Int J Ethics Soc 2025; 7(1): 25-34

10.22034/ijethics.7.1.25

INTERNATIONAL JOURNAL OF Ethics & Society

Original Article

Analysis of the Obstacles to Ethical Use of Artificial Intelligence

Abbas Shoul*, Mohammad Amini, Fatemeh Arabpour
Department of Industrial Management, Faculty of Administrative Sciences and Economy, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran.

Corresponding Author: Abbas Shoul, Department of Industrial Management, Faculty of Administrative Sciences and Economy, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran. E-mail: shoul@vru.ac.ir

Received 15 Aug 2024 **Accepted** 03 Oct 2024 **Online Published** 01 May 2025

Abstract

Introduction: Given the expansion of numerous applications of artificial intelligence in the daily activities of human societies and the ethical obstacles that may arise in this direction, the need for a detailed and comprehensive examination of these issues is felt more than ever. Therefore, this study was developed with the aim of examining the challenges of ethical use of artificial intelligence technology.
Material and Methods: This study is applied in terms of purpose and descriptive-survey in terms of data collection. Library methods, searching through electronic sources, and field study were used to collect data. In the field method, a questionnaire was used. The research community is experts and specialists in the field of artificial intelligence, 15 of whom were selected through purposive sampling (snowball sampling). By reviewing the literature, 6 obstacles related to the ethical use of artificial intelligence were identified; the DEMATEL method and Excel and MATLAB software were used to analyze the data and evaluate the cause-and-effect relationships between the obstacles.
Results: The results show that among the 6 identified barriers (lack of accountability, lack of transparency, lack of legal regulations, possibility of bias and discrimination, violation of data privacy, and violation of social justice and livelihood of individuals), the barrier of lack of transparency, followed by the lack of legal regulations, are considered to be the most important and influential barriers to the ethical use of AI. Also, the barriers of data privacy violation and violation of social justice and livelihood of individuals were identified as the most influential barriers among the aforementioned barriers.
Conclusion: The findings of this study indicate the importance and causal relationships governing the existing barriers to the ethical use of AI and can help experts and managers in this field in designing AI systems and making decisions in order to comply with ethical principles.
Keywords: *Artificial Intelligence, Ethics-Based Application, Ethical Factors*

How to Cite: Shoul A, Amini M, Arabpour F. Analysis of the obstacles to ethical use of artificial intelligence, Int J Ethics Soc. 2025; 7(1): 25-34. doi: 10.22034/ijethics.7.1.25

INTRODUCTION

Artificial intelligence is an umbrella concept that draws on many disciplines, including computer science, business, engineering, biology, psychology, mathematics, statistics, logic, philosophy, and linguistics. Artificial intelligence refers to systems that exhibit intelligent behavior by analyzing their environment and taking actions with a degree of autonomy to achieve specific goals [1]. The complexity and capabilities of artificial intelligence make it unique and controversial. Artificial intelligence is rapidly expanding and has many applications. Artificial intelligence is now a part of our lives, influenced by our dependence on technology. This technology has great potential to enhance human capabilities, improve decision-making processes, and promote social well-being and productivity.

Technology companies and governments are harnessing the potential of AI in areas as diverse as healthcare, national security, self-driving cars, data analytics and predictions, marketing and advertising, inventory optimization, and the use of robots and systems in manufacturing processes.

In recent years, AI has transformed from a valuable asset that increased human productivity in various sectors to a technology that carries potentially grave threats. Its rapid evolution, once promising, has now raised widespread ethical concerns about transparency, privacy, bias, accountability, and other human rights issues; this has alarmed technology industry activists, who claim that the current rapid development of AI systems is taking place without regard or assurance of the positive impact or the safety of their applications, and poses potential risks to humanity. These concerns are a hopeful sign that organizations and their teams are at least aware of the ethical concerns surrounding the use of AI. But they also indicate a lack of enforceable regulations, laws, and punitive measures for ethical violations in this area. Unfortunately, the rapid advancement of AI technology has occurred much faster than the development of guidelines and a full understanding of ethical aspects, leading to a significant gap between the potential threats and benefits of these innovations.

Research has shown that the future belongs to organizations that actively use the benefits of AI to increase the efficiency of their activities. The successful implementation of digital technologies and automation, and consequently the success of the company, depends on resolving the ethical contradictions associated with AI [2-6]. Currently, the problems related to the development of ethical standards are acute not only at the stage of using AI systems, but also at the stage of software development. Researchers [7] believe that reliable and responsible AI systems will be crucial for the present and future of our society. Researchers [8] stated that it is naive to believe that algorithms are perfect in both implementation and design to make better decisions for all stakeholders. Researchers [9] state that algorithms are not neutral by themselves but can actually create ethical consequences. Researchers [10] identified the social impacts that AI can bring as the most important issue for societies. The use of AI is currently facing strong public criticism due to ethical risks or challenges. Studies have shown that if these risks are not addressed, the benefits that can be derived from it will remain a mirage. There are various concerns related to the ethical issues of using AI, including bias [11], violation of labor rights [12], lack of accountability [13], injustice [14], and privacy violations [4]. Since a prerequisite for addressing these issues is to clarify the key concepts involved, the following is an examination of the characteristics and factors that significantly influence the ethical use of AI.

Responsibility: One of the main concerns about the ethical use of AI is the issue of responsibility. Accountability begins before the implementation and use of the algorithm. Researchers [9] argue that developers who design decision-making algorithms also design how accountability is delegated. The challenge of accountability and responsibility can arise at different levels of AI decision-making. When an AI system fails to perform a specific task correctly, determining who should be held accountable can be challenging. This situation can lead to what is known as the problem of too many hands [15]. In the context of AI, a negative outcome can occur for a variety of reasons, including errors in the programming code, incorrect input data, incorrect operations, or other factors. The question that arises is: Who should be held accountable for the undesirable outcome? Is it the programmers who wrote the code, the person who provided the input data, or the end users

who used the AI system? Identifying the responsible party is complicated by the multiple actors involved in the use of AI systems [16].

Transparency: In the guidelines provided by the AI HLEG (European Commission High Level Expert Group on Artificial Intelligence), transparency is identified as one of seven key requirements for AI technology. Transparency is described as the ability to understand a model without providing technical details; as well as the extent to which the end user can access information that has been intentionally disclosed, which enhances their understanding [10]. AI systems have a higher degree of opacity than other technical systems. Many AI systems act as black boxes, which makes their decision-making processes hidden and opaque. For individual outputs of a particular AI system, we do not have available explanations as to why the AI produces this particular output. Similarly, it is difficult to predict the output of AI systems given new inputs; this is exacerbated by the fact that small changes to a system's inputs can have large effects on the output. Moreover, even their developers have difficulty predicting what the AI system will do and why due to the lack of transparency [15, 17]. This lack of transparency can create significant challenges for managers who need to justify decisions to stakeholders, a fundamental ethical consideration. The ambiguous nature of these systems often leaves users confused about the decision-making mechanisms. Due to their ambiguous nature, which humans are unable to interpret, AI may evolve without human supervision and guidance. Their lack of transparency raises a set of ethical and political problems about the nature of public and private decision-making that may be (intentionally or unintentionally) obscured by the use of AI systems [16, 18, 19]. Such a lack of transparency can impair the ability to trust and verify the results produced by these algorithms, raising concerns about their reliability and ethical use [10].

Regulatory regulation: The rapid growth of AI applications, coupled with their potential to significantly impact society, has led to calls for new legislation around the world [20, 21]. Regulation of AI is an issue of increasing importance, and various approaches have been proposed in the literature to define regulatory frameworks [17, 22]. Due to the lack of central international regulations and rules for the development of AI, the use of AI will be high-risk. Many of the existing guidelines for the use of AI have been created by technology companies, which calls into question their legitimacy. Therefore, it is better for these guidelines to be developed by an external authority and monitor. It is worth noting that some experts believe that companies should act independently in the application of these guidelines so as not to hinder the promotion of innovation [23], while others call for stricter regulations, especially in critical sectors such as healthcare. Thus, as this technology becomes increasingly integrated into everyday life, protecting individual rights and ensuring social welfare should be a priority.

Bias and discrimination: Bias and unfairness represent other important challenges in AI systems across domains. These issues often arise from historical data that reflect social biases and are used as the basis for AI algorithms (15, 17, 24). Some of today's technologies that leverage AI generate or reinforce bias and discrimination against some of the most vulnerable members of society, such as the poor, people with disabilities, ethnic and racial minorities, women, children, the elderly, immigrants, and people at the intersection of these identities [12, 25–28]. For example, facial recognition systems that are primarily trained on one racial or ethnic group may misidentify individuals from other groups and reinforce biases. Non-discrimination, a fundamental human right, is often violated by AI

systems [29] because they inadvertently perpetuate biases rooted in historical data. In areas such as lending, employment, and healthcare, AI models may reinforce discrimination against demographic groups and create inequalities [19, 30]. It is therefore crucial to program and train AI systems without human biases.

Data privacy: Most AI systems require large sets of data to function. With this comes the issue of data privacy. Privacy in the context of AI refers to the protection of individuals' personal information and ensuring that their sensitive information is not inappropriately used or disclosed without their consent [31]. The impact of AI on privacy can be viewed in terms of consent, transparency, and the extent to which private data is collected and accessed. Traditional approaches to privacy protection may not be able to adequately address the complexities of data analysis, which often involves processing large data sets to identify patterns, trends, and correlations. In addition, the growth of data sharing methods raises concerns about user consent and data ownership [15]. AI system development relies heavily on large amounts of data, including personal data. This data needs to be managed properly to prevent misuse and malicious use [10]. AI increasingly relies on unauthorized data collection methods, including data scraping, unauthorized collection of personal information, and illegal surveillance. This poses significant risks, from privacy breaches to data misuse and identity theft. The real impact of such breaches goes beyond data misuse; they directly affect individuals' lives and choices, raising significant ethical and social concerns. Also, with the increasing use of AI in surveillance systems, individuals' privacy could be severely compromised [12]. While privacy laws exist that aim to protect consumers and address privacy concerns, without proper privacy regulations, consumer data can be misused for political or commercial purposes without consent [23]. But these laws only apply to certain regions of the world, leaving a void for central global laws to protect privacy [32].

Social justice and livelihoods: The widespread adoption of AI technologies has challenged the ethical, sociological, and political boundaries of the human rights framework. Operations controlled by AI must be conducted in a way that respects human dignity. Attention must be paid to the human being who creates and influences the technology, not the technology itself. Therefore, focusing on preserving human dignity helps ensure that the use of AI will yield satisfactory outcomes. The expansion of AI applications has led to increased inequality between and within countries [17], poverty [28], social unrest [5], changing economic contexts [10], and business practices [21]. With the increasing use of automation and AI in various industries, there are concerns that many jobs may become obsolete, which could lead to significant job losses and violate the right to work and a livelihood with dignity [12]. While AI increases efficiency, it must be balanced with safeguarding workers' jobs and livelihoods. By the mid-2030s, one-third of all jobs could be automated, primarily affecting those with lower levels of education [17]. In addition, the development of autonomous weapons and drones has raised significant ethical concerns related to human rights. These weapons could harm innocent people and violate the right to life [12, 21]. Equitable access to AI is also crucial, so that the benefits of AI are accessible to all, regardless of their background, socioeconomic class, or geographic location [17]. Unequal access can perpetuate educational inequalities among students and affect future career prospects.

According to the research literature, barriers to the ethical use of AI were identified, which include lack of accountability (A), lack of transparency (B), lack of legal regulations (C),

possibility of bias and discrimination (D), violation of data privacy (E), and violation of social justice and livelihood of individuals (F). In this study, causal relationships between barriers to the ethical use of AI are identified and determined, so that experts and managers can observe ethical considerations in the design, implementation, deployment, and monitoring of AI systems by considering influential and important factors.

MATERIAL AND METHODS

The present study is applied research in terms of purpose and descriptive-survey in terms of data collection. In this study, library methods, electronic resource searches, and field studies were used to collect data. Since the present study is descriptive, a questionnaire was used in the field method. In this study, after identifying 6 aspects of the barriers to the ethical use of artificial intelligence, the CVR content validity index was used to assess their validity. In this study, the members of the expert panel determined the relevance of each feature based on a 3-option spectrum (essential, useful, not essential, and not useful) and then this index was measured according to the formula $CVR = \frac{N_e - N/2}{N/2}$. In this formula, N indicates the number of experts and Ne indicates the number of experts who assessed the desired option as essential. The minimum acceptable value for the CVR index is 0.79. In this study, the validity of all identified barriers was above 0.8. The statistical population of the study is experts and specialists in the field of artificial intelligence, to whom we used purposive sampling due to lack of access. Identifying obstacles and determining the causal relationships governing them requires a survey of experts and specialists in the relevant field. The expert panel consists of 15 experts and specialists in artificial intelligence and university professors with a focus on artificial intelligence, who were selected through snowball sampling. Given the expert-oriented nature of the information in the DEMATEL technique questionnaire, two university professors with expertise in the field of artificial intelligence were initially selected, and through them and the use of snowball sampling, other experts were introduced, ultimately bringing the number of expert panel members to 15. Of the 15 panel members, 11 were university faculty members and the rest were technical experts working in technology companies in the field of artificial intelligence. All panel members had at least a master's degree in related fields, and each of them had more than 5 years of work experience in this field. Then, in order to determine the level of effectiveness and impact of each of them, the relevant data was first collected through a survey among the panel members (15 people). The numerical values for the phrases "no impact" were zero, "low impact" was 1, "medium impact" was 2, "high impact" was 3, and "very high impact" was 4, with each number determining the level of impact of criterion i on j. Then, the DEMETEL technique was used according to the following steps.

Step 1: Determine the indicators / Step 2: Determine the cause-and-effect relationships by experts (direct impact). When the viewpoint of several experts is used, the arithmetic mean of the opinions is calculated and the matrix of direct relationships is formed. / Step 3: Determine the intensity of direct and indirect relationships between the indicators using the following steps:

a) Normalize the matrix of direct relationships (M matrix) according to the following equation:

$$M'_{ij} = \frac{M_{ij}}{\max\left(\sum_{j=1}^{n} M_{ij}, \sum_{i=1}^{n} M_{ij}\right)}$$

b) Calculating the total relationship matrix (T-matrix): In this step, the existing relative intensity matrix is formed from direct and indirect relationships according to the following formula.

$$T = M + M^2 + M^3 + \cdots + M^n = \frac{M(I - M^n)}{I - M} = \frac{M}{I - M} = M(I - M)^{-1}$$

c) For each of the indices, the row sum of the elements (D-i) and the column sum of the elements (R-i) are obtained.

d) For each of the indices, the sum and difference of D-i and R-i are calculated: $D_i + R_i$, $D_i - R_i$

If Di-R-i for the C index is greater than zero, then the C-i index is an effective index (its level of effectiveness is greater than its effectiveness) and if Di-R-i for the C-i index is less than zero, then the C-i index is an effective index (its level of effectiveness is greater than its effectiveness).

The test-retest method was used to measure the reliability of the DEMATEL questionnaire. To achieve this goal, the questionnaire was sent to five experts who were able to access it again, twice, three weeks apart, and the correlation value between the responses in the first and second stages was higher than 0.8. Therefore, it can be said that the reliability of the questionnaire is acceptable. The collected data was processed using the DEMATEL method in the Microsoft Excel software environment, version 2019, and the calculation of the total relationship matrix was carried out using MATLAB software.

RESULTS

In this study, first, indicators related to the obstacles to the ethical use of artificial intelligence were identified in 6 aspects: lack of accountability (A), lack of transparency (B), lack of legal regulations (C), possibility of bias and discrimination (D), violation of data privacy (E), and violation of social justice and livelihood of individuals (F). Then, using the opinions of experts and their (arithmetic) averages in the form of the initial comparison matrix in Table 1, and after normalizing it, the total relationship matrix (T) was calculated according to Table 2.

Table 1: Initial comparison matrix

	A	B	C	D	E	F	sum
A	0	3.53	3.86	2.46	2.93	2.8	15.58
B	3.47	0	3.67	3.87	3.13	2.87	17.01
C	3.8	3.93	0	2.6	3.6	2.8	16.73
D	2.6	2.27	2.4	0	2.87	4	14.14
E	2.87	2.8	2.87	3.13	0	3.93	15.6
F	2.8	2.87	2.6	2.8	3.8	0	14.87
sum	15.54	15.4	15.4	14.86	16.33	16.4	

To normalize the initial comparison matrix, it is necessary to divide all the numbers in Table 1 by max max $(\sum_{j=1}^{n} M_{ij}, \sum_{i=1}^{n} M_{ij})$ which is equal to 17.01 here. After normalization, the total relationship matrix is obtained using the relation $T = M(I - M)^{-1}$ (Table 2).

Table 2: Total Relationship Matrix (T)

	A	B	C	D	E	F
A	1.7917	1.9519	1.9641	1.8507	2.0178	2.0135
B	2.0817	1.8987	2.0755	2.0275	2.1530	2.1475
C	2.0818	2.0744	1.8859	1.9612	2.1571	2.1272
D	1.7555	1.7304	1.7344	1.5639	1.8438	1.8953
E	1.9142	1.8990	1.9004	1.8607	1.8520	2.0447
F	1.8456	1.8365	1.8247	1.7842	1.9658	1.7877

To obtain significant relationships between research variables, it is necessary to determine the threshold for the values of the total relationship matrix. Here, after the agreement of the experts, the threshold value of 2.0271 (the third quartile of the values of the total relationship matrix) was obtained. If the coefficient of two variables in the total relationship matrix is greater than the threshold, it means that the relationship between the two variables is significant. Table 3 shows the significant relationships between the variables. Cells with a value of 1 indicate a significant relationship between the two variables.

Table 4: Table 3: Significant relationships between variables

	A	B	C	D	E	F
A	0	0	0	0	0	0
B	1	0	1	1	1	1
C	1	1	0	0	1	1
D	0	0	0	0	0	0
E	0	0	0	0	0	1
F	0	0	0	0	0	0

Next, for each of the indices, the row sum of the elements (D-i) and the column sum of the elements (R-i) are obtained. Table 4 shows the horizontal sum (D) and vertical sum (R) as well as the calculation of D+R and D-R.

Table 4: Calculations of D and R

	A	B	C	D	E	F
D_i	11.5897	12.3839	12.2876	10.5233	11.471	11.0445
R_i	11.4705	11.3909	11.385	11.0482	11.9895	12.0159
$D_i + R_i$	23.0602	23.7748	23.6726	21.5715	23.4605	23.0604
$D_i - R_i$	0.1192	0.993	0.9026	-0.5249	-0.5185	-0.9714

According to the values in Table (4) and based on the opinions of experts, the relationships between the 6 identified variables were plotted in figure (1). In addition, this chart represents the causal relationships between the variables.

According to the values in Table 4, the variable "lack of transparency" (B) is considered the most influential variable. Also, the variables "lack of legal regulations" (C) and "lack of accountability" (A) are in the next priorities in terms of impact (row D_i in Table 4). The variables "violation of social justice and livelihood of individuals" (F) and "violation of data privacy" (E) are considered the most influential variables, respectively (row R_i in Table 4). As can be seen in Figure 1, the variables "lack of transparency" (B) and "lack of legal regulations" (C) are considered important variables because the sum of the impact and effectiveness of these variables is greater than other variables ($D_i + R_i$). These two variables have the greatest interaction and relationship with other variables. It is necessary to explain that the row (D_i-R_i) shows the intensity of the impact of a component. If its value is positive for a component; The component has an influential nature and if it is negative, it will have an affected nature. Accordingly, among the six components examined, three components "lack of

accountability" (A), "lack of transparency" (B) and "lack of legal regulations" (C) have an influential nature. In contrast, the impact of the three variables "possibility of bias and discrimination" (D), "violation of data privacy" (E) and "violation of social justice and livelihood of individuals" (F) is less than their impact (D_i-R_i), so they are among the affected variables and are the effects of other components.

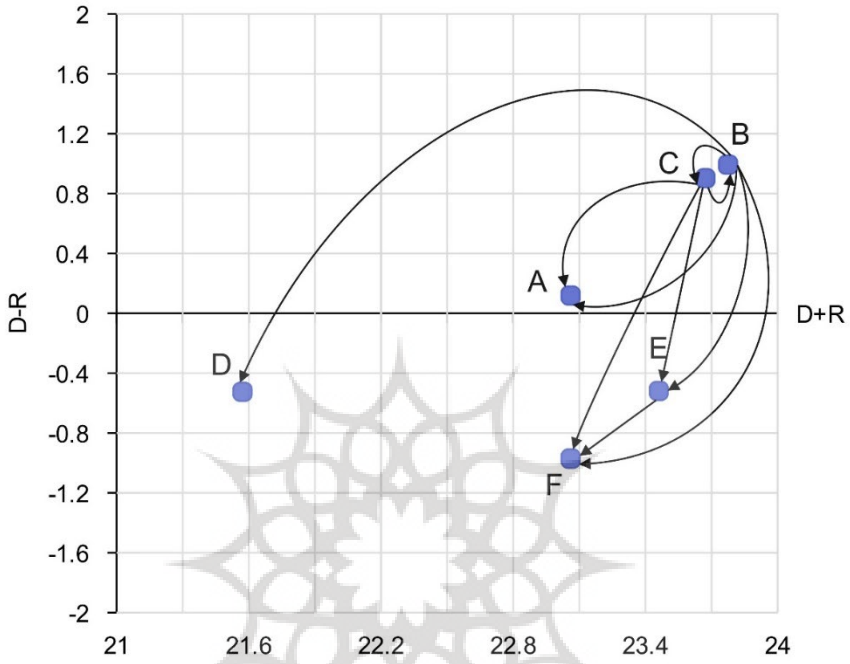

Figure 1: Degree of impact and influence of identified barriers to ethical use of AI

DISCUSSION

In this study, based on a literature review and discussions with AI technology experts, six variables were identified. The findings showed that the variable of lack of transparency is the most influential variable and has a strong impact on creating and reinforcing other barriers. When AI algorithms operate without transparency, they may make unfair decisions that benefit a particular group and ignore the rights of others. These results are consistent with the findings of studies [1, 17, 29]. The potentially morally reinforcing nature of transparency (as an enabler of other moral principles) may partly explain its high importance [5]. Next, the barrier of lack of legal regulations has a strong impact on other barriers. This factor is consistent with studies [12, 21] that state that due to the lack of legal regulations, grounds for discrimination, data privacy violations, and violations of social justice of individuals have been created. Emphasizing the existence of a legal vacuum is a precautionary measure to ensure that AI developers and deployers are under control [32]. Violation of social justice and livelihoods is another obstacle examined in this study. This factor was identified as the most influential obstacle and is most affected by the obstacles of lack of transparency and lack of legal regulations, which is in line with research [19, 29]. This study can help researchers and activists in this field to better understand the existing challenges. Examining the causal relationships between different obstacles can help to better understand how these obstacles affect

each other. This analysis can show how one obstacle can lead to the emergence or exacerbation of other obstacles. Accordingly, it is necessary for experts to focus more on the influential variables (lack of transparency and lack of legal regulations). Since these variables have a significant effect on other variables, eliminating these variables will lead to the elimination of other obstacles. It is essential that designers of AI systems receive the necessary training to build AI ethically, and that potential AI users are also prepared to deal with this technology ethically. This can help improve public trust in AI. The present study can help increase public awareness and education about the importance of ethical use of AI. This can lead to changes in attitudes and behaviors at the community level. The results of the study can help policymakers and legislators to develop appropriate laws and regulations for the ethical use of AI. These laws can help protect the rights and privacy of individuals. This study can also serve as a valuable resource for researchers in the field of AI and ethics and pave the way for further research in this area. The findings of this study provide an idea of the importance of barriers to ethical use of AI, which may help AI professionals design AI systems based on ethical principles. Although the results highlight the importance of some ethical principles, our review revealed that there are conflicts between some ethical principles. For example, addressing bias requires granting more control over personal data, which conflicts with the principle of respect for privacy. Therefore, it is suggested that risk-benefit assessments be considered in future research. The present study, like other studies, has limitations. The snowball sampling method can lead to sampling bias, as the selected individuals may tend to represent individuals with similar views and experiences. This can lead to a reduction in the diversity of views and opinions. Despite its many advantages, the DEMATEL technique has limitations. For example, the results are highly dependent on expert judgments and may be influenced by individual biases. Also, setting a threshold for relationships can lead to different results. The opinions and perspectives of experts may be influenced by the specific cultural and geographical contexts studied in the research. This can limit the generalizability of the results to other cultures and regions.

CONCLUSION

Despite the great potential of AI technology to improve human lives in numerous ways, there are also growing concerns about the risks and hazards that such systems pose as they become increasingly embedded in the daily activities of human societies, highlighting the importance of understanding their drawbacks and how to deal with them. Furthermore, our limited ability to guide and predict the outcomes of AI systems exacerbates their threats. The results suggest that by promoting transparency, accountability, and establishing legal regulations in the development of AI, we can ensure that the impact of AI is ethical and protects the rights and well-being of all individuals.

ETHICAL CONSIDERATIONS

Ethical issues (such as plagiarism, conscious satisfaction, misleading, making and or forging data, publishing or sending to two places, redundancy and etc.) have been fully considered by the writers.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interests.

REFERENCES

1. Farisco M, Evers K, Salles A. Towards establishing criteria for the ethical analysis of artificial intelligence. Science and Engineering Ethics, 2020; 26(5): 2413-25. Doi: 10.1007/s11948-020-00238-w
2. Ramezani M, Feizi-Derakhshi MR. Machine ethics: ethical challenges and strategies in artificial intelligent and

superintelligence. Ethics in Science and Technology, 2014; 8(4): 35-43. (In Persian).
3. Mohamed S, Png M-T, Isaac W. Decolonial AI: decolonial theory as sociotechnical foresight in artificial intelligence. Philosophy & Technology, 2020; 33(4): 659-84. Doi: 10.1007/s13347-020-00405-8
4. Roberts H, Cowls J, Morley J, Taddeo M, Wang V, Floridi L. The Chinese approach to artificial intelligence: an analysis of policy, ethics, and regulation. Ai & Society, 2020; 36(1): 59-77. Doi: 10.1007/s00146-020-00992-2
5. Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. Nature Machine Intelligence, 2019; 1(9): 389-99. Doi: 10.1038/s42256-019-0088-2
6. Ganege S, Laksiri A, Liyanage A. Harmony of science, technology and ethics. Int. J. Ethics Soc, 2024; 6(1): 2. Doi: 10.22034/ijethics.6.1.10
7. Díaz-Rodríguez N, Del Ser J, Coeckelbergh M, de Prado ML, Herrera-Viedma E, Herrera F. Connecting the dots in trustworthy Artificial Intelligence: From AI principles, ethics, and key requirements to responsible AI systems and regulation. Information Fusion, 2023; 99: 101896. Doi: 10.1016/j.inffus.2023.101896
8. Marabelli M, Newell S, Handunge V. The lifecycle of algorithmic decision-making systems: Organizational choices and ethical challenges. The Journal of Strategic Information Systems, 2021; 30(3). Doi: 10.1016/j.jsis.2021.101683
9. Martin K. Ethical implications and accountability of algorithms. Journal of Business Ethics. 2018; 160(4): 835-50. Doi: 10.1007/s10551-018-3921-3
10. Stahl BC, Brooks L, Hatzakis T, Santiago N, Wright D. Exploring ethics and human rights in artificial intelligence–A Delphi study. Technological Forecasting and Social Change, 2023; 191: 122502. Doi: 10.1016/j.techfore.2023.122502
11. Murikah W, Nthenge JK, Musyoka FM. Bias and ethics of AI systems applied in auditing-a systematic review. Scientific African, 2024; 25(5): e02281. Doi: 10.1016/j.sciaf.2024.e02281
12. Polok B, el-Taj H, Rana AA. Balancing potential and peril: the ethical implications of artificial intelligence on human rights. Multicultural Education, 2023; 9(6). Doi: 10.2139/ssrn.4484386
13. Vinuesa R, Azizpour H, Leite I, Balaam M, Dignum V, Domisch S, et al. The role of artificial intelligence in achieving the Sustainable Development Goals. Nature Communications, 2020; 11(1): 233.
14. Barredo Arrieta A, Díaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 2020; 58(3): 82-115. Doi: 10.1016/j.inffus.2019.12.012
15. Barbosa F, Valadares L. Artificial intelligence: a claim for strict liability for human rights violation. Revista de Direito Internacional, 2023; 20(2).
16. Bathaee Y. The artificial intelligence black box and the failure of intent and causation. Harvard Journal of Law & Technology, 2017; 31: 889.
17. Alahmed Y, Abadla R, Ameen N, Shteiwi A. Bridging the gap between ethical AI implementations. International Journal of Membrane Science and Technology, 2023; 10(3): 3034-46. Doi: 10.15379/ijmst.v10i3.2953
18. Andrada G, Clowes RW, Smart PR. Varieties of transparency: exploring agency within AI systems. AI & Society, 2022; 1-11. Doi: 10.1007/s00146-021-01326-6
19. Qadir J, Islam MQ, Al-Fuqaha A. Toward accountable human-centered AI: rationale and promising directions. Journal of Information, Communication and Ethics in Society, 2022; 20(2): 329-42. Doi: 10.1108/JICES-06-2021-0059
20. Tzimas T. The need for an international treaty for AI from the perspective of human rights. Scientia Moralitas-International Journal of Multidisciplinary Research, 2019; 4(1): 73-91.
21. Bakiner O. The promises and challenges of addressing artificial intelligence with human rights. Big Data & Society, 2023; 10(2). Doi: 10.1177/20539517231205476
22. Zhang B, Anderljung M, Kahn L, Dreksler N, Horowitz MC, Dafoe A. Ethics and governance of artificial intelligence: Evidence from a survey of machine learning researchers. Journal of Artificial Intelligence Research, 2021; 71: 591-666. Doi: 10.1613/jair.1.12895
23. Buiten M, De Streel A, Peitz M. The law and economics of AI liability. Computer Law & Security Review, 2023; 48: 105794. Doi: 10.1016/j.clsr.2023.105794
24. Talha Khalil A, Shinwari Z K. The dual nature of ai; a double-edged sword amid digital transformation; an ethical analysis. Int. J. Ethics Soc, 2024; 6(2): 1-8. Doi: 10.22034/ijethics.6.2.1
25. Larson BN. Gender as a variable in natural-language processing: Ethical considerations. Association for Computational Linguistics, 2017; 1-11. Doi: 10.18653/v1/W17-1601
26. Koolen C, van Cranenburgh A. These are not the stereotypes you are looking for: bias and fairness in authorial gender attribution. In Proceedings of the First Ethics in NLP workshop. Association for Computational Linguistics (ACL), 2017; 12-22. Doi: 10.18653/v1/W17-1602
27. O'neil C. Weapons of math destruction: How big data increases inequality and threatens democracy. Crown. 2017.
28. Eubanks V. Automating inequality: how high-tech tools profile, police, and punish the poor. New York: Picador, St Martin's Press. Law, Technology and Humans, 2018; 162-164. Doi: 10.5204/lthj.v1i0.1386
29. Buijsman S, Klenk M, van den Hoven J. Ethics of artificial intelligence. Cambridge Handbook on the Law, Ethics and Policy of AI. Cambridge University Press. 2023.
30. Tat E, Bhatt DL, Rabbat MG. Addressing bias: artificial intelligence in cardiovascular medicine. The Lancet Digital Health, 2020; 2(12): e635-6. Doi: 10.1016/S2589-7500(20)30249-1
31. Bartneck C, Lütge C, Wagner A, Welsh S. Introduction to ethics in robotics and AI. Springer Nature. 2021. Doi: 10.1007/978-3-030-51110-4
32. Zhang N, Wang L, Wilson I. Explaining the hacking of society's information systems from the point of view of ethical theories. Int. J. Ethics Soc, 2024; 6(2): 9-19. Doi: 10.22034/ijethics.6.2.9
33. Benkler Y. Don't let industry write the rules for AI. Nature. 2019; 161-161. Doi: 10.1038/d41586-019-01413-1