



HAL
open science

Encouraging the Ethical Use of Artificial Intelligence: Risks and Recommendations across four Business Sectors

Lamia Berrah, Myriam Benfares, Marie-Isaline Bontemps

► **To cite this version:**

Lamia Berrah, Myriam Benfares, Marie-Isaline Bontemps. Encouraging the Ethical Use of Artificial Intelligence: Risks and Recommendations across four Business Sectors. AIR 2024: Artificial Intelligence Revolutions, Université Jean Monnet, Oct 2024, Roanne, France. hal-04770469

HAL Id: hal-04770469

<https://hal.science/hal-04770469v1>

Submitted on 6 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Encouraging the Ethical Use of Artificial Intelligence: Risks and Recommendations across four Business Sectors

Myriam Benfares
Université Savoie Mont Blanc
Annecy, France
myriam.benfares@etu.univ-smb.fr

Marie-Isaline Bontemps
Université Savoie Mont Blanc
Annecy, France
Marie-Isaline.Bontemps@etu.univ-smb.fr

Lamia Berrah
LISTIC, Université Savoie Mont Blanc
Annecy, France
lamia.berrah@univ-smb.fr
0000-0002-5165-6953

Abstract—The use of Artificial Intelligence (AI) is expanding fast and in all business sectors, offering and promising technical performance that exceeds human abilities. However, uncontrolled, this use can also threaten human well-being (replacement, demotivation, discrimination, dependency, etc.). In this sense, this paper analyzes the ethical risks of using AI in four diversified key business sectors: Human Resources Management, Healthcare, Entertainment and Education. From testimonials in literature, the aim of this exploratory study is to identify first the specific ethical risks for each sector and then to propose recommendations to address these ethical issues. Unethical uses of the data used and produced by AI present a variety of ethical risks that require urgent response. For instance, for the Human Resources Management sector, developing an ethical framework is recommended, promoting transparency and making public funding conditional on meeting strict ethical criteria. In the Healthcare sector, training staff involving ethics committees is suggested, as well as, developing overall guidelines, making reimbursement conditional on ethical compliance and establishing independent ethical evaluation bodies. For the Entertainment sector, recommendations include raising awareness among content creators, involving citizen ethics committees, defining an industry ethics charter and labeling AI-generated works. Finally, for the Education sector, creating AI ethics committees, is proposed requiring transparency from providers, sharing best practices, training students and ensuring access to educational data. Generalizing and systematizing the approach will be the next step in this preliminary analysis.

Keywords—Artificial Intelligence (AI), Ethics, Risks, Business sectors, Recommendations

I. INTRODUCTION

Artificial Intelligence (AI) has become omnipresent in our daily lives, transforming all the business sectors related to modern society. Industry, healthcare, finance, education commerce, etc. have all undergone radical changes in the way they operate and the services they offer. The “4.0” label, in reference to the 4th industrial revolution [1], is thus the label for this transformation and the aim of companies and organizations. AI technology-based applications ranging from virtual assistants and recommendation systems to autonomous cars and medical diagnostics illustrate these changes and the great potential of digitalization in general. The use of advanced technologies in such businesses is providing many advantages that can be observed throughout, for instance:

- The technical specificities that these technologies facilitate, being increasing the product/service quality and value and shortening the deliveries.
- The traceability of data, facilitating the management and enhancing reactivity.
- The relief of humans from tedious or repetitive tasks.
- The reduction of the costs associated with the autonomy of the systems.

However, the autonomy which characterizes the so-called systems of the Future as well as the associated profusion of data produced can lead to poorly controlled, inappropriate or even diverted or hazardous decisions which therefore lead to potential ethical risks and also threat to performance [2] [3]. Human replacement and loss of skills, algorithmic discrimination, data privacy mismanagement, automated judgement, lack of transparency and surveillance policies, are examples of the ethical risks possibly encountered and of which there is a growing awareness. Such risks are naturally different, according to the considered technologies, like virtual reality, IoT and technologies that can learn by itself, like machine learning systems and Generative Artificial Intelligence [4].

Faced with these challenges, the following question (Q) arises: *How to mitigate the ethical issues associated with the use of AI in managing businesses?* Indeed, the speed and acceleration of this use induces a need for an urgent, concrete and operational answer to the associated question of ethics. Even though frameworks and standards that have been defined to circumscribe the use of advanced technologies in general, the proposals remain conceptual, highly focused, almost descriptive and partially handling the operational use of these advanced technologies. Indeed, among the most significant frameworks and guidelines, one can recall the Corporate Social Responsibility (CSR) which essentially associates financial and non-financial aspects and covers the strategic, business and legal aspects of the organizations [5], the social ISO26000 standard and the European General Data Protection Regulation (GDPR) which concerns the commitment of the processing carried out on the data by the collector [6]. Therefore, much remains to be done to ensure that AI technologies are used responsibly and fairly. It is thus essential to develop practical and effective recommendations to guide developers, policy makers and users in the ethical implementation of AI.

Parts of the work presented in this paper are carried out in the context of the “Ethical Risk Assessment and Management for Industry of the Future” ANR ETHICS40 (ANR-236CE1060014) project, funded by the French ANR research program.

This study subscribes to the general theme of the consideration of ethical risks in the use of what is referred to as AI. The aim of this research is to provide decision-makers with operational tool for systematically integrating these risks into their decision-making. In this sense and as a first step, this paper proposes an exploratory study that consists of analyzing some business sectors, identifying the ethical risks associated with the use of AI technologies and suggesting recommendations in order to mitigate these risks. The idea is to validate the hypotheses at the basis of this need for an operational tool, i.e. that: *ethical risks do indeed exist, whatever the business sector under consideration (H1)*, and that: *these risks can be prevented or mitigated by compliance with appropriate recommendations (H2)*. This subject is complementary to the study currently being carried out on the management of ethical risks in the performance management of industrial systems of the future [7] and aims to explore other business sectors for the sake of genericity.

Hence, with the hypotheses (H1) and (H2), this paper deals with the following sub-question (Q') of (Q): *What are the ethical risks associated with the use of AI and the recommendations for mitigating them?* To answer this question, a focus is made on four specific AI use cases, namely: Human Resources Management, Healthcare, Entertainment and Education. Through a comparative analysis of these business sectors, the idea is also to explore the main common and specific ethical issues that need to be addressed. For that purpose, Section II shortly brings some elements concerning ethics and AI ethical issues in general, while Section III focuses, on the basis of testimonies found in the literature, on the four considered business sectors and the comparative analysis of their associated ethical risks. Recommendations are therefore deduced in Section IV with regards to the ethical risks identified, by respective sector. These preliminary results are then discussed in Section V. And, finally, concluding remarks and perspectives end the study.

II. ARTIFICIAL INTELLIGENCE AND ITS ETHICAL ISSUES

A. Ethics: definition and paradigms

Ethics is a moral philosophy concept that is delicate to easily grasp. Defining this concept, has been a subject of discussion in philosophy, sociology, law, science (and more recently data sciences), etc. for centuries with many paradoxes at the heart of it. However, the following general definition that simply captures the essence of ethics is adopted here: *“the strive for the good life, with oneself and others, in just/fair institutions”* (in French: « *Une vie bonne, avec et pour autrui, dans des institutions justes* ») [8].

Several philosophical schools of thought consider ethics from different points of view, leading to the establishment of various paradigms. The two main paradigms that are easier to implement are respectively: deontology (where decisions are made using immutable ethical rules) and consequentialism - including utilitarianism - (where decisions are made using rules evaluated in terms of possible ethical consequences) paradigms [9] [10] [11]. Therefore, ethical issues are situations that raise moral and virtue questions or dilemmas about what is right or wrong, just or unjust, fair or unfair, etc. [12] [13].

B. Artificial Intelligence

Multidisciplinary academics and industrialists as well as politicians are currently investigating AI and its exploitation. For the sake of respectively, illustration, simplicity and comprehensiveness, the vision of the European Parliament is considered here [14], in which AI is defined as: *“technology that allows machines to simulate human intelligence by learning and adapting from data”* [14]. Therefore, AI identifies *“the ability of a machine to display human-like capabilities such as reasoning, learning, planning and creativity”* [14]. AI encompasses a wide range of technologies, including machine learning, speech recognition, computer vision and natural language processing. As an analysis of the various technologies associated with AI is outside the scope of this study, the discussion will be confined in this study to AI as a concept covering all the advanced technologies currently known.

Moreover, a wide range of business sectors are using AI, in various degrees. Among these sectors and as an illustration, one can mention:

- Agriculture: AI can be used to optimize irrigation, fertilization and pest control.
- Education: AI can be used to personalize learning, provide feedback to students, and identify students who are struggling.
- Environment: AI can be used to monitor the state of the environment, track climate change and develop sustainable solutions.
- Finance: AI can be used to detect fraud, manage risk and offer financial advices.
- Justice: AI can be used to analyze evidence, predict offender recidivism and identify potential cases of discrimination.

In this sense, AI can bring many benefits to society in general. However, it also raises challenges, particularly ethical ones. Even if it is not yet clear how this will work in practice, there is general awareness of the need to use it responsibly and ethically, to guarantee not only human well-being but also the performance of organizations [15].

AI benefits are currently well-known and cover all the sectors. Among them [14] [15] (Fig. 1):

- Better healthcare: AI can be used to develop new drugs and treatments, improve disease diagnosis, and provide personalized patient care.
- Safer and cleaner transportation: AI-powered autonomous vehicles can reduce road accidents and CO₂ emissions.
- More efficient manufacturing: AI can be used to optimize production processes, reduce waste and improve product quality.
- Cheaper and more sustainable energy: AI can help develop renewable energy sources and manage power grids more efficiently.

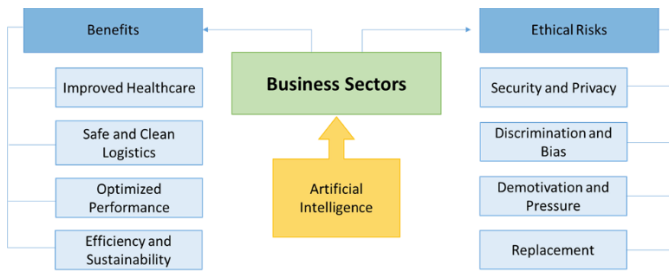


Fig. 1. Some benefits and ethical risks of using AI

Despite these surpassing features, AI also presents major ethical risks, such as [14] [15] [16] (Fig. 1):

- Security and Privacy Risks: AI systems can be used to hack computers, steal personal data, and even cause physical harm.
- Discrimination and bias: If AI systems are not designed and used responsibly, they can perpetuate existing discrimination in society.
- Demotivation: The performance of AI and its over-human capabilities can create pressure and overwork for humans.
- Job loss: AI-powered automation could lead to job losses in some sectors.

Therefore, an expanding set of business sectors are concerned with ethics, essentially because of the conjunction of the digital and physical worlds where humans interact with complex artificial systems. Managing ethical risks is a lever for human well-being as well as performance. However, identifying specifically such risks remains a challenge that is addressed in the following section, through the analysis of four particular business sectors and corresponding literature testimonies.

III. COMPARATIVE ANALYSIS OF ETHICAL RISKS ACROSS BUSINESS SECTORS

To go further in considering the ethical risks associated with an uncontrolled use of AI, four business sectors have been considered in this study, namely:

- Human Resources Management.
- Healthcare.
- Entertainment.
- Education.

Several arguments can be given concerning the choice of these four business sectors, which are naturally not intended to be exhaustive but merely exploratory, as mentioned before. Indeed, these sectors are at the forefront of society, becoming areas where AI is increasingly predominant. Moreover, these are sectors that are directly related to humans, the interactions between AI and humans being significant. Finally, everyone is familiar with AI technologies and has worked in at one time or another.

In addition, due to their diversity, the selected business sectors potentially offer a variety of uses for AI and the associated ethical risks. Lastly, the focus on these sectors is less, compared with industry, for example.

The proposed investigation approach for validating the (*H1*) and (*H2*) hypotheses and their corollary sub-question (*Q'*) (§ section I) is based on the collection of some testimonies given in the literature in line with the subject of ethics in the use of AI (by sector). These testimonies are used to guide the thinking around the following steps:

- Identify some ethical risks as well as their major consequences.
- Draw up a typology of ethical risks.
- Propose some recommendations for the mitigation of the identified ethical risks.
- And then discuss these first results obtained.

Fig. 2 below summarizes these different steps.

And for the sake of illustration, Table 1 hereafter gives three major ethical risks respectively associated with the use of AI in the business sectors mentioned before. Naturally, the proposed discussion remains very modest with regard to the considered risks.

By reviewing these business sectors together, it is possible to gain a deeper and more nuanced understanding of the implications of AI in various aspects of society. The following subsection translates what emerges from the succinct analysis of different use-case literature review.

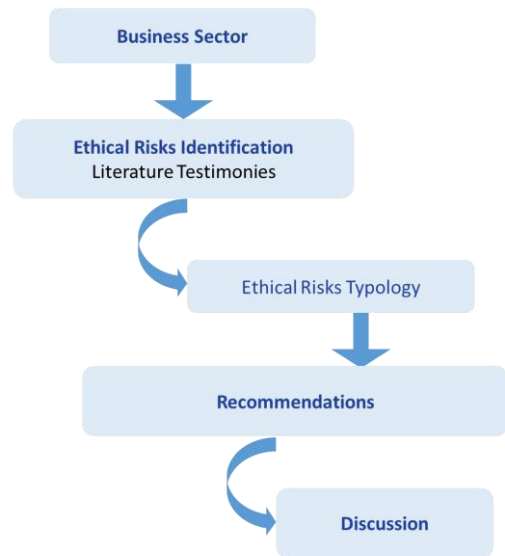


Fig. 2. Investigation methodology

TABLE I. USE CASES OF AI PRESENTING ETHICAL RISKS

Business Sector	Risk 1	Risk 2	Risk 3
Human Resources Management	Systemic discrimination	Lack of transparency	Violation of privacy
Healthcare	Data confidentiality	Bias in diagnosis and treatment	Wrongful responsibility and decision-making
Entertainment	Damage to reputation	Disinformation and manipulation	Infringement of copyright and intellectual property rights

Education	Bias in grading systems	Over-reliance on technology	Violation of privacy
-----------	-------------------------	-----------------------------	----------------------

A. Human Resources Management Sector

The use of AI in the Human Resources Management sector can lead to systemic discrimination. For example, Amazon's recruiting algorithm, abandoned in 2018, systematically discriminated against women because it learned from mostly male resumes, thus perpetuating existing gender biases [17]. In the case of Amazon, the imbalance between the number of employees identified as male and female is attributed in part to this gender bias. These bias influences judgments and decisions, hindering the advancement of women. Thus, the data used by the AI system may be biased towards male candidates, reflecting this imbalance. Additionally, applicants may experience transparency issues when rejected by automated systems without clear explanations, raising questions about the fairness of the process. Indeed, [17] explains that the use of AI can lead to a lack of transparency towards candidates, when the use of AI and automated systems is not proactively communicated to candidates. AI systems used for recruitment can analyze candidates' personal data from various online sources (social networks, blogs, etc.) without their explicit consent. This raises concerns about privacy and the protection of sensitive personal data.

Let us mention that only by discussing and elaborating only one risk with regard to the Human Resource Management Sector, taking reference only from one reference which is Amazon example, the analysis lacks completeness. Therefore, a more thoughtful point to be considered as a risk needs to be investigated from many references and situations.

B. Healthcare Sector

In the Healthcare sector, the use of AI requires access to sensitive medical data, which poses significant risks to patient data privacy and security. In this sense, [18] provides an exhaustive survey of privacy risks in what can be referred as Healthcare 4.0. For instance, ethical risks may appear when using non-anonymized medical data to train AI algorithms, or when AI is used to detect breast cancer from mammographic images. If training images contain patient identifying information, it would be possible to communicate model results to individuals, thereby violating their privacy. The use of data anonymization processes and federated learning techniques are thus recommended to be used in order to mitigate these risks.

Moreover, when AI is used to aid medical diagnosis, biases that are present in historical training data may appear, which are often derived from majority populations. This can lead to less accurate or inappropriate diagnoses and treatment recommendations for patients from minority or under-represented groups [19].

Furthermore, in this context of increasing the automation of medical decision-making, the issue of liability in the event of an AI diagnostic error remains unclear, raising legal and ethical concerns, with regards to the question: "who is responsible: the AI developer, the healthcare professional, or the institution?" [20]. As an unfortunate illustration, let us mention the case of the patient who, in 2021, was administered an incorrect dose of an anti-cancer drug by an AI dosage assistance system, causing him to die from an overdose. In such incidents,

responsibility was shared between the developer of the AI system, which had not adequately tested the system for this specific type of medication, and the hospital, which had not implemented safety protocols for the use of the AI system.

C. Entertainment Sector

In the Entertainment industry, one major risk is related to the deep fakes phenomenon [21] [22], which can seriously damage individual's reputation. For example, deep fakes have been used to create compromising videos of celebrities, causing irreversible damage to their public image. Let us mention for instance the famous deep fake video, which, in 2019, compromised the actor Tom Hanks, showing him making racist and offensive statements. The video was quickly shared on social networks and caused a public outcry. Tom Hanks was forced to issue a statement denying the statements made in the video and claiming that it was a deep fake. This incident illustrates the potential of deep fakes to be used to damage a person's reputation [23], and by extension of organizations.

Deep fakes can also be used to spread false political information, potentially influencing election results and public perception. For example, in 2020, a deep fake video of Donald Trump was released online, showing him conceding defeat in the presidential election. The video was shared by Donald Trump supporters on social networks and was presented as evidence that the election had been rigged. This video is an example of how deep fakes can be used to spread disinformation and manipulate public opinion [24]. It is thus easy to imagine the possible consequences of this type of use of AI on intra- and inter-organizational relations, and the mistrust and pressure that can result.

Additionally, in the same spirit, unauthorized use of a person's voice and likeness by AI systems may result in copyright and intellectual property violations, such as in the case of generated advertising by AI using without authorization the image of a well-known actors or personalities. For example, in 2021, a deep fake video of the actress Scarlett Johansson was circulated online, showing her in an advertisement for a cosmetic product she had never approved. The video was created by an artist without Scarlett Johansson's consent and was shared widely on social networks. This incident illustrates the potential of deep fakes to be used to infringe copyright and image rights [25], that can be used, by extension, to circumvent instructions or impose excessive workloads, for example.

D. Education Sector

In the Education sector, automated grading systems can be biased, favoring certain groups of students due to biases in the training data [26]. For example, automated grading systems have been criticized for systematically favoring students from advantaged backgrounds. For instance, in 2020, a study revealed that an automated grading system used in an American university discriminated against black students. The system was more likely to predict that black students would fail their courses than white students, even when the two groups of students had comparable grades and academic backgrounds. This example illustrates the potential of automated grading systems to perpetuate existing biases and undermine equity in education. Not only is this unfair, but it can also lead to demotivation and low self-esteem in those who suffer from it.

Excessive and unbalanced use of AI in the Education sector can also reduce human interaction, limiting the development of students' social and emotional skills [27]. For example, a 2021 study highlighted that students who took a course entirely online had lower levels of social interaction and emotional skills than students who took the same course face-to-face. This study suggests that an over-reliance on technology in education can be detrimental to students' social and emotional development. Since it is precisely humans' emotions and social nature that differentiate them from AI, this use can lead to a loss of creativity and motivation.

Additionally, collecting and analyzing personal data about students without their consent raises significant ethical risks regarding data privacy and security. For example, in 2018, it was revealed that an educational technology company in the United States was collecting data on students without their consent or that of their parents. These data included information about their learning habits, academic achievements and even facial expressions. The Company then used these data to enrich its data base and develop targeted marketing products. This incident illustrates the potential for AI systems to violate privacy if the data are not managed responsibly [28]. This naturally applies whatever the sector.

E. Synthesis

As a synthesis of this exploratory analysis, one can observe that common ethical risks across these business sectors include the perpetuation of bias and discrimination, as seen in the Human Resources Management sector, the Healthcare sector and the Education sector. Data confidentiality and privacy are also major concerns in the Human Resources Management sector, the Healthcare sector, and the Education sector. Challenges around transparency and accountability are pervasive, making it difficult to understand and react to the decisions AI makes and then to assign responsibility for errors.

On the other hand, each business sector presents its specific risks. In the Entertainment sector, the unique issues surrounding deep fakes and digital content manipulation are of particular concern, leading to reputational damage and misinformation. In the Healthcare sector, responsibility for medical decisions and inequitable treatment are critical ethical issues. In the Education sector, overreliance on technology and its impact on students' social skills are major ethical concerns.

Once this preliminary identification of the ethical risks associated with an uncontrolled use of AI in the four considered business sectors has been completed, the following section proposes now some recommendations to manage these risks.

IV. ETHICAL RECOMMENDATIONS

In the light of what has been seen in the previous section, this section is devoted to proposing thus some preliminary recommendations, by business sector, for managing the ethical risks associated with the use of AI.

A. Human Resources Management Sector

- Develop a sector-specific ethical reference framework and certification dedicated to AI ethics for technology companies, in partnership with professional federations.

- Make access to public funding for AI research conditional on compliance with strict ethical criteria.
- Promote transparency by making it compulsory for AI systems on the market to be fully documented and for users to be able to clearly understand how they work.

B. Healthcare Sector

- Train all healthcare and administrative staff in the ethical issues specific to AI in the medical field (assisted diagnosis, confidentiality of health data, etc.).
- Involve multidisciplinary hospital ethics committees right from the design stage of AI projects in healthcare (medical imaging, intelligent patient records, etc.).
- Develop national overall guidelines for good ethical practice in the use of AI in healthcare establishments, under the aegis of the health authorities.
- Make reimbursement for medical procedures involving AI conditional on compliance with strict ethical criteria.
- Set up external and independent organizations or bodies responsible for regularly assessing AI systems to ensure that they meet high ethical standards and are used responsibly and safely in the medical sector.

C. Entertainment Sector

- Make content creators (video games, films, music, etc.) aware of the ethical risks of using Generative AI (bias, stereotypes, copyright, etc.).
- Involve citizen ethics committees in the design of entertainment experiences exploiting AI (theme parks, virtual/augmented reality, etc.).
- Define a sectoral ethical charter for the responsible use of AI in the creative and cultural industries.
- Include modules on the ethics of creative AI in training courses for artistic and entertainment professions.
- Set up an ethical evaluation and labelling system for works generated by AI (music, books, films, etc.).

D. Education Sector

- Create an AI Ethics Committee comprised of students, teachers, parents, administrators and outside experts. This committee would be responsible for advising the school or school district on ethical issues related to the use of AI and ensuring thus that AI technologies are used in a responsible and ethical manner.
- Require AI-based education technology providers to provide detailed reporting on the data used to train the algorithms, including its origin, quality and representativeness.
- Create platforms for educators and researchers to share experiences and best practices in using AI in education, thereby promoting learning and continuous improvement.
- Provide courses or workshops to students on understanding AI-based educational technologies, including how to recognize and challenge algorithmic bias.

- Establish a secure online portal where students can access their educational data collected by AI systems.
- Use crowdsourcing and collaborative learning techniques to identify and correct biases and gaps in data sets.

V. DISCUSSION

The recommendations proposed in this study to address ethical issues related to the use of AI are obviously not exhaustive. The choice made here to select the business sectors of Human Resources Management, Healthcare, Entertainment, and Education is based on a deep understanding of the social and individual impacts of AI. As mentioned before (§ Section III), these business sectors were chosen because of their significant influence on individuals' daily lives and the structure of society itself. AI technologies are playing an increasingly important role in these sectors, transforming recruitment processes, healthcare, leisure activities, and learning. By exploring them together, we shed light on the common and specific ethical risks encountered when integrating AI, such as privacy protection, algorithmic discrimination, and social justice.

Therefore, this multidisciplinary exploratory approach enables a holistic analysis of the ethical implications of the expanding current use of AI, providing a solid foundation for informed decision-making and responsible policies in the deployment of these advanced technologies.

The proposed comparative approach across these use cases has made it possible to identify both common ethical risks (bias, data privacy, lack of transparency) and sector-specific challenges. This has helped to formulate concrete recommendations, tailored to the realities of each business sector. While not exhaustive, this analysis underscores the importance of involving all stakeholders (citizens, experts, authorities) from the inception of AI systems to prevent ethical lapses. Training, awareness, transparency, and accountability are indispensable cross-cutting axis.

However, these study has limitations and can only be seen as a preliminary basis for developing an operational tool for managing the ethical risks associated with the use of AI. Moreover, delving into certain aspects such as environmental impacts or geopolitical considerations of AI is necessary to provide coherent and holistic answers, ethical use of AI being a matter for the whole of society. It will also be necessary to extend the study to different business sectors and decision levels, and to take into account the needs and specificities of each sector and level. Furthermore, as the ethical and regulatory landscape is rapidly evolving, some of the recommendations may become obsolete and have to be regularly updated.

Despite these limitations, this work subscribes with the validity of the hypotheses (*H1*) and (*H2*), and also contributes to answering the sub-question (Q') (§ Section I), by structuring reflection around concrete diversified use cases and proposing pragmatic and operational action points for a more responsible and equitable use of AI in our society. It now remains to generalize the approach for a sound identification of ethical risks, the definition of recommendations and best practices, and the integration of those, into an operational tool. In line with previous works on these methodological aspects [29], such a tool will allow a systematic handling of the ethical risks to be made. The established recommendations will then

need to be transformed into ethical, deontological or consequentialist, rules (§ Section II.A) [30].

VI. CONCLUSION AND PROSPECTS

AI and its data have the potential to positively transform many aspects of our lives, but it is important to remain aware of the ethical risks associated with it. In this sense, this study explored these risks associated with an uncontrolled use of AI in various business sectors. It is part of a more general reflection on the consideration of ethical risks in performance management. Four business sectors have been therefore selected for an initial exploratory analysis, making it possible to identify some of the major ethical risks, when using AI, and their consequences, based on the testimonies reported in the literature on the subject. An initial typology was proposed, distinguishing common ethical risks from those specific to the business sector, and was followed by a number of recommendations for managing these risks.

The approach adopted now requires to be generalized and formalized, with the aim of developing first an operational tool for systematically taking these ethical risks into account in decision-making. This development should be consistent with a deeper exploration of the existing frameworks of ethical AI usage, in particular the overall proposals of the institutions, such as that of the European Parliament, which is currently working on developing a regulatory framework for AI to ensure that it is developed and used responsibly and ethically, by including rules on AI safety, transparency, accountability and ethics.

Another prospect of this work will concern the definition, the operationalization, the implementation and the validation of the recommendations to mitigate the ethical risks, insuring their adaptability to the considered cultural and legal contexts.

REFERENCES

- [1] G. Büchi, M. Cugno and R. Castagnoli, "Smart factory performance and Industry 4.0", 2020, *Technological Forecasting and Social Change*, 150, 119790.
- [2] J. Borenstein and A. Howard, "Emerging challenges in AI and the need for AI ethics education", 2021, *AI and Ethics*, vol. 1, 1, pp.61-65.
- [3] D. L. Amoroso, "Ethical Dilemmas in Artificial Intelligence", 2022, *Journal of Business Ethics*, pp. 1-15.
- [4] F. García-Peñalvo and A. Vázquez-Ingelmo, "What Do We Mean by GenAI? A Systematic Mapping of The Evolution, Trends, and Techniques Involved in Generative AI", 2023, *International Journal of Interactive Multimedia and Artificial Intelligence*, 8, pp. 7-16.
- [5] A. Weller, "Exploring practitioners' meaning of "ethics," "compliance," and "corporate social responsibility" practices: A communities of practice perspective", 2020, *Business & Society*, 59(3), pp. 518-544.
- [6] GDPR General Data Protection Regulation, Complete Guide to GDPR Compliance, 2023, [Online]. Available: <https://gdpr.eu/>.
- [7] J.-F. Jimenez, L. Berrah, D. Trentesaux and C. Chapel, "Towards the ethical awareness integration on industrial performance management systems", *IFAC-PapersOnLine*, 55(10), pp. 3232-3237.
- [8] P. Ricoeur, "Soi-même comme un autre", 1990, *Sciences humaines*. Seuil.
- [9] I. Kant, "Groundwork for the Metaphysics of Morals", 1785, Königsberg.
- [10] L. T. Bergmann, L. Schlicht, C. Meixner, P. König, G. Pipa, S. Boshammer and A. Stephan, "Autonomous vehicles require socio-political acceptance-an empirical and philosophical perspective on the problem of moral decision making", 2018, *Frontiers in behavioral neuroscience*, vol. 12, 31.
- [11] S. Karnouskos, "Self-Driving Car Acceptance and the Role of Ethics", 2020, *IEEE Transactions on Engineering Management*, vol. 67, 2, pp. 252-265.

- [12] H. Mihai Bădău, "Les enjeux éthiques de la communication de l'information d'actualité sur les blogs et réseaux sociaux : Une proposition de charte", 2018, *Revue française des sciences de l'information et de la communication*, vol. 12, no. 2, pp. 1-17.
- [13] IEEE Advancing Technology for Humanity, "IEEE code of ethics", 2020, [Online], <https://www.ieee.org/content/dam/ieee-org/ieee/web/org/about/corporate/ieee-code-of-ethics.pdf>
- [14] European Parliament, "What is artificial intelligence and how is it used?", 2020, [Online]. Available: <https://www.europarl.europa.eu/topics/en/article/20200827STO85804/what-is-artificial-intelligence-and-how-is-it-used>
- [15] L. Berrah, D. Trentesaux and J.-F. Jimenez, "Ethics in Performance Management of Future Industrial Systems: Weak and Strong View", 2024, *Service-Oriented, Holonic and Multi-Agent Manufacturing Systems for Industry of the Future*, Proceedings of SOHOMA 2023, pp. 457-468, Springer International Publishing.
- [16] K. Siau and W. Wang, "Artificial intelligence (AI) ethics: ethics of AI and ethical AI", 2020, *Journal of Database Management (JDM)*, vol. 31, 2, pp. 74-87.
- [17] A. Tursunbayeva, C. Pagliari, S. Di Lauro and G. Antonelli, "Ethical Implications of AI Systems for Human Resource Management: A Systematic Literature Review", 2021, *Ethics and Information Technology*, pp. 1-18.
- [18] J.-J. Hathaliya and S. Tanwar, "An exhaustive survey on security and privacy issues in Healthcare 4.0", 2020, *Computer Communications*, vol. 153, pp. 311-335,
- [19] A. Rajkomar, M. Hardt, M. D. Howell, G. Corrado and M. H. Chin, "Bias and Fairness in AI Systems for Healthcare", 2022, *IEEE Reviews in Biomedical Engineering*.
- [20] H. Smith, "Clinical AI: opacity, accountability, responsibility and liability", 2021, *AI & Society*, vol. 36, pp. 535-545.
- [21] E. Meskys, J. Kalpokiene, P. Jurcys, and A. Liaudanskas, "Regulating Deep Fakes: Legal and Ethical Considerations", 2019, *Journal of Intellectual Property Law & Practice*, vol. 15, 1, pp. 24-31.
- [22] J. Bosch and A. Johnson, "A Comparative Study of Deepfake Regulatory Regimes in the United States, the European Union, and China", 2022, *Georgetown Law Technology Review*, 01.
- [23] P. Sprenger, "Tom Hanks Dental Ad: AI Version Is Fake", 2023, October 2, *The Guardian*.
- [24] B. Nyhan et al "The Threat of Deepfakes to Democracy", 2019, *Science*, Vol. 366, No. 6441 (2019), pp. 1087-1092.
- [25] *Hollywood Reporter*, February 2021.
- [26] K. R. Koedinger and A. T. Corbett, "Algorithmic Bias in Automated Assessment and Grading", 2019, *Educational Technology Research and Development*, vol. 67, No. 4, pp. 541-567.
- [27] M. B. Jones and N. F. Johnson, "The Impact of Online Learning on Social and Emotional Development", 2021, *Journal of Computer Assisted Learning*, vol. 37, No. 4, pp. 485-495.
- [28] J. DiGrazia, and C. Stewart, "Protecting Student Privacy", 2018 December 19, *The New York Times*.
- [29] D. Trentesaux, D., "Ensuring ethics of cyber-physical and human systems: a guideline", 2022, *Service Oriented, Holonic and Multi-Agent Manufacturing Systems for Industry of the Future: Proceedings of SOHOMA LATIN AMERICA 2021*, pp. 223-233, Springer International Publishing.
- [30] L. Berrah and D. Trentesaux, "Decision-Making in Future Industrial Systems: Is Ethics a New Performance Indicator?", 2021, *Service Oriented, Holonic and Multi-Agent Manufacturing Systems for Industry of the Future: Proceedings of SOHOMA 2020*, pp. 231-245, Springer International Publishing.