Risks of Artificial Intelligence-Based Decision Support and Decision-Making Systems in Executive-Level Decision-Making in Companies – A Literature Review[1]

Gergely Kappel *

https://doi.org/10.15170/PJIEL.2024.2.3

Contents

* PhD student, University of Pécs, Faculty of Law. ORCID ID: https://orcid.org/0009-0004-2442-8850

[1] SUPPORTED BY THE ÚNKP-23-3-I. NEW NATIONAL EXCELLENCE PROGRAM OF THE MINISTRY FOR CULTURE AND INNOVATION FROM THE SOURCE OF THE NATIONAL RESEARCH, DEVELOPMENT AND INNOVATION FUND.

The present paper is based on the manuscript in Hungarian entitled: 'Mesterséges intelligencia alapú döntéstámogató és döntéshozó rendszerek kockázatai a vállalatok vezetői szintű döntéshozatalában – Szakirodalmi áttekintés'. It is expected to be published in the Hungarian law journal Pro Futuro (online ISSN 2063-2754, printed ISSN: 2063-1987).

Abstract

The study examines the risks associated with artificial intelligence (AI) based decision-making and decision-support systems in the decision-making processes of company executives, as well as small and medium-sized enterprises. Due to global trends and digital advancements, company management increasingly faces complex decisions, which AI-based decision-making and decision-support systems may well be suited to support. However, this carries several risks, and the study aims to identify the legal, ethical, and business risks associated with the use of such AI systems, with a particular focus on the decisions made by company executives. The analysis is based on a literature review, which will ultimately be compared with survey responses found in the *AI Index Reports* published annually by Stanford University.

Keywords: artificial intelligence, data protection, automated decision-making, risk management, literature review

I. Introduction

Due to the interconnected nature of the global economy and prevailing global trends, corporate leadership is increasingly required to make complex decisions that align with sustainable development and financial stability. From the perspective of corporate operations, this complex environment presents a range of problems and challenges, which are further exacerbated by the often stringent demands of shareholders and other stakeholders. These expectations are frequently accompanied by significant pressure, leading to potential financial losses for the company—and ultimately for the shareholders—if poor decisions are made. On average, inadequate or suboptimal decision-making at the management level costs companies at least 3% of their profits. The opportunities created by digital advancements, such as artificial intelligence (AI) -based decision-making and decision-support systems, can greatly facilitate the smooth resolution of obstacles that require rapid and complex decision-making. It is no surprise, then, that a growing number of companies are striving to integrate various AI-based systems into their daily operations.[2]

The *IBM Institute for Business Value*, in collaboration with *Oxford Economics*, conducted a study examining the decision-making processes of chief executive officers (CEOs) in relation to one of today's most highly praised innovations: AI. As part of the research, more than 3,000 CEOs were surveyed across over 30 countries and 24 industries. The study revealed, among other findings, that 43% of CEOs utilize some form of generative AI in strategic decision-making. Furthermore, 66% of board members and 64% of investors and creditors are encouraging CEOs to accelerate the corporate implementation of AI.[3] The *2024 AI Index Report* published by *Stanford University* reveals that investor interest in AI has been a key trend for several years, with global corporate investments reaching USD 189.2 billion in 2023. Of this amount, USD 95.99 billion came from private investments in the global market. The investment landscape is predominantly led by the United States, accounting for USD 67.22 billion, followed by China with USD 7.76 billion, the United Kingdom with USD 3.78 billion, and Germany in fourth place with USD 1.91 billion. Despite a decline from the peak investment levels seen in 2021,[4] the report forecasts an exponential growth in

[2] Mark Purdy and A. Mark Williams, 'How AI Can Help Leaders Make Better Decisions Under Pressure' (*Harvard Business Review*, 26 October 2023) <https://hbr.org/2023/10/how-ai-can-help-leaders-make-better-decisions-under-pressure> accessed 31 May 2024.

[3] IBM Institute for Business Value, 'CEO decision-making in the age of AI' (*IBM*, 2023) <https://www.ibm.com/downloads/cas/1V2XKXYJ> accessed 31 May 2024.

[4] According to analyses, the global private investment figure was close to USD 130 billion in 2021. See: Nestor Maslej, Loredana Fattorini, Raymond Perrault, Vanessa Parli, Anka Reuel, Erik Brynjolfsson, John Etchemendy, Katrina Ligett, Terah Lyons, James Manyika, Juan Carlos Niebles, Yoav Shoham, Russell Wald, and Jack Clark, 'The AI Index 2024 Annual Report'

AI-based software and solutions in the coming years. Notably, the investments in generative AI have surged dramatically. While in 2022, global private investments in generative AI were around USD 3 billion, they skyrocketed to USD 25.23 billion in 2023. This figure represents 26% of all private AI investments, underscoring the substantial and growing interest in this particular area of AI.[5]

With the substantial increase in investment in the field of generative AI, it is no surprise that corporate usage is also on the rise. Naturally, it is not only the number of generative AI solutions that is expanding in the realm of business applications, but it is clear that generative AI will play a dominant role in the coming business years. As the usage increases, the range of risk scenarios associated with AI-based solutions is also widening. In my view, a significant portion of these risk scenarios emerges from the application of AI-based decision-making and decision-support systems. Nevertheless, I find it necessary to accurately define the risk scenarios (particularly the most frequently occurring ones), as this can help companies better prepare for mitigating these risks. There are several ways to identify risks, but I would highlight two main approaches. Firstly, risks can be identified by examining the academic literature and expert analyses, from which we can infer the most commonly addressed risks. Secondly, risks can be identified through surveys of companies, including small and medium-sized enterprises, based on their experiences. This study, as a literature review, follows the first identification method; however, I also compare the risks frequently mentioned in the literature with those identified through surveys reported annually in the AI Index Report.

The structure of this study is as follows: The first section outlines the research process, detailing (i) data sources and research strategy, and (ii) exclusion criteria. The second section addresses the identification of risks, covering (i) the identification of risks based on the reviewed literature; (ii) the evolution of risk prioritization during the examined period; and (iii) a comparison of the results with responses provided by companies. The third section presents proposed solutions for the identified risks, as found in the reviewed academic works. The fourth section includes the conclusions drawn from the research, and finally, the fifth chapter discusses the limitations of the study.

II. Research objectives and methodology

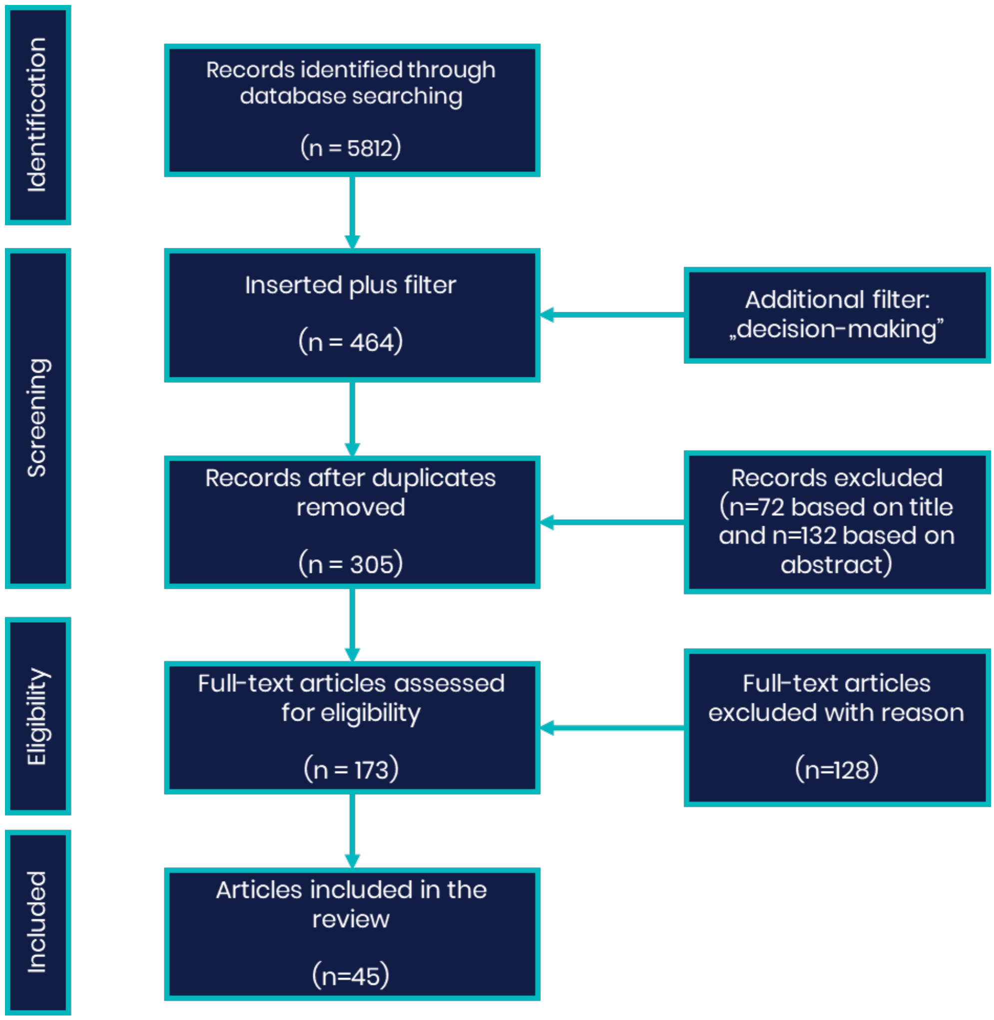
The objective of this research is to identify the legal, ethical, and business risks arising from the use of AI-based decision-making and decision-support systems,

(2024) Stanford University Human-Centered Artifical Intelligence <https://aiindex.stanford.edu/wp-content/uploads/2024/04/HAI_2024_AI-Index-Report.pdf> accessed 31 May 2024.

[5] ibid.

with a focus on decision-making at the executive level of companies, and somewhat of small and medium-sized enterprises. Given that the reviewed literature applies different legal frameworks and specific definitions, the current study refrains from providing explicit definitions of *company* and *small and medium-sized enterprise*. Therefore, this study does not follow the definitions set out in Hungarian law or those found in European Union legal frameworks and qualification criteria. For the purposes of this research, I adopt an abstract approach when referring to companies and small and medium-sized enterprises. Furthermore, my analysis is not based on Hungarian law but is conducted at an abstraction level that, in my view, best supports the objectives of the research. Nevertheless, I believe that proper identification and management of risks can prevent reputational and financial damage to businesses and can also allow for clearer delineation of responsibilities within various legal relationships. Essentially, the central question of this research is what are the most common risks identified in the literature concerning AI-based decision-making and decision-support systems. During the research, I aimed to uncover the risks and proposed solutions identified in the literature and contrast these with the risks perceived by the surveyed organizations. As a result, this study presents opinions, rather than specific recommendations.

The research process consisted of five steps (see Figure 1). In the first step, I created keywords, using OpenAI's ChatGPT service to generate synonyms for each keyword. In the second step, based on the generated keywords, I conducted a search in the Scopus and Web of Science databases, as these were the most suitable for applying the filtering criteria and performing complex keyword searches, which are described later. Initially, I intended to examine the topic within a broader context; however, due to the large volume of results, I applied additional filters and narrowed the scope of the investigation. In the third step, I eliminated duplicate entries. Next, I reviewed the abstracts of the remaining articles and further excluded those deemed irrelevant to the research, based on the exclusion criteria outlined later. In the fifth step, I reviewed the remaining articles, narrowing the selection to open-access academic works.



Figure/Table 1 – Steps of the Research Process.

1. Data Sources and Research Strategy

The research underlying this study was developed based on the following criteria and strategy, with the first step being the identification of a comprehensive list of keywords.[6] In the course of the research, I applied an interdisciplinary

[6] The first alternative elements of a list of keywords: Legal; regulation; Law; Contract law; Consumer law; Business law; Corporate law; Employment law; Intellectual property law; Tax law; Environmental law; Regulatory compliance; Privacy law; Cybersecurity law; Commercial litigation; Antitrust law; Banking and finance law; Real estate law; Labor law; Mergers and acquisitions; Bankruptcy law; Intellectual property rights; Data protection; Employment contracts; Trademark law; Copyright law; Product liability; Healthcare law; Estate planning; Antitrust regulations; Immigration law; Legal compliance; Taxation; Insurance law; Securities regulation; Environmental compliance; Corporate governance; Contract negotiations; Regulatory affairs; Trade secrets ; Compliance management; Employment disputes; Data privacy; Licensing agreements; Labor disputes; Financial regulations; Corporate transactions; Business contracts; Intellectual

approach, and therefore, I did not limit the selection of scholarly works based solely on legal studies. The search was conducted on 3 November 2023, in the two databases: *Scopus* and *Web of Science*. For Scopus, the following filtering criteria were applied: (i) final, (ii) in English, (iii) publications, books, book chapters, and reviews/critics. For Web of Science, the criteria were: (i) in English, (ii) publications, books, and reviews/critics. The temporal scope of the study included publications from 2013 to 2023 in both databases.[7] The resulting list of scholarly works consisted of 5,812 entries, which initially served a broader identification purpose. Recognizing that the number of scholarly articles on the broader research topic has grown exponentially in recent years, and this trend continues to this day, I applied an additional keyword, namely *decision-making*, as a filter. This narrowed the list to 464 entries.

The next major phase of the filtering process was narrowing down the found literature based on the titles (the exclusion criteria related to this are outlined below). After filtering based on the titles, the results were narrowed to 392 scholarly works, which was further reduced to 305 after eliminating duplicates. This was followed by additional filtering based on the review of abstracts (the exclusion criteria are listed below), which resulted in a narrowing of the literature review to 173 works. During the complete review of these 173 entries, additional scholarly works were excluded from the study because they were not open access or access was restricted, and I did not have permission to access them. Finally, I conducted a detailed review of 137 scholarly works, of which 45 were found to be relevant to my research objectives.

As the final phase of the research, I filtered the 45 scholarly works to identify the most frequently examined risks by the authors, and then drew my conclusions from this analysis, which are elaborated in Chapter III of this study.

property protection; Tax compliance; Legal risk management; Contract disputes; Legal counsel; Corporate litigation; Corporate governance; Legal compliance; Corporate ethics; Corporate responsibility; Legal advisory; Corporate policies; Legal department; Legal regulations; Corporate legal framework; Legal issues in business; Legal risk assessment.
The second alternative elements of a list of keywords: company; corporation; firm; business; enterprise; organization; concern; institution; agency; establishment; venture; house; conglomerate; consortium; partnership; agency; firm. The third alternative elements of a list of keywords: Artificial Intelligence; AI.

[7] In my view, the proliferation of artificial intelligence-based applications and their interdisciplinary examination have increased exponentially in recent years. Taking this into account, I limited my research to the 10 years preceding the start of my study.

2. *Exclusion criteria*

During the research, the exclusion criteria included any scholarly works that focused on AI-based systems used in the public sector or the judiciary. Additionally, any works that were purely focused on IT, mathematical, physical, or similar considerations and studies were excluded. Another exclusion criterion was the literature on self-driving cars and autonomous systems. I only examined articles, books, etc., that analyzed the risks of AI-based decision-making and decision-support systems in such a way that they were relevant to supporting the decisions of leaders within companies in the corporate, as well as small and medium-sized business sectors. Therefore, among the reviewed works, there are some that concern the healthcare sector but also identify risks that could be significant in decision-making at the corporate and small and medium-sized business levels.

III. Identification of risks

In this chapter, the study presents the risks that have most frequently appeared in the reviewed scholarly works in relation to AI-based tools. The research areas of the examined works show a varied picture, and for the purpose of the review, Appendix 1 of the study contains the research areas of the individual works as well as their publication dates. Below, I describe the various risks and their definitions as found in the literature, as well as the frequency of their occurrence based on the publication years of the scholarly works. The risks were identified based on the fact that the individual works examine them to some extent within the scope of the given work. However, if a risk is only mentioned briefly, it was excluded from the analysis. Among the identified risks are some that are not primarily legal risks but could become legal in nature as a secondary effect. One such example is the risk of inaccuracy, which could cause harm to companies, including small and medium-sized enterprises, thus requiring legal remedies for resolution.

1. *Identification of Risks*

Based on the reviewed scholarly works, the following risks and risk areas were identified most frequently: (i) bias; (ii) discrimination; (iii) fairness; (iv) transparency; (v) explainability; (vi) interpretability; (vii) intelligibility; (viii) reliability; (ix) lack of trust; (x) data protection; (xi) cybersecurity; (xii) access to data; (xiii) inaccuracy; (xiv) robustness; (xv) accountability; (xvi) lack of legal framework. These risks often overlap in the scholarly works, so it is common for one risk to be identified as an element of another risk, as a synonym for it, or as a consequence of a different risk

1.1 *Bias*

The definition of bias is often omitted in the scholarly works, as it is treated as self-evident. However, there are some works that refer to it as the unjust favoritism or prejudice toward or against someone or something.[8] In the case of AI-based decision-making and decision-support systems, there are several forms in which bias can manifest. These include bias arising from the training data,[9] bias resulting from the input data,[10] bias caused by biased variables introduced into the algorithm by the developer,[11] historical and systematic bias,[12] cognitive bias[13] and so on. Bias as a risk appears most frequently in the examined scholarly works, with 36 out of the 45 works addressing or touching upon this issue (see: Figure/Table 2).

1.2. Discrimination

As mentioned above, from the perspective of discrimination, various risks and approaches also appear in scholarly works. On one hand, discrimination as a risk factor can be traced back to discriminatory design flaws that arise during the development of the algorithm, leading to adverse differentiation.[14] Other authors emphasize that, in AI-based systems, a breeding ground for discrimination is the use of discriminatory input data.[15] Some authors argue that the necessary subjective decisions related to machine learning lead to the discriminatory nature of AI-based decision-making and decision-support systems. In this context, they mention aspects such as the collection and handling of training data, the design of the model, and so on.[16]

---

[8] Deepika Chhillar and Ruth V. Aguilera, 'An Eye for Artificial Intelligence: Insights Into the Governance of Artificial Intelligence and Vision for Future Research' (2022) 61 Business & Society 1197.

[9] Daniel Schrönberger, 'Artificial intelligence in healthcare: a critical analysis of the legal and ethical implications' (2019) 27 International Journal of Law and Information Technology 171; Kristin N. Johnson, 'Automating the Risk of Bias' (2019) Tulane Public Law Research Paper No. 19-12, <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3486723> accessed 3 November 2023.

[10] Johnson (n 9).

[11] Tetyana Krupiy, 'A vulnerability analysis: Theorising the impact of artificial intelligence decision-making processes on individuals, society and human diversity from a social justice perspective' (2020) 38 Computer Law & Security Review 1.

[12] Vidushi Marda, 'Artificial intelligence policy in India: a framework for engaging the limits of data-driven decision-making' (2018) 376 Philosophical Transactions of the Royal Society A 1.

[13] Johnson (n 9).

[14] Elif Kiesow Cortez and Nestor Maslej, 'Adjudication of Artificial Intelligence and Automated Decision-Making Cases in Europe and the USA' (2023) 14 European Journal of Risk Regulation 457.

[15] Marvin van Bekkum and Frederik Zuiderveen Borgesius, 'Using sensitive data to prevent discrimination by artificial intelligence: Does the GDPR need a new exception?' (2023) 48 Computer Law & Security Review 1.

[16] Andrew D. Selbst, 'Negligence and AI's Human Users' (2020) UCLA School of Law, Public Law Research Paper No. 20-01 <https://papers.ssrn.com/sol3/papers.cfm?abstract_

1.3. Fairness

The issue of fairness predominantly appears in a broad sense in the reviewed scholarly works. In some instances, the realization of fairness is considered particularly important in order to prevent bias and discriminatory effects in AI-based systems.[17] Some authors examine the human-developed elements that influence the fairness of AI systems. These include the method used to create the AI system (e.g., the learning model, etc.), the algorithm, as well as the physical technological infrastructure.[18] In addition to the above, the issue of fairness is examined from several other aspects in the reviewed scholarly works.

1.4. Transparency

Transparency is the second most examined area in relation to the risks of AI-based decision-making and decision-support systems according to the reviewed works. The concept of transparency is consistently difficult to define, and as such, the approaches in the various works differ. The situation is further complicated by the fact that transparency carries different meanings for different stakeholders. For example, some authors suggest that, from the developer's perspective, transparency involves understanding whether the algorithm is functioning correctly, in order to resolve any emerging errors or contradictions. From the user's perspective, transparency refers to the attribute or condition of knowing what the system is doing, why it is doing it, and what led to a particular decision.[19] Others view transparency as a fundamental element of trust in AI-based systems and approach it from the perspective of informing the stakeholders involved.[20]

1.5. Explainability

A key point of investigation in the context of explainability is the so-called *black-box* effect and the fact that the explainability of a properly organized AI system significantly increases user-level trust. Furthermore, it encourages the identification process of decision-making, allowing us to interpret the causes behind a

id=3350508> accessed 3 November 2023.

17 Lorwai Tan, David Tivey, Helena Kopunic, Wendy Babidge, Sally Langley and Guy Maddern, 'Part 1: Artificial intelligence technology in surgery' (2020) 90 ANZ Journal of Surgery 2409.

18 Charlotte Tschider, 'Beyond the Black Box' (2021) 98 Denver Law Review 683.

19 Heike Felzmann, Eduard Fosch Villaronga, Christoph Lutz and Aurelia Tamò-Larrieux, 'Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns' (2019) 6 Big Data & Society 1.

20 Rozita Dara, Seyed Mehdi Hazrati Fard and Jasmin Kaur, 'Recommendations for ethical and responsible use of artificial intelligence in digital agriculture' (2022) 5 Frontiers in Artifical Intelligence 1.

given output.[21]

1.6. Interpretability

Interpretability as a risk is also closely related to the "black-box effect" associated with AI-based systems. The risk lies in the fact that understanding the functioning of an AI model often encounters difficulties. Essentially, it is tied to the technical realization of the decision-making process. Just like explainability, interpretability is also a fundamental element in building a reliable AI system.[22] Furthermore, by increasing interpretability, we also enhance the transparency of the AI system, which means that these two risks are interconnected.[23]

1.7. Intelligibility

The lack of understandability as a risk also stems from the black-box effect, considering that often the designers of the systems themselves are unable to explain the exact reasons behind the decisions.[24] It is a similar risk category to that arising from the lack of explainability or interpretability, yet the reviewed literature evaluates them separately.

1.8. Reliability

According to some authors, reliability refers to the ability of the AI-based system to indicate when it is likely to fail or become inoperable.[25] Others address it in the context that the decisions made by the algorithm must be reliable, especially when critical decisions need to be made, such as executing stock market transactions.[26] When examining from the perspective of output reliability, the reliability is heavily dependent on the quantity and quality of the data used. A lack of proper data can lead to the artificial intelligence-based decision-making

---

21 Sajid Ali, Tamer Abuhmed, Shaker El-Sappagh, Khan Muhammad, Jose M. Alonso-Moral, Roberto Confalonieri, Riccardo Guidotti, Javier Del Ser, Natalia Díaz-Rodríguez and Francisco Herrera, 'Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence' (2023) 99 Information Fusion 1.

22 Ali et al. (n 21).

23 Maryan Rizinski, Hristijan Peshov, Kostadin Mishev, Lubomir T. Chitkushev, Irena Vodenska and Dimitar Trajanov, 'Ethically Responsible Machine Learning in Fintech' (2016) 10 IEEE Access, 97531.

24 Schönberger (n 9).

25 Dara (n 20).

26 Toan Huu Bui and Van Phuoc Nguyen, 'The Impact of Artificial Intelligence and Digital Economy on Vietnam's Legal System' (2023) 36 International Journal for the Semiotics of Law 969.

and decision-support system providing unreliable output predictions.[27]

1.9. Lack of Trust

Building trust in AI-based systems is a current topic in the literature, as it is seen as a necessary prerequisite for these systems to fulfill their roles. Several authors thus define trust as an attainable goal, with elements such as legality, ethics, as well as technical and social reliability, being key components.[28] There are authors who examine how transparency—especially the level of understanding—affects trust in the system.[29] From the user's perspective, particularly when implementing an AI-based system into executive-level decision-making within a company, a critical factor is how much trust the company's leaders have in the system and how successfully it becomes part of their daily decision-making process. If the AI system's design does not prioritize fostering trust, this can have negative consequences for the company in terms of costs and innovation. It is possible that an AI system is integrated into management based on a corporate decision, but the lack of trust in the system can lead to its actual use being delayed or abandoned. Therefore, a significant risk in the implementation of AI systems is whether a system that either promotes or hinders trust is integrated into executive decision-making.

1.10. Data Protection

In this context, most of the literature examines the framework defined by the 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 (GDPR),[30] particularly in relation to the implementation of personal data protection within AI-based decision-making and decision-support systems. However, other frameworks derived from different data protection regulations also appear in this area. Furthermore, in this risk category, the establishment of inadequate data management practices also emerges as a significant risk.[31] Ac-

[27] Tschider (n 18).

[28] Muzaffer Eroğlu and Meltem Karatepe Kaya, 'Impact of Artificial Intelligence on Corporate Board Diversity Policies and Regulations' (2022) 23 European Business Organization Law Review 541.

[29] Ali et al. (n 21).

[30] van Bekkum (n 15); Michelle Seng Ah Lee, Jennifer Cobbe, Heleen Janssen and Jatinder Singh, 'Defining the scope of AI ADM system risk assessment' in Eleni Kosta, Ronald Leenes and Irene Kamara (eds.) *Research handbook on EU data protection law* (Research Handbooks in European Law 2022).

[31] Michael Hilb, 'Toward artificial governance? The role of artificial intelligence in shaping the future of corporate governance' (2020) 24 Journal of Management and Governance 851.

cording to some works, the risks stemming from data protection arise from the large volume of data used, as most AI-based decision-making and decision-support systems rely on large databases for prediction.[32]

1.11. Cybersecurity

Cybersecurity as a risk is self-evident in these AI-based systems, given that they use large amounts of data for their operation. Naturally—though in sector-specific ways—many of these data sets contain sensitive, sometimes special data or other business-critical information. For example, some authors focus on the healthcare sector, emphasizing the importance of cybersecurity. They stress that the relevance of this risk is exceptionally high, as evidenced by the fact that, in 2021, healthcare cyberattacks affected around 45 million people in the United States.[33]

1.12. Access to Data

Given that AI-based decision-making and decision-support systems require vast amounts of data, the issue of access to this data becomes critical. With the right type and quantity of data, the occurrence of other risks, such as bias or accuracy issues, can be mitigated. In fact, a lack of access to data can easily lead to discriminatory decisions. A notable example is India, where the private sector has limited access to other market or public databases. As a result, individuals from disadvantaged groups, such as those identified by gender, caste, or geographical location, may become victims of discriminatory decisions.[34]

1.13. Inaccuracy

The issue of accuracy and the risks arising from its absence are also prominently featured in the reviewed literature. More than half of the works examined address the impact of this risk and its relationship to other risks. For example, some authors discuss how biased training datasets significantly affect the accuracy of the output of AI systems, and how improper training of machine

[32] Javed Iqbal, Diana Carolina Cortés Jaimes, Pallavi Makineni, Sachin Subramani, Sarah Hemaida, Thanmai Reddy Thugu, Amna Naveed Butt, Jarin Tasnim Sikto, Pareena Kaur, Muhammad Ali Lak, Monisha Augustine, Roheen Shahzad and Mustafa Arain, 'Reimagining Healthcare: Unleashing the Power of Artificial Intelligence in Medicine' (2023) 15 Cureus 1.

[33] Daniele Veritti, Leopoldo Rubinato, Valentina Sarao, Axel De Nardin, Gian Luca Foresti and Paolo Lanzetta, 'Behind the mask: a critical perspective on the ethical, moral, and legal implications of AI in ophthalmology' (2023) 262 Graefe's Archive for Clinical and Experimental Ophthalmology 975.

[34] Marda (n 12).

learning-based models can lead to inaccurate results.[35] Obviously, inaccurate decisions can have negative consequences for an AI-based decision-making or decision-support system, which—like other risks—can result in both financial and reputational damage to a company or a small- and medium-sized enterprise.

1.14. Robustness

The reviewed literature often addresses the robustness of AI systems as a necessary principle, but it generally does not provide a detailed definition or elaborate on its components. However, some authors explain that by robustness, they mean the system's ability to maintain the quality of its performance even under changing conditions.[36] Thus, the lack of stability is closely related to the lack of reliability of the system as well.

1.15. Accountability

The issue of accountability is discussed in some scholarly works together with transparency, considering the latter as a prerequisite for the former's realization.[37] The principle of accountability in AI-based decision-making and decision-support systems suggests that these systems should be able of explaining the output decisions and providing the underlying reasoning behind them.[38] In terms of responsibility, the enforcement of accountability is essential for AI-based systems, as their decisions can directly impact individuals.[39]

1.16. Lack of Legal Framework

The absence or inadequacy of a legal framework is an evident risk in the application of AI-based decision-making and decision-support systems. However, this risk factor is relatively rarely discussed in the reviewed literature. Out of the 45 reviewed works, only 4 addressed this specific risk factor. A proper legal framework can help increase trust in the use of a given AI system, which includes

[35] Veritti et. al. (n 33).

[36] Ali et al. (n 21).

[37] Marda (n 12).

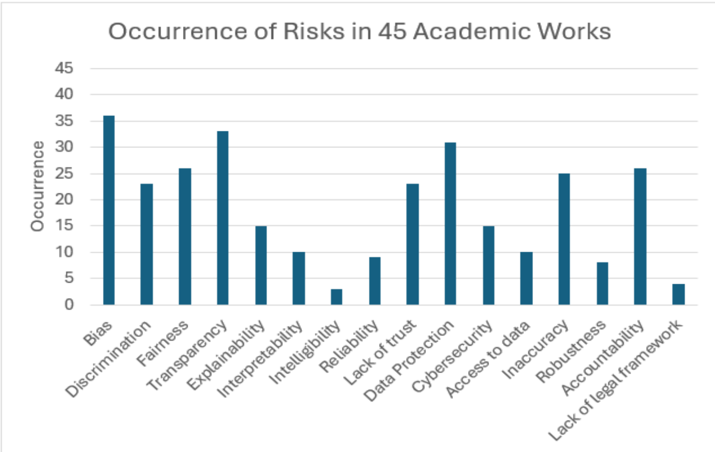
[38] Alžběta Krausová and Václav Moravec, 'Disappearing Authorship: Ethical Protection of AI-Generated News' (2022) 13 Journal of Intellectual Property, Information Technology, and Electronic Commerce Law 1.

[39] Rata Rokhshad, Maxime Ducret, Akhilanand Chaurasia, Teodora Karteva, Miroslav Radenkovic, Jelena Roganovic, Manal Hamdan, Hossein Mohammad–Rahimi, Joachim Krois, Pierre Lahoud and Falk Schwendicke, 'Ethical considerations on artificial intelligence in dentistry: A framework and checklist' (2023) 135 Journal of Denistry 1.

clarifying the responsibility issues arising from its usage.[40]

2. Evolution of Risks

The above risks appeared in different ways during the examined period in the reviewed works. The three most frequently addressed risks are (i) bias, (ii) transparency, and (iii) data privacy. However, it should be noted that some risks were occasionally assessed by the authors as subcategories of other risks. For example, the elements of a reliable (i.e., trustworthy) AI-based system include fairness, transparency, interpretability, explainability, and robustness.[41] In my view, it also happens that some of the risk names are treated as synonyms rather than distinct risks. Based on this, it is difficult to examine the occurrence rate of the risks from an objective perspective. Nonetheless, Figure/Table 2 contains the number of occurrences of each risk in the 45 reviewed works. The earliest publication in the examined literature is from 2016, while the most recent one is from 2023.



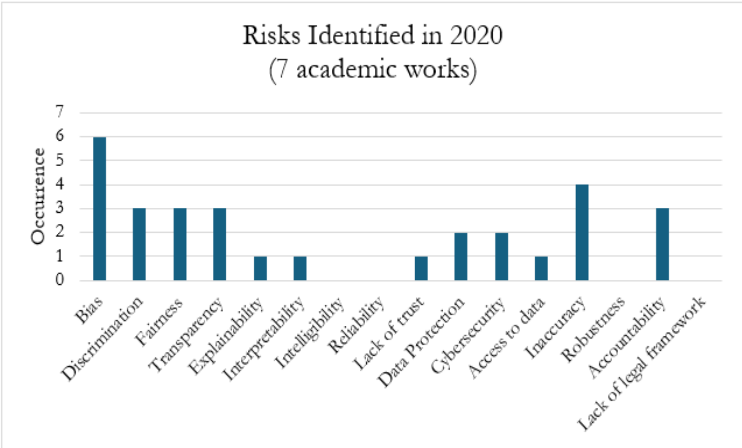
Figure/Table 2 - Occurrence of Risks in 45 Academic Works.

The above risks appeared differently in the reviewed works during the examined period. The three most commonly addressed risks were: (i) bias, (ii) transparency, and (iii) data protection. However, it is necessary to mention that some risks were occasionally considered as subcategories of other risks. For example, a trustworthy AI system element is fairness, transparency, interpretability, explainability, and stability. In my opinion, it also occurs that certain risk terms are

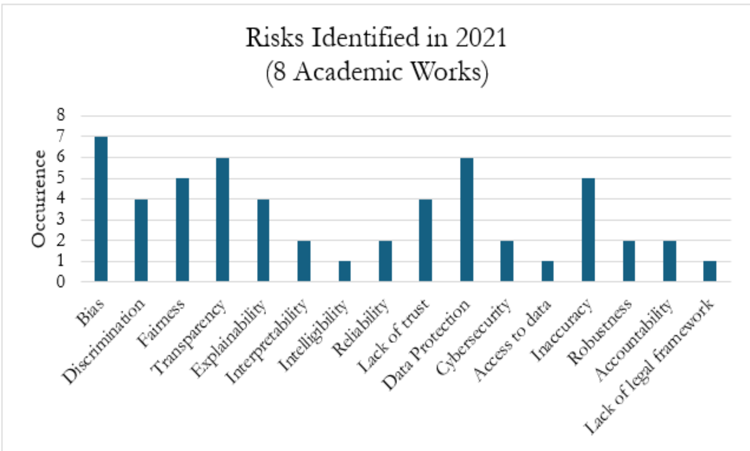
[40] Zhilian Huang, Mithun Mohan George, Yi-Roe Tan, Karthiga Natarajan, Emily Devasagayam, Evonne Tay, Abi Manesh, George M Varghese, Ooriapadickal Cherian Abraham, Anand Zachariah, Peiling Yap, Dorothy Lall and Angela Chow, 'Are physicians ready for precision antibiotic prescribing? A qualitative analysis of the acceptance of artificial intelligence-enabled clinical decision support systems in India and Singapore' (2023) 35 Journal of Global Antimicrobial Resistance 1.

[41] Ali et al. (n 21).

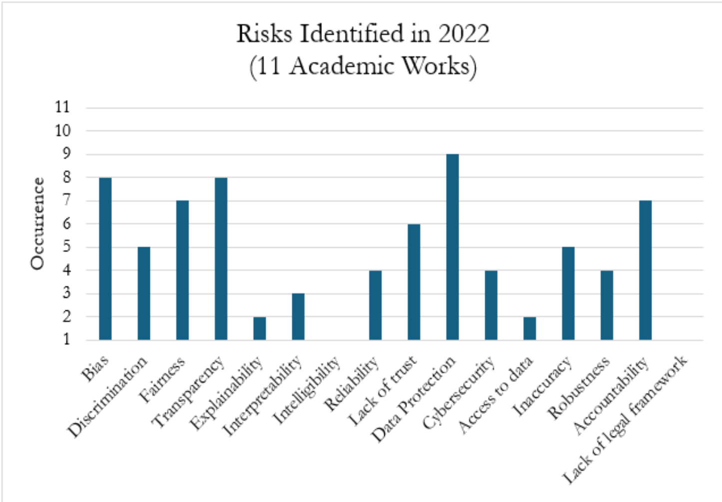
evaluated as synonyms, rather than separate risks. Based on this, it is difficult to objectively examine the occurrence ratio of the risks. Nevertheless, Figure/Table 2 presents the number of occurrences of each risk examined in the 45 works. The earliest published work in the review is from 2016, while the most recent one is from 2023.



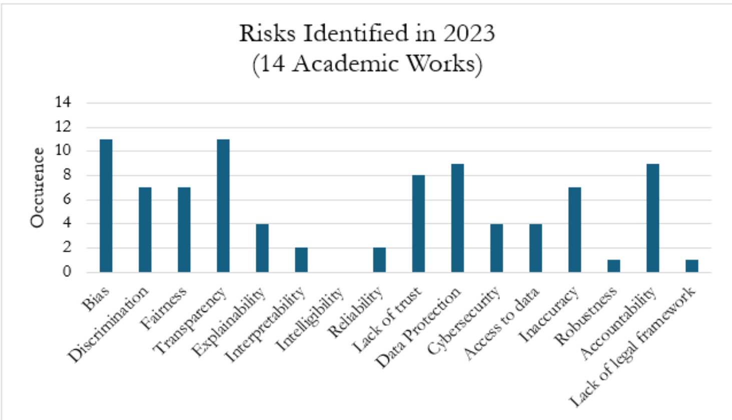
Figure/Table 3 – Risks Identified in 2020.



Figure/Table 4 – Risks Identified in 2021.



Figure/Table 5 – Risks Identified in 2022.



Figure/Table 6 – Risks Identified in 2023.

The analysis indicates a trend in the changing focus of issues that concern academic experts the most regarding AI-based decision-making and decision-support systems. However, the reviewed literature often addresses specific risk elements with minimal depth, resulting in a lack of detailed analysis for most of these risks. Furthermore, at the definitional level, the identification and clear definition of these risks are almost entirely absent. This raises the question of whether the risk elements examined by academic researchers align with the areas considered critical by market players. The following section attempts to answer this question.

3. Comparing results

To determine whether the risk areas most examined by scientific experts align with the risk areas that market participants – namely, corporations and small and medium-sized enterprises – consider important and relevant, I base my analysis on the annually published AI Index Report, as referenced in the introduction. According to a study conducted by *McKinsey & Company* and featured in the 2021 AI Index Report, a survey of 1,872 companies in 2019 found that cybersecurity was the most relevant AI-related risk (according to 62% of respondents). This was followed by regulatory compliance (50%), explainability (45%), protection of personal data (39%), organizational reputation (35%), workforce displacement (34%), fairness and justice (including bias, according to 26% of respondents), and so on.[42] By 2022, this trend had not changed significantly. According to respondents, 59% still identified cybersecurity as the most relevant risk, followed by regulatory compliance (45%), protection of personal data (40%), explainability (37%), organizational reputation (32%), fairness and justice (30%), and workforce displacement (28%).[43]

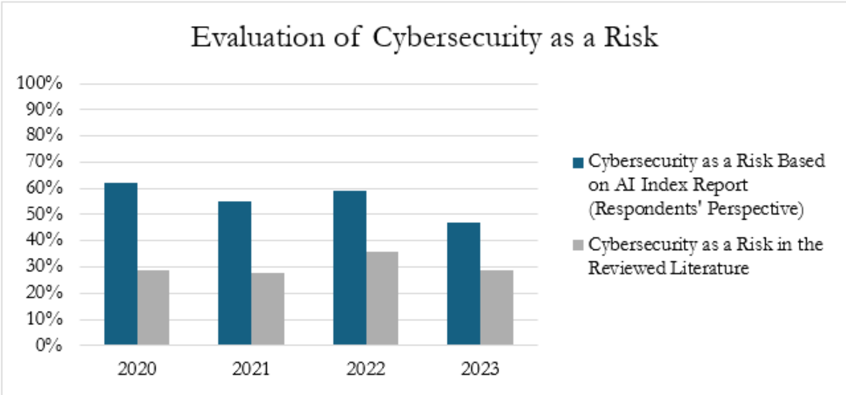
In contrast to the previous years, the 2024 AI Index Report indicates that, based on responses from over 1,000 organizations across 20 countries and 19 industries, the greatest perceived risk is related to the protection and management of personal data. This is followed by concerns about the reliability of AI systems, their security (including cybersecurity), transparency, and fairness.[44]

It is evident that the risks examined by scientific experts present a distinctly different picture compared to the responses from market organizations and stakeholders. An exception to this is the risk related to data protection, which remains a top priority according to the 2024 AI Index Report. However, biases—considered under the category of fairness in the AI Index Report—are viewed as a relatively less significant risk (see, for example, Figures/Tables 7-9).

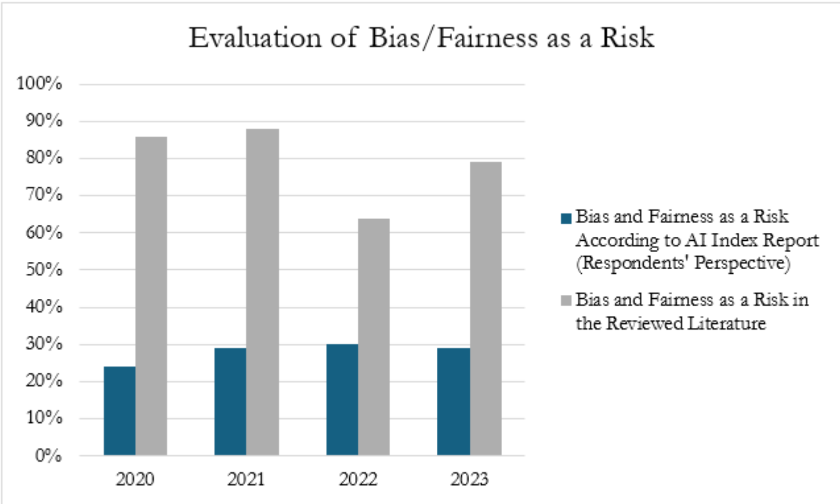
[42] Daniel Zhang, Saurabh Mishra, Erik Brynjolfsson, John Etchemendy, Deep Ganguli, Barbara Grosz, Terah Lyons, James Manyika, Juan Carlos Niebles, Michael Sellitto, Yoav Shoham, Jack Clark, and Raymond Perrault, 'The AI Index 2021 Annual Report' (2021) Stanford University Human-Centered Artifical Intelligence <https://aiindex.stanford.edu/wp-content/uploads/2021/11/2021-AI-Index-Report_Master.pdf> accessed 12 July 2024.

[43] Nestor Maslej, Loredana Fattorini, Erik Brynjolfsson, John Etchemendy, Katrina Ligett, Terah Lyons, James Manyika, Helen Ngo, Juan Carlos Niebles, Vanessa Parli, Yoav Shoham, Russell Wald, Jack Clark, and Raymond Perrault, 'The AI Index 2023 Annual Report' (2023) Stanford University Human-Centered Artifical Intelligence <https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf> accessed 12 July 2024.

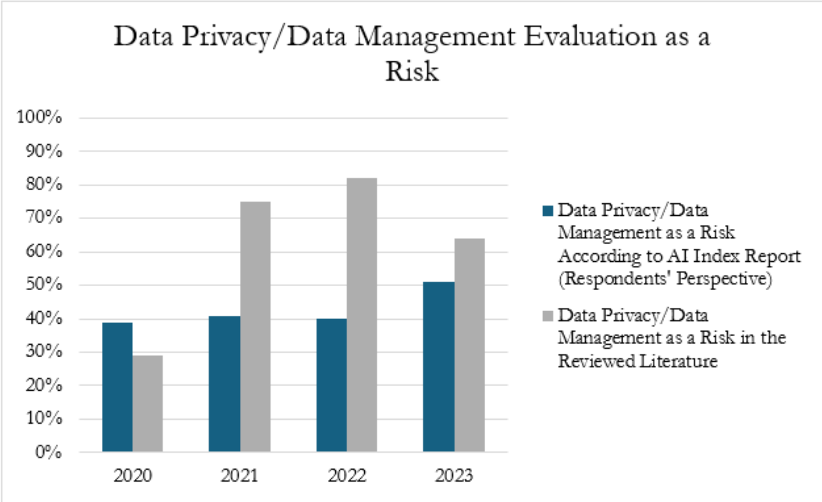
[44] Maslej et al. (n 4).



Figure/Table 7 – Evaluation of Cybersecurity as a Risk.



Figure/Table 8 – Evaluation of Bias/Fairness as a Risk.



Figure/Table 9 – Evaluation of Data Privacy/Data Management as a Risk.

IV. Proposed solutions to address the risks

Among the reviewed literature, several works also present potential solutions to address the aforementioned risks. From these, the following strategies—selected based on subjective evaluation—appear to be suitable for mitigating and managing the identified risks effectively.

In my view, corporate leadership must place significant emphasis on minimizing the aforementioned risks or, at the very least, take comprehensive measures to achieve such minimization. Beyond preventing financial losses for the company, implementing an effective risk management protocol can also safeguard corporate reputation during the implementation of AI-based decision-making and decision-support systems. To effectively minimize these risks, it is advisable, based on the reviewed literature, to first assess the planned AI system from a contextual perspective. This involves examining the industry's regulatory environment, especially if the AI application targets a highly regulated sector with specific compliance requirements. Additionally, it is essential to evaluate the potential negative impacts of erroneous AI decisions, such as the scope of impact, the vulnerable groups affected, whether the consequences are internal or external from the company's standpoint, the potential effects on human rights, the reversibility of the impact, and the expected duration of any negative outcomes. From a procedural standpoint, it is necessary to consider the technical aspects related to the complexity of the system and its degree of interoperability. In terms of business considerations, factors such as the company's risk appetite, the extent of human intervention, and alternative protocols for damage control must be analyzed. On the technological side, it is important to determine the type of algorithm used, whether third-party involvement (e.g., external IT support) is necessary, the level of transparency achieved, the expected accuracy, and the speed of the system's learning process. From the data management/ data base perspective, it is crucial to examine the dataset used for input-output processing, whether it includes personal data (or special categories of data), the presence of anonymization measures, as well as the size and coverage of the database.[45]

If management has mapped out the aforementioned factors, it is worth considering which type of AI-based system should be implemented to support the given operational decision-making process. In other words, how much decision-making power should be allocated to human judgment versus the AI-based system. In this regard, we can consider three main categories. Moving in the order of decreasing human autonomy, we first have the so-called *human-in-the-loop* solution, where the AI-based system serves as a simple support tool. In this case, the independence of the AI system is low, and management is fully involved in the decision-making process, thus bearing the responsibility as well. A more autonomous solution is the *human on-the-loop* model, where management's role is limited to approving or rejecting the decision. Here, the AI system operates with a higher degree of independence, but the final decision still lies with management. Finally, we can talk about the *human-out-of-the-loop* system, which is fully automated, meaning that management does not participate in the decision-making process. Based on the above, if the decisions require rapid resolution, the outcome of the decision is unlikely to involve human rights violations, and the decision-making process is frequently repeated, it is advisable to implement a human out-of-the-loop system. In cases where the decision-making process may potentially involve human rights concerns or is generally complex and intricate, it is recommended to use either a human in the loop or, if appropriate, a human on the loop system.[46] In my view, the corporate AI implementations that will succeed in the coming years are those preceded by comprehensive risk analysis and risk management efforts. This is essential for management to not only reduce the likelihood of potential risks but also to minimize the occurrence of damages during day-to-day operations.

When implementing AI-based decision-making and decision-support systems, as well as during their application within the European Union, it is necessary to comply with the regulatory environment shaped by two major legal frameworks. Firstly, due to the automated decision-making process—if it involves the processing of personal data—it is essential to meet the specific requirements outlined in the GDPR. One of the specific requirements is that data controllers using automated decision-making must inform data subjects, during the prelimi-

[45] Lee (n 30).

[46] Stanislav Hristov Ivanov, 'Automated decision-making' (2023) 25 Foresight 1.

nary information process, about the fact of automated decision-making, the logic employed, and its significance and potential consequences. This same obligation to inform applies when the data subject exercises his or her right of access.[47] The GDPR provides additional rights to data subjects in cases where decisions based solely on automated data processing (i.e., automated decision-making) would have legal effects or similarly significantly affect them. In such cases, the data subject may choose to opt out of this type of data processing. However, there are exceptions where this right cannot be exercised. For example, if the decision is based on the data subject's explicit consent, or if it is necessary for entering into or performing a contract between the data subject and the data controller, etc. Nevertheless, even in these scenarios, the data subject has the right to request human intervention from the data controller, express their viewpoint, and contest the decision.[48] A question may arise as to which of the three types of automated decision-making and decision-support systems mentioned above are subject to this specific rule. At first glance, it is clearly the human out-of-the-loop model that requires the application of these special GDPR provisions. However, the Court of Justice of the European Union, in case C-634/21 SCHUFA Holding (Scoring), concluded that it also qualifies as an automated individual decision if the data controller decides on a contract with a client by employing a third-party service provider that uses automated decision-making to determine the positive or negative outcome of the contract, and the data controller automatically adopts this decision.[49] As a result, even the human-in-the-loop model may fall under the scope of automated decision-making regulated by the GDPR. Therefore, during corporate integration, special attention must be given to this aspect. It is advisable to establish procedures that ensure decisions made by human-in/on-the-loop-based systems are not classified as automated decision-making under the GDPR. This approach can help mitigate legal risks and ensure compliance with data protection regulations.

On the other hand, the regulatory framework for compliance is provided by the EU Regulation on Artificial Intelligence (AI Act).[50] For AI-based decision-making and decision-support systems used by companies, including small and medi-

[47] 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 [2016] OJ L119/1, art. 13 and 15.

[48] ibid, art 22.

[49] C-634/21 *SCHUFA Holding (Scoring)* [2023] ECLI:EU:C:2023:957.

[50] 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 [2024] OJ L 2024/1689.

um-sized enterprises, the AI Act sets different compliance requirements based primarily on the category of the system, and secondarily on the role of the organization (e.g., as a provider or user). For instance, if a decision-support system is employed for monitoring employee performance and involves profiling, it will be classified as a high-risk AI system. In such cases, the system must meet stringent compliance criteria, including transparency obligations, risk management processes, and human oversight mechanisms, due to the potential impact on individual rights and freedoms.[51] In this case, organizations must comply with requirements such as implementing a risk management system, establishing effective data governance, providing technical documentation, maintaining records, ensuring transparency, and offering adequate information to users. Although compliance with these requirements will only be mandatory starting from August 2, 2027,[52] it is advisable to begin integrating these measures into corporate or small and medium-sized enterprises' AI systems now. It is worth noting, however, that in addition to the aforementioned requirements, the AI Act imposes many other compliance criteria. Therefore, it is crucial to identify the specific regulatory requirements and frameworks tailored to the particular AI system being integrated. This proactive approach will help ensure that the system meets all relevant legal obligations and is well-prepared for future compliance audits.

V. Conclusion and suggestions for further research

Based on the results of the research, I conclude that there is a greater need for examining risks related to the cybersecurity, data management, and reliability of AI-based systems for market players, i.e., companies, as well as small and medium-sized enterprises, rather than focusing on bias and other risks described above. Furthermore, there is a strong demand for a deeper, possibly sector-specific investigation of individual risks, which could contribute to the risk management of market players. One of the risks associated with AI-based systems that is missing is environmental protection. Additionally, an ESG (Environmental, Social, Governance) perspective is also lacking in the examination of these types of AI-based systems. Based on the above, I also identify the impact of these supporting systems on the workforce as an area for further investigation. Moreover, I consider it useful to review the risk database related to AI published by researchers from the *Massachusetts Institute of Technology* in August 2024,[53] with a

[51] ibid, art. 6, para 3.

[52] ibid, sec. 2 and art. 113(c).

[53] Peter Slattery, Alexander K. Saeri, Emily A. C. Grundy, Jess Graham, Michael Noetel, Risto Uuk, James Dao ,Soroush Pour, Stephen Casper, Neil Thompson, 'A systematic evidence review and common frame of reference for the risks from artificial intelligence' (2024) <https://www.researchgate.net/publication/383089263_The_AI_Risk_Repository_A_Comprehensive_Meta-Review_Database_and_Taxonomy_of_Risks_From_Artificial_Intelligence?channel=doi&linkId=66bc0c43299c327096c752dc&showFulltext=true accessed> 23 August 2024.

focus on AI-based decision-making and decision-support systems used by managers of companies, as well as small and medium-sized enterprises.

VI. Limitation

I would like to highlight the following limitations regarding the current research:

- In many cases, the scholarly articles do not define the specific risk, i.e., what exactly is meant by it, which sometimes leads to the same risk being identified under two different names.
- Given the vast amount of literature on AI that is published monthly, the literature review as a genre can only provide a snapshot of the current body of work on AI-related risks.
- The literature review may not be entirely suitable for examining the subject from a social science perspective, as it is subject to the author's subjectivity.

Appendix 1 - The Reviewed Scholarly Works

Name of the author(s)	Title of the article	Scope of the study	Date of publication
Maryanrizinski, Hristijan Peshov, Kostadin Mishev, Lubomir T. Chitkushev, Irena Vodenska, And Dimitar Trajanov	*Ethically Responsible Machine Learning in Fintech*	Ethical challenges in the fintech sector, particularly bias, discrimination, differentiated pricing, conflicts of interest, and data privacy	2016
Vidushi Marda	*Artificial intelligence policy in India: a framework for engaging the limits of data-driven decision-making*	Recommendation of a framework for understanding the impacts of AI, focusing on the three main stages of introducing machine learning: the data, model, and application stages	2018

Daniel Schönberger	*Artificial intelligence in healthcare: a critical analysis of the legal and ethical implications.*	The decision-making capabilities of AI technologies	2019
Kristin N. Johnson	*Automating the Risk of Bias*	Increased gender inclusion in the development of AI technologies and its impacts	2019
Heike Felzmann, Eduard Fosch Villaronga, Christoph Lutz and Aurelia Tamó Larrieux	*Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns*	The significance of the GDPR-based transparency requirement in the case of AI and automated decision-making systems	2019
Tetyana (Tanya) Krupiy	*A vulnerability analysis: Theorising the impact of artificial intelligence decision-making processes on individuals, society and human diversity from a social justice perspective*	Social issues related to the application of AI decision-making processes	2020
Lorwai Tan, David Tivey, Helena Kopunic, Wendy Babidge, Sally Langley and Guy Maddern	*Artificial intelligence technology in surgery*	The significance of AI in surgery	2020
Sergio Alberto Gramitto Ricci	*Artificail Agents in corporate boardrooms*	The use of AI in corporate boards	2020
Helen Smith, Kit Fotheringham	*Artificial intelligence in clinical decision-making: Rethinking liability*	Possible outcomes of *negligence* lawsuits filed against clinicians and software development companies regarding the use of AI-based systems with human clinical oversight	2020

Christopher M. Bruner	*Distributed ledgers, artificial intelligence and the purpose of the corporation*	The impact of emerging technologies from both positive and normative perspectives, focusing on how these developments might influence debates regarding corporate objectives in the context of publicly listed companies	2020
ANDREW D. SELBST	*Negligence and ai's Human Users*	Examining four complications arising from the unique nature of AI in relation to negligence	2020
Michael Hilb	*Toward artificial governance? The role of artificial intelligence in shaping the future of corporate governance*	How the continuous development and adaptation of AI impacts the practice of corporate governance	2020
Jocelyn Maclure	*AI, Explainability and Public Reason: The Argument from the Limitations of the Human Mind*	Interpretation of the explainability problem of AI and highlighting its ethical significance	2021
Charlotte A. Tschider	*Beyond the "black box"*	Examination of transparency and explainability	2021

Marcus Buckmann, Andy Haldane and Anne-Caroline Hüser	*Comparing minds and machines: implications for financial stability*	Does human or artificial intelligence better support a stable financial system? - Aspects of human and artificial intelligence decision-making behavior	2021
Ricardo Francisco Reier Forradellas and Luis Miguel Garay Gallastegui	*Digital Transformation and Artificial Intelligence Applied to Business: Legal Regulations, Economic Impact and Perspective*	The impact of AI and digital transformation on business	2021
Fiorella Operto	*Elements of Roboethics*	The ethical, legal, and societal implications of robotics, with particular focus on advanced robotics applications	2021
Björn Lundgren	*Ethical machine decisions and the inputselection problem*	The role of factual uncertainty in moral decision-making	2021
Friedrich Hamadziripi, Howard Chitimira	*The Integration and Reliance on Technology to Enhance the Independence and Accountability of Company Directors in South Africa*	The integration of technology and the reliance on technology is intended to be discussed in order to improve corporate governance principles in developing countries, such as South Africa	2021

Tae Wan Kim, Bryan R. Routledge	*Why a Right to an Explanation of Algorithmic Decision-Making Should Exist: A Trust-Based Approach*	It examines algorithmic decision-making and the right to explanation	2021
Leo H. Chiang, Birgit Braun, Zhenyu Wang, Ivan Castillo	*Towards artificial intelligence at scale in the chemical industry*	Application of AI in the chemical industry	2022
Oscar Rodríguez-Espíndola Ali Emrouznejad	*Analysis of the adoption of emergent technologies for risk management in the era of digital manufacturing*	The application of big data, AI, cloud computing, and blockchain in risk management	2022
Therese Enarsson, Lena Enqvist, Markus Naarttijärvi	*Approaching the human in the loop – legal perspectives on hybrid human/algorithmic decision-making in three contexts*	Different forms of hybrid decision-making systems	2022
Jingchen Zhao	*Artificial Intelligence and Corporate Decisions: Fantasy, Reality or Destiny*	The role of AI in corporate boards	2022
Deepika Chhillar and Ruth V. Aguilera	*An Eye for Artificial Intelligence: Insights Into the Governance of Artificial Intelligence and Vision for Future Research*	The intersection of corporate governance and AI	2022
Lee, M.S.A.; Cobbe, J.; Janssen, H.; Singh, J.	*Defining the scope of AI ADM system risk assessment*	It examines the ambiguity of definitions inherent in AI and their relationship with automated decision-making, as well as the potential consequences for the organizational understanding of system risks	2022

Alžběta Krausová and Václav Moravec	*Disappearing Authorship Ethical Protection of AI-Generated News from the Perspective of Copyright and Other Laws*	It demonstrates that the current Czech legal environment does not encourage accountability, responsibility, and transparency	2022
Muzaffer Eroğlu, Meltem Karatepe Kaya	*Impact of Artificial Intelligence on Corporate Board Diversity Policies and Regulations*	The potential impacts of AI on corporate board diversity policies and regulations	2022
Rozita Dara, Seyed Mehdi Hazrati Fard and Jasmin Kaur	*Recommendations for ethical and responsible use of artificial intelligence in digital agriculture*	Ethical challenges of AI application in agriculture, including fairness, transparency, accountability, sustainability, privacy protection, and robustness	2022
Agata Leszkiewicz, Tina Hormann and Manfred Krafft	*Smart business and the social value of AI*	Analysis of AI benefits and costs for a B2B company and its internal, external, and societal stakeholders	2022
Toan Huu Bui, Van Phuoc Nguyen	*The Impact of Artificial Intelligence and Digital Economy on Vietnam's Legal System*	The integration of AI and digital transformation into various applications and their regulations	2022

Zhilian Huang, Mithun Mohan George, Yi-Roe Tan, Karthiga Natarajan, Emily Devasagayam, Evonne Tay, Abi Manesh, George M. Varghese, Ooriapadickal Cherian Abraham, Anand Zachariah,	*Are physicians ready for precision antibiotic prescribing? A qualitative analysis of the acceptance of artificial intelligence-enabled clinical decision support systems in India and Singapore*	AI-based decision support systems used in the health-care sector	2023
Elif Kiesow Cortez and Nestor Maslej	*Adjudication of Artificial Intelligence and Automated Decision-Making Cases in Europe and the USA*	Legal cases related to AI and automated decision-making	2023
Anna van der Gaag; Robert Jago; Ann Gallagher; Kostas Stathis; Michelle Webster; and Zubin Austin	*Artificial Intelligence in Health Professions Regulation: An Exploratory Qualitative Study of Nurse Regulators in Three Jurisdictions*	The potential role and value of AI technologies in regulatory practice	2023
Stanislav Hristov Ivanov	*Automated decision-making*	Analysis of decision-making approaches related to AI: human in the loop, human on the loop, human out of the loop	2023
Daniele Veritti, Leopoldo Rubinato, Valentina Sarao, Axel De Nardin, Gian Luca Foresti, Paolo Lanzetta	*Behind the mask: a critical perspective on the ethical, moral, and legal implications of AI in ophthalmology*	The dangers, controversial aspects, and consequences of the application of AI in ophthalmology and other areas related to medicine	2023

Ralitsa Raycheva, Kostadin Kostadinov, Elena Mitova, Nataliya Bogoeva, Georgi Iskrov, Georgi Stefanov and Rumen Stefanov	*Challenges in mapping European rare disease databases, relevant for ML-based screening technologies in terms of organizational, FAIR and legal principles: scoping review*	Identification of the key challenges arising during the mapping of European databases for rare diseases, which are relevant from the perspectives of organizational, FAIR, and legal principles in relation to machine learning-based screening technologies	2023
Tanya Brigden, Colin Mitchell, Elizabeth Redrup Hill, Alison Hall	*Ethical and legal implications of implementing risk algorithms for early detection and screening for oesophageal cancer, now and in the future*	Identification of ethical and legal issues related to the application of an esophageal cancer risk prediction tool in primary care	2023
Rata Rokhshad, Maxime Ducret, Akhilanand Chaurasia, Teodora Karteva, Miroslav Radenkovic, Jelena Roganovic, Manal Hamdan, Hossein Mohammad-Rahimi, Joachim Krois, Pierre Lahoud, Falk Schwendicke	*Ethical considerations on artificial intelligence in dentistry: A framework and checklist*	Providing a framework and a checklist for evaluating AI applications used in dentistry from this perspective	2023

Andrew Mowbray, Philip Chung, Graham Greenleaf	*Explainable AI (XAI) in Rules as Code (RaC): The DataLex approach*	It examines the fundamental necessity of explainability and transparency in the implementation of *Rules as Code* and explores the similarity of this requirement with the concept of *explainable AI*	2023
Sajid Ali, Tamer Abuhmed, Shaker El-Sappagh, Khan Muhammad, Jose M. Alonso-Moral, Roberto Confalonieri, Riccardo Guidotti, Javier Del Ser, Natalia Díaz-Rodríguez, Francisco Herrera	*Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence*	It examines evaluation methods, available tools, *explainable AI datasets*, and other related aspects	2023
Tetyana (Tanya) Krupiy and Martin Scheinin	*Disability Discrimination in the Digital Realm: How the ICRPD Applies to Artificial Intelligence Decision-Making Processes and Helps in Determining the State of International Human Rights Law*	The impact of AI decision-making processes on individuals with disabilities	2023

Javed Iqbal, Diana Carolina Cortés Jaimes, Pallavi Makineni, Sachin Subramani, Sarah Hemaida, Thanmai Reddy Thugu, Amna Naveed Butt, Jarin Tasnim Sikto, Pareena Kaur, Muhammad Ali Lak, Monisha Augustine, Roheen Shahzad, Mustafa Arain	*Reimagining Healthcare: Unleashing the Power of Artificial Intelligence in Medicine*	The impact of AI on healthcare and the importance of ethical and balanced integration	2023
Tyler R Ray, Ryan T Kellogg, Kyle M Fargen, Ferdinand Hui, Jan Vargas	*The perils and promises of generative artificial intelligence in neurointerventional surgery*	It examines the dangers and promises of generative AI in neurointerventional surgery	2023
Marvin van Bekkum, Frederik Zuiderveen Borgesius	*Using sensitive data to prevent discrimination by artificial intelligence: Does the GDPR need a new exception?*	The GDPR rules regarding special categories of personal data hinder the prevention of AI-based discrimination	2023