Page 3852 - 3860 **DOI:** *https://doij.org/10.10000/IJLMH.1110301*

INTERNATIONAL JOURNAL OF LAW MANAGEMENT & HUMANITIES

[ISSN 2581-5369]

Volume 8 | Issue 3

2025

Artificial Intelligence and Ethics

CHAITY[1]

ABSTRACT

This paper explores the ethical challenges and responsibilities that arise with the growing use of Artificial Intelligence (AI), especially in the legal field. AI has the power to mimic human intelligence and is now being used to make legal decisions, but its use raises important concerns around fairness, accountability, and transparency. These concerns are grouped under the concept of FATE, which stands for 'Fairness, Accountability, and Transparency in Ethics'. The paper discusses how faulty data can bring about to biased results, how it is difficult to hold anyone accountable when AI systems make mistakes, and how many AI systems lack the transparency needed for people to understand how decisions are made. Various countries and global institutions have proposed ethical guidelines to manage AI use responsibly, including India's NITI Aayog, the European Union, UNESCO, the UAE, and China. However, most of these guidelines are non-binding and symbolic. Lastly, the paper suggests a shift toward Human-Centered AI (HCAI), which focuses on supporting human values, rights, and dignity to ensure the development of a truly trustworthy AI.

Keywords: *Artificial Intelligence, Ethics, Fairness, Accountability and Transparency.*

I. INTRODUCTION

Artificial Intelligence is transforming the world at an unprecedented pace, reshaping industries, services, daily life and societal structures with its ability to mimic human intelligence through learning, reasoning and adaptation all the while challenging traditional paradigms of human capability, effortlessly. As is seen in almost every field, AI is increasingly being incorporated into the legal arena as well, due to its ability to process information, recognise pattern and apply reasoning to reach a conclusion, all at lightning speed and accuracy.

The intersection of AI and law also raises significant ethical and regulatory concerns. AI may adopt from biased data and code, leading to unfair legal outcomes. The lack of transparency in AI decision-making challenges accountability and raises questions about trust in automated legal processes, making it difficult to justify AI-driven recommendations in a legal context. The lack of transparency and comprehensibility in AI systems is a major barrier. The

[1] Author is a Student at Hidayatullah National Law University, Raipur, India.

 [ISSN 2581-5369]

reliability of AI-generated results is called into doubt since advanced algorithms and other 'black-box' models frequently produce precise forecasts without offering concise justifications for their conclusions.

Over the years, AI powered systems have become more well-known because of their enormous potential to generate economic value and lessen burden by automating repetitive tasks. While there are many advantages to the growing application of AI, there are also important ethical issues that need to be resolved. Critical challenges that need constant attention include bias, accountability, openness, privacy, etc. In the past, technologies like "nuclear power, cars, or plastics, have caused ethical and political discussion, usually only when some damage is done."[2] The same must be avoided when it comes to AI and its widespread adoption, so as to avoid any harm at all. Naturally, the aspect of ethics and AI needs to be explored due to its increasing capacities and its seeming competence at tasks formerly restricted to the human realm, which raises significant questions about the impact this technology has on the justice system.

II. FATE OF AI

Ethics are important in the field of AI given its ever increasing influence in multiple domains. Ethical principles help "discern between right and wrong, and how to optimize the beneficial impact of AI while reducing risks and adverse outcomes".[3] Ethical factors are crucial in order to ensure that AI tools are 'transparent, equitable and safe'. Consequently, the concept of 'FATE' in AI technology has emerged which stands for *Fairness, Accountability & Transparency in Ethics*, where "*fairness* represents the quality or state of being fair, signifying impartiality and honesty, conforming to established rules, and being of average and acceptable quality". The term "*accountability* represents the quality or state of being accountable, signifies being answerable and explainable". Lastly, "*transparency* represent the quality or state of being transparent, signifying being readily understood and having the accessibility of information".[4] These are further discussed in detail below.

FAIRNESS: fairness in AI refers to "the absence of bias or discrimination in AI systems, which can be challenging to achieve due to the different types of bias that can arise in these

[2] *Ethics of Artificial Intelligence and Robotics*, STANFORD ENCYCLOPAEDIA OF PHILOSOPHY, (Mar. 17, 2025), https://plato.stanford.edu/entries/ethics-ai/

[3] *What Is AI Ethics?*, IBM (Mar. 17, 2025), https://www.ibm.com/think/topics/ai-ethics.

[4] Bahar Memarian & Tenzin Doleck, Fairness, Accountability, Transparency, and Ethics (FATE) in Artificial Intelligence (AI) and higher education: A systematic review, 5 COMPUTERS & EDUCATION: ARTIFICIAL INTELLIGENCE, 6 (2023).

 [ISSN 2581-5369]

systems".[5] It is a critical issue because AI systems often reflect and amplify existing societal biases and discrimination. Although bias and fairness are inter-related concepts, they standout in concrete ways, such as the fact that bias may be inadvertent but fairness is usually a purposeful and thoughtful objective, consciously aimed for. Bias occurs "when an algorithm's output consistently and systematically deviates from the true value or from what would be predicted in the absence of bias". In contrast, fairness is the lack of prejudice or favouritism against any person or group because of some traits, such as gender, age, race, or religion.

Bias can stem from the data used to train AI models, which may be imbalanced, incomplete, or historically prejudiced. If an AI system is fed dataset that is biased, it may produce prejudiced results, excessively disadvantaging some sections of society. For example, bias and prejudice can be seen in USA's COMPAS system, where "facial recognition technology has been shown to have higher error rates for people with darker skin tones due to a lack of diverse training data", reflecting a systemic bias against a racial minority.[6] These "biased AI systems have the potential to restrict personal liberties and strengthen social power structures" and the effect of these systems on human agency and autonomy must also be taken into account. For instance, a system that employs AI in the process of recruiting can unfairly reject applicants from underrepresented groups, making it harder for them to find work and make a positive contribution to society. Similarly, hiring algorithms trained on past hiring decisions may favour one demographic over another, reinforcing discrimination rather than eliminating it.

To guarantee that these systems are just and equal for all users, bias in AI must be recognised and addressed. In addition to the AI system itself, the people who develop and implement it must also be held accountable if the system exhibits prejudice and generates discriminatory results. Careful consideration of the datasets utilised to train AI systems and frequent checks to identify and reduce bias are necessary to ensure fairness. Debiasing of training datasets can be done to ensure that it is representative of maximum possible population, specially marginalised and minority groups. This is essential to reduce the possibility of biased results and to promote a wider range of perspectives leading to the development of a system which can stand the scrutiny that ensures fairness.[7]

[5] Dr. A. Sreelatha & Dr. Gyandeep Choudhary, Exploring The Use of AI In Legal Decision Making: Benefits and Ethical Implications, WOXEN UNIVERSITY, (Mar. 17, 2025), https://woxsen.edu.in/research/white-papers/exploring-the-use-of-ai-in-legal-decision-making-benefits-and-ethical-implications/.
[6] Matthias Spielkamp, *Inspecting Algorithms For Bias*, MIT TECHNOLOGY REVIEW, (Mar. 16, 2025), https://www.technologyreview.com/inspecting-algorithms-for-bias/.
[7] SREELATHA, *supra* note 4.

 [ISSN 2581-5369]

ACCOUNTABILITY: accountability, in the general sense stands for the "ability to determine whether a decision was made in accordance with procedural and substantive standards and to hold someone responsible if those standards are not met".[8] In the context of ethics, it refers to the responsibility and liability associated with the use of AI in legal decision-making, advisory roles, and judicial processes. It involves determining who is responsible when AI systems make errors, cause harm, or produces unfair outcomes. This is important as AI is becoming more autonomous and capable of decision-making, raising the queries about "who should be held responsible when AI makes an error or causes legal harm?".[9]

The question of accountability has led some legal scholars and policymakers to wonder if AI should be recognised as a distinct legal entity to fix responsibility, which entails granting it a legal status similar to that of natural persons (humans) or legal entities (such as corporations).[10] This would mean AI could have rights, responsibilities, and potentially even liabilities under the law. With that in mind, the Parliament of Europe endorsed a "resolution calling on its Commission to consider creating a specific legal status for robots in the long run, established the status of electronic persons responsible for making good any damage they may cause".[11] The Parliament, in the year 2017, "issued a report that recommended setting up a status of 'electronic person' for highly sophisticated and autonomous robots".[12] However, there are still a lot of ambiguities, like whether this strategy would not interfere with society's ability to function and how much it might be used by different groups to escape responsibility. Additionally, the question arises as to how accountability be placed since there is "no soul to be damned, and no body to be kicked".[13]

TRANSPARENCY: transparency involves "making the AI system's decision-making process visible to users".[14] It helps understand how an AI system makes decisions and is commonly known as *Explainability*.[15] Most, complex AI models, such as deep learning algorithms, do not offer transparency or explain how they reached a conclusion, function as 'black boxes' i.e. making precise predictions without offering concise justifications. This lack of openness calls into doubt the reliability and responsibility of AI-generated results in the legal field. Here,

[8] Finale Doshi-Velez et al., *Accountability of AI Under the Law: The Role of Explanation*, BERKMAN KLEIN CENTER WORKING GROUP ON AI INTERPRETABILITY (2017).
[9] WEINBERGER, *supra* note 7.
[10] Fatima Rizq Moustafa, *Towards Recognition of the Legal Personality of Artificial Intelligence (AI): Recognizing Reality and Law*, Vol. 19(1) IJCJS 271, 274 (2024).
[11] Simon Chesterman, *Artificial Intelligence and the Limits of Legal Personality,* Vol. 69 ICLQ 819, 821 (2020).
[12] FATIMA, *supra* note 9.
[13] SIMON, *supra* note 10.
[14] Emilio Ferrara, *Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies*, Vol. 6(3) SCI, (2024).
[15] *What is AI Transparency?*, IBM, (Mar. 20, 2025), https://www.ibm.com/think/topics/ai-transparency.

 [ISSN 2581-5369]

"explainability becomes crucial in the legal system as it enables legal professionals, litigants, and the public to understand and evaluate the reasoning behind AI-generated decisions".[16]

Merely knowing the result, without any explanation as to how it was arrived at leads to dissatisfaction, which could hamper public's trust and raise questions about biasness in the reasoning of AI. It is critical to comprehend at least the key elements that contributed to the output before making a decision based on it. This can be fixed by utilising AI models that are explainable in nature i.e. *White Box* models that offer comprehensible justifications for their forecasts. By adding easily understandable explanations, actions can be justified in a clear and understandable manner, increasing openness and allowing legal experts to assess if AI-generated results are suitable.

Together, *Fairness, Accountability and Transparency* form the backbone of ethical principles and aims to set standards relating to AI, ensuring that "the use of AI should not violate human autonomy; that AI's benefits should outweigh its risks; and AI's benefits and risks should be distributed equally and fairly".[17]

III. STANDARDS OF ETHICS FOR AI

Recognizing the importance of ethics for AI, various standards have been set by different countries and global institutions, guiding the development and deployment of AI along the lines of moral values and principles that are aligned with human rights and give primacy to human well-being. Some of these standards and guidelines are briefly discussed below.

INDIA: NITI Aayog has provided a set of 7 "Principles for Responsible Management of Artificial Intelligence" in India, which includes the "Principles of Safety and Reliability, Equality, Inclusivity and Non-discrimination, Privacy and security, Transparency, Accountability and lastly, protection and reinforcement of positive human values".[18] It prescribes a discretionary approach, which offers an "enabling environment for fostering a responsible AI ecosystem in India" by recommending a strategy, manner, and degree of principle implementation. In order to stay up to date with technological advancements, the measures can be appropriately adjusted based on the unique risk associated with various AI applications.

EUROPE: The "European Union's Expert Group on AI presented Ethics Guidelines for Trustworthy Artificial Intelligence" in the year 2019, where it suggested that a 'trustworthy

[16] SREELATHA, *supra* note 4.
[17] BAHAR, *supra* note 3.
[18] *Responsible AI*, NITI AAYOG, (Mar. 20, 2025), https://www.niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf.

 [ISSN 2581-5369]

AI' should be "lawful and respect all applicable laws and regulations, ethical and respecting ethical principles and values and technically robust both from a technical perspective while taking into account its social environment". They also laid down 7 additional requirements the AI systems must fulfill in order to be "deemed trustworthy, which includes human oversight, technical robustness, privacy and data governance, transparency, fairness, well-being, and accountability".[19]

UNESCO: The UN body came out with the Recommendation on the Ethics of Artificial Intelligence, in the year 2021, which aims to take an approach to AI which puts human at its center, where its values and principles ensure the "respect, protection and promotion of human rights, fundamental freedoms and human dignity".[20] It also provides 11 keys policy action areas where practical application of the values and principles are recommended.

UNITED ARAB EMIRATES: The UAE, in the year 2022, has introduced a set of 8 guidelines under its AI Ethics: Principles & Guidelines, which are geared towards "achieving the ethical design and deployment of AI systems in both public and private sectors". The principles include "explainable, transparent, accountable, fair, privacy preserving, sustainable and environmentally friendly, human centered and lastly, robust, safe and secure".[21]

CHINA: The National Governance Committee for the New Generation Artificial Intelligence, in the year 2021, released the Ethical Norms for New-Generation Artificial Intelligence, which "aims to integrate ethics into the entire lifecycle of AI, to provide ethical guidelines for natural persons, legal persons, and other related organisations engaged in AI-related activities".[22] The ethical norms include "Enhancing the well-being of humankind, Promoting fairness and justice, Protecting privacy and security, Ensuring controllability and trustworthiness, Strengthening accountability and Improving ethical literacy".[23]

While many ethical standards and principles governing AI exist, it must be noted that almost all of these are merely recommendations with no real sanction to ensure conformity. Given the non-binding nature of these guidelines, they fail to make any significant changes in the way AI is developed and deployed, remaining mere symbolic policy gesture.

---

[19] *Ethics guidelines for trustworthy AI*, EUROPEAN COMMISSION, (Mar. 20, 2025), https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.
[20] *Recommendation on the Ethics of Artificial Intelligence*, UNESCO DIGITAL LIBRARY, (Mar. 20, 2025), https://unesdoc.unesco.org/ark:/48223/pf0000381137.
[21] *AI Ethics: Principles & Guidelines,* UAE MINISTER OF STATE FOR ARTIFICIAL INTELLIGENCE, (Mar. 20, 2025), https://ai.gov.ae/wp-content/uploads/2023/05/MOCAI-AI-Ethics-EN.pdf.
[22] *The Ethical Norms for the New Generation Artificial Intelligence*, INTERNATIONAL RESEARCH CENTER FOR AI ETHICS AND GOVERNANCE, (Mar. 21, 2025), https://ai-ethics-and-governance.institute/2021/09/27/the-ethical-norms-for-the-new-generation-artificial-intelligence-china/.
[23] ETHICAL NORMS, *supra* note 21.

 [ISSN 2581-5369]

IV. WAY FORWARD

Development of a trustworthy AI warrants a shift in the way this technology is designed i.e. with productivity in mind, which is a fair consideration. However, a human-centric element must also be added in the mix to ensure its optimisation on a wider scale. This can be done by adopting a Human Centered AI (HCAI) approach which priorities human needs, values, and well-being in the design, development, and deployment of AI systems. It focuses on a 'human+AI' approach, creating a collaborative environment that enhances human capabilities, promotes fairness, and aligns with ethical principles, rather than simply automating tasks or replacing human involvement.

In essence, the values that are expected from AI systems must follow the humanistic approach and also adhere to the vision of the Supreme Court of India, which has time and again reiterated the "prevailing morality of our country to be based on the principle of Constitutional morality". Stressing on adherence to "constitutional morality over social morality", which encompasses the "values of a diverse and inclusive society while remaining faithful to other constitutional principles".[24] Recognizing the need for responsible development and use of these technologies, with a robust ethical foundation to support it, AI systems hold the capacity to be the disruptive change it is visioned to be, aiding in the identification, understanding, reasoning and response, associated with the various legal process, revolutionising the ideas of justice.

*****

[24] NITI AAYOG, *supra* note 17.

 [ISSN 2581-5369]

V. REFERENCES

1. *Ethics of Artificial Intelligence and Robotics*, STANFORD ENCYCLOPAEDIA OF PHILOSOPHY, (Mar. 17, 2025), https://plato.stanford.edu/entries/ethics-ai/
2. *What Is AI Ethics?*, IBM (Mar. 17, 2025), https://www.ibm.com/think/topics/ai-ethics.
3. Bahar Memarian & Tenzin Doleck, Fairness, Accountability, Transparency, and Ethics (FATE) in Artificial Intelligence (AI) and higher education: A systematic review, 5 COMPUTERS & EDUCATION: ARTIFICIAL INTELLIGENCE, 6 (2023).
4. Dr. A. Sreelatha & Dr. Gyandeep Choudhary, Exploring The Use of AI In Legal Decision Making: Benefits and Ethical Implications, WOXEN UNIVERSITY, (Mar. 17, 2025), https://woxsen.edu.in/research/white-papers/exploring-the-use-of-ai-in-legal-decision-making-benefits-and-ethical-implications/.
5. Matthias Spielkamp, *Inspecting Algorithms For Bias*, MIT TECHNOLOGY REVIEW, (Mar. 16, 2025), https://www.technologyreview.com/inspecting-algorithms-for-bias/.
6. Finale Doshi-Velez et al., *Accountability of AI Under the Law: The Role of Explanation*, BERKMAN KLEIN CENTER WORKING GROUP ON AI INTERPRETABILITY (2017).
7. Fatima Rizq Moustafa, *Towards Recognition of the Legal Personality of Artificial Intelligence (AI): Recognizing Reality and Law*, Vol. 19(1) IJCJS 271, 274 (2024).
8. Simon Chesterman, *Artificial Intelligence and the Limits of Legal Personality,* Vol. 69 ICLQ 819, 821 (2020).
9. Emilio Ferrara, *Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies*, Vol. 6(3) SCI, (2024).
10. *What is AI Transparency?*, IBM, (Mar. 20, 2025), https://www.ibm.com/think/topics/ai-transparency.
11. *Responsible AI*, NITI AAYOG, (Mar. 20, 2025), https://www.niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf.
12. *Ethics guidelines for trustworthy AI*, EUROPEAN COMMISSION, (Mar. 20, 2025), https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.
13. *Recommendation on the Ethics of Artificial Intelligence*, UNESCO DIGITAL LIBRARY, (Mar. 20, 2025), https://unesdoc.unesco.org/ark:/48223/pf0000381137.

 [ISSN 2581-5369]

14. *AI Ethics: Principles & Guidelines,* UAE MINISTER OF STATE FOR ARTIFICIAL INTELLIGENCE, (Mar. 20, 2025), https://ai.gov.ae/wp-content/uploads/2023/05/MOCAI-AI-Ethics-EN.pdf.

15. *The Ethical Norms for the New Generation Artificial Intelligence*, INTERNATIONAL RESEARCH CENTER FOR AI ETHICS AND GOVERNANCE, (Mar. 21, 2025), https://ai-ethics-and-governance.institute/2021/09/27/the-ethical-norms-for-the-new-generation-artificial-intelligence-china/.

*****

 [ISSN 2581-5369]