://socialprotection.org/discover/publications/ai-social-protection

Open AI. (2023). GPT-4 Technical Report. In Submitted.

Patsakis, C., & Lykousas, N. (2023). Man vs the machine: The Struggle for Effective Text Anonymisation in the Age of Large Language Models. Preprint. https://arxiv.org/abs/2303.12429v1

Roshchyk, I., Oliinyk, O., Mishchuk, H., & Bilan, Y. (2022). IT Products, E-Commerce, and Growth: Analysis of Links in Emerging Market. *Transformations in Business & Economics, 21*(1), 209-227.

Ross, C., & Swetlitz, I. (2017). IBM pitched Watson as a revolution in cancer care. It's nowhere close. moz-extension://b3f4dfef-8263-4daa-9eff-df6d2dfbba2b/enhanced-reader.html?openApp&pdf=https%3A%2F%2Fwww.preventcancer.org%2Fwp-content%2Fuploads%2F2018%2F06%2FIBM_pitched_Watson_as_a_revolution_in_cancer_care.pdf

Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., Chaudhary, V., Young, M., Crespo, J.-F., & Dennison, D. (2015). Hidden Technical Debt in Machine Learning Systems. *In Advances in Neural Information Processing Systems* (Vol. 28).

Shevlane, T., Farquhar, S., Garfinkel, B., Phuong, M., Whittlestone, J., Leung, J., Kokotajlo, D., Marchal, N., Anderljung, M., Kolt, N., Ho, L., Siddarth, D., Avin, S., Hawkins, W., Kim, B., Gabriel, I., Bolina, V., Clark, J., Bengio, Y., … Dafoe, A. (2023). Model evaluation for extreme risks. Preprint. https://arxiv.org/abs/2305.15324v2

Strickland, E. (2019). IBM Watson, heal thyself: How IBM overpromised and underdelivered on AI health care. *IEEE Spectrum*, 56(4), 24–31. https://doi.org/10.1109/MSPEC.2019.8678513

Suresh, H., & Guttag, J. (2021, October 5). A Framework for Understanding Sources of Harm throughout the Machine Learning Life Cycle. ACM International Conference Proceeding Series. https://doi.org/10.1145/3465416.3483305

Torralba, A., & Efros, A. (2011, May). Unbiased look at dataset bias. In CVPR 2011 (Pp. 1521-1528). IEEE. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5995347

United Nations. (2015). Transforming our world : the 2030 Agenda for Sustainable Development. https://www.refworld.org/docid/57b6e3e44.html

United Nations. (2021). Recommendation on the Ethics of Artificial Intelligence (Issue November). https://doi.org/10.7551/mitpress/14102.003.0010

United Nations. (2022). United Nations Activities on Artificial Intelligence (AI). https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52018DC0237

United Nations. (2023). Building fertile ground for data science in Uganda. United Nations Global Pulse. https://medium.com/un-global-pulse/building-fertile-ground-for-data-science-in-uganda-a950dfd3ca0b

White House. (2022). Blueprint for an AI Bill of Rights. Making automated systems work for the American peopleTEMS WORK FOR THE AMERICAN PEOPLE. https://www.whitehouse.gov/ostp/ai-bill-of-rights

Zhou, B., Li, B., Qi, P., Liu, B., Di, S., Liu, J., Pei, J., & Yi, J. (2023). Trustworthy AI: From Principles to Practices; Trustworthy AI: From Principles to Practices. *ACM Computing Surveys*, 55(9), 177. https://doi.org/10.1145/3555803

Authors' Note

All correspondence should be addressed to
Miroslavas Seniutis
Institute of Sociology and Social Work, Vilnius University, Lithuania
miroslavas.seniutis@fsf.vu.lt

Human Technology
ISSN 1795-6889
https://ht.csr-pub.eu