Journal of Legal Studies, International Refereed Peer-Reviewed Journal
ISSN 2321-1059, Vol. 13, Issue I, January 2025
www.journaloflegalstudies.co.in

Bias and Fairness in AI-Driven Legal Systems: Ethical and Legal Considerations

Kolawole Afuwape[1]
&
Olanrewaju Oladokun Adeojo[2]

Abstract

The use of artificial intelligence (AI) in the legal system has the capacity to improve efficiency, consistency, and accessibility. There are, however, issues around bias and fairness in AI-led legal decision-making that pose profound ethical and legal issues. The research discusses the way biases in training data, algorithmic structure, and human intervention affect the fairness of AI-based legal systems. It investigates issues such as discriminatory outcomes, non-transparency, and accountability in automated decision-making. Through case studies of AI use in legal settings, such as predictive policing, sentencing suggestions, and automated contract review, the study points to cases where bias has resulted in unfair legal outcomes. It also examines regulatory protocols and ethical standards intended to reduce such bias while maintaining adherence to core legal principles such as due process and non-discrimination. The study also assesses the function of human oversight in reducing AI bias, considering the efficacy of interventions like algorithmic auditing, explainability mandates, and fairness-promoting design methods. It addresses the conflict between the efficiency of AI-based systems and the necessity of judicial discretion, highlighting the significance of preserving ethical protection in automated legal decision-making. Ultimately, this study is a contribution to current discussions of responsible AI design and legal ethics, offering policy suggestions to ensure fairness, accountability, and trust in AI-powered legal systems. It emphasizes the need for cross-disciplinary work among technologists, legal experts, and policymakers to develop AI frameworks that protect justice, equality, and fundamental rights in the changing context of legal automation.

Keywords: *AI bias; legal ethics; fairness in AI; algorithmic accountability; automated decision-making*

Introduction

The swift convergence of AI in various fields, ranging from transportation to health and finance, has revolutionized decision-making processes and efficiency of operations.[3] Nonetheless, this dependency on AI raises new legal challenges, especially accountability and sanctions. AI embodies attributes that relate to human intelligence and executes it via numerical systems.[4] These can form, learn, predict, analyze, draw conclusions, and even correct themselves. It is implemented to solve an array of medicine predictions, planning, imagining pictures, voice

[1] ***Lecturer, Jindal Global Law School, O.P. Jindal Global University, Sonipat, Haryana, India, kolawole.afuwape@jgu.edu.in***
[2] ***Associate Professor, Department of Private Law, Olabisi Onabanjo University, Ago-Iwoye, Ogun State, Nigeria, adeojo.lanre@oouagoiwoye.edu.ng***

[3] Li, Y. H., Li, Y. L., Wei, M. Y., & Li, G. Y. (2024). Innovation and challenges of artificial intelligence technology in personalized healthcare. *Scientific reports*, *14*(1), 18994.

[4] Sarker, I. H. (2022). AI-based modeling: techniques, applications and research issues towards automation, intelligent and smart systems. *SN computer science*, *3*(2), 158.

recognition and learning a specific aspect. AI systems apply a database for training to forecast better results and help in the resolution of other problems at high precision.[5]

AI is all about algorithms and data. Decision-making capability of AI is accurate only based on how well it has been trained and by utilizing authentic and unbiased data.[6] Unethical and unfair repercussions are built into critical decision-making if the data employed to train is infused with racial, gender, communal, or ethnic biases.[7] These will likely be more exacerbated, since most AI systems will still be trained on bad data.

Human beings outsource a lot of our choices to AI algorithms. We consume services that, with the use of AI algorithms, impact our choices. Public authorities and corporations also increasingly depend upon algorithms to make long-term choices.[8] For instance, they are relying on AI algorithms to analyze and forecast crime. The main mechanism in an algorithm is the spread of recommendations and decisions based on data analysis.[9] Therefore, algorithms are increasingly dictating our personal and collective decisions. The word algorithm has various interpretations based on diverse viewpoints.[10]

This algorithm usually applies to a computer's collection of rules and instructions for responding to a collection of data to solve an issue[11], although it might even have a broader application such as, for example, the instructions for baking a cake, i.e., what we commonly call a cooking recipe. In recent years, however, the name algorithm is usually applied to AI algorithms. Artificial intelligence algorithms are a category of technology that is capable of learning as well as determining what would be the most suitable rules to solve an issue based on some data and a target output.[12] Furthermore, AI algorithms have significantly enhanced their performance by gaining mature cognitive abilities, including perception, reasoning, and decision-making.[13] But

---

[5] Mehedi, I. M., Hanif, M. S., Bilal, M., Vellingiri, M. T., & Palaniswamy, T. (2024). Remote sensing and decision support system applications in precision agriculture: Challenges and possibilities. *Ieee Access*.

[6] Mohanarajesh, K. (2024). Develop New Techniques for Ensuring Fairness in Artificial Intelligence and ML Models to Promote Ethical and Unbiased Decision-Making.

[7] Donvir, A., & Sharma, G. (2025, January). Ethical Challenges and Frameworks in AI-Driven Software Development and Testing. In *2025 IEEE 15th Annual Computing and Communication Workshop and Conference (CCWC)* (pp. 00569-00576). IEEE.

[8] Zhao, J., & Gómez Fariñas, B. (2023). Artificial intelligence and sustainable decisions. *European Business Organization Law Review*, *24*(1), 1-39.

[9] Fernandez, M., Bellogin, A., & Cantador, I. (2024, May). Analysing the effect of recommendation algorithms on the spread of misinformation. In *Proceedings of the 16th ACM Web Science Conference* (pp. 159-169).

[10] Akter, S., Dwivedi, Y. K., Sajib, S., Biswas, K., Bandara, R. J., & Michael, K. (2022). Algorithmic bias in machine learning-based marketing models. *Journal of Business Research*, *144*, 201-216.

[11] Zhang, H., Ma, Y., Yuan, K., Khayatnezhad, M., & Ghadimi, N. (2024). Efficient design of energy microgrid management system: a promoted Remora optimization algorithm-based approach. *Heliyon*, *10*(1).

[12] Aldoseri, A., Al-Khalifa, K. N., & Hamouda, A. M. (2023). Re-thinking data strategy and integration for artificial intelligence: concepts, opportunities, and challenges. *Applied Sciences*, *13*(12), 7082.

[13] Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. *Smart Learning Environments*, *11*(1), 28.

very far from this picture of objectivity and neutrality, several studies have pointed out several potential issues with the use of algorithms, most notably the fact that they can be biased.[14]

Biases are systematic judgment and decision-making errors that happen in information processing and interpretation.[15] Most of these biases that are now being found in AI, are discriminative social biases, including racism, socioeconomic class, and even sexism.[16] In fact, AI has been most criticized for the amplification and influence of discriminative social biases, including gender bias.[17]

Understanding Bias in AI-Driven Legal Systems

AI bias is systematic. AI system errors that result in biased or unequal outcomes.[18] This may involve problems like inaccurate predictions, excessive false negatives or decision-making that disproportionately impact marginalized communities. This is based on biased assumptions made during its development and deployment, such as during AI data collection and preparation.

A simple illustration would be machine learning algorithms employed in hiring and using AI. An HR model can be modeled using biased historic data that targets applicants from specified schools or ones with specific merit. This tends to result in hiring that fuels inequality and diminishes diversity and talent, becoming a major setback in building a just and welcoming workplace. And legal issues including discrimination lawsuits, as well. Algorithmic bias can be created at any phase of an AI system's lifetime: data procurement, data labeling, model building, AI research and deployment, resulting in an "unfair" AI system.[19]

AI and ML systems rely on the quality of training data.[20] Data in our world is sometimes thought of as a source of truth, but it's not that simple. How data is created, built and interpreted changes over time, mirroring changes in the world. At times, this data might not reflect certain populations accurately, resulting in biased results. At other times, datasets can be representative but also exhibit historical bias and current biases that can be propagated into the AI system.[21]

---

[14] Varsha, P. S. (2023). How can we manage biases in artificial intelligence systems–A systematic literature review. *International Journal of Information Management Data Insights*, *3*(1), 100165.
[15] Cherry, K. (2024). How cognitive biases influence the way you think and act. *URL: https://www. verywellmind. com/what-is-a-cognitive-bias-2794963 (date of access: 23.12. 2024)*.
[16] Stypinska, J. (2023). AI ageism: a critical roadmap for studying age discrimination and exclusion in digitalized societies. *AI & society*, *38*(2), 665-677.
[17] O'Connor, S., & Liu, H. (2024). Gender bias perpetuation and mitigation in AI technologies: challenges and opportunities. *AI & SOCIETY*, *39*(4), 2045-2057.
[18] Ferrara, E. (2024). The butterfly effect in artificial intelligence systems: Implications for AI bias and fairness. *Machine Learning with Applications*, *15*, 100525.
[19] Narula, S., Afaq, A., Nagar, S., & Chaudhary, M. (2025). Transformative Potential and Ethical Challenges of Generative AI in E-Commerce: Data Bias, Algorithm Bias. In *Responsible Implementations of Generative AI for Multidisciplinary Use* (pp. 317-336). IGI Global.
[20] Zha, D., Bhat, Z. P., Lai, K. H., Yang, F., Jiang, Z., Zhong, S., & Hu, X. (2025). Data-centric artificial intelligence: A survey. ACM Computing Surveys, 57(5), 1-42.
[21] **Oguntibeju**, O. O. (2024). Mitigating artificial intelligence bias in financial systems: A comparative analysis of debiasing techniques. *Asian Journal of Research in Computer Science*, *17*(12), 165-178.

Humans could shape AI systems in many ways, like choosing which data sets to include or exclude, deciding how to label the data, deciding on training, and controlling feedback loops.[22]

Comprehending bias in AI involves knowledge of coexisting definitions of the term bias set within the context of AI development and deployment.[23] In the absence of context, the word bias can be used to describe an uneven preference for a group or an individual due to preconceived attitudes or notions. The use of the words explicit and implicit separates intentional from unintentional bias, respectively.

Cognitive bias can be defined as systematic departure from objective judgment.[24] Intuitive reasoning based on heuristics, or mental shortcuts to reduce information processing and decision-making, is susceptible to cognitive bias and diagnostic errors.[25] Cognitive biases built into model construction can reinforce existing health inequities.[26] Social bias may also arise from the absence of diversity in data sets used to train a model or assumptions during the model construction process, enhancing inherent social and cultural bias.[27]

Psychology has a long history of research in cognitive biases and has made vast strides in the examination of how biases affect human judgments and decision-making.[28] It has even been suggested that psychology's most valuable contribution to society is the discovery of effective debiasing interventions.[29] Thus, the input from cognitive psychology may be also significant to the research of how to eliminate biases that affect AI by way of human interference.[30] There is a lot of research that has established that human thinking is subject to a skewed perception of reality. Illusions are cognitive biases which lead individuals to mistakenly interpret visual information.[31]

---

[22] Zha, D., Bhat, Z. P., Lai, K. H., Yang, F., Jiang, Z., Zhong, S., & Hu, X. (2025). Data-centric artificial intelligence: A survey. *ACM Computing Surveys*, *57*(5), 1-42.

[23] Brauner, P., Hick, A., Philipsen, R., & Ziefle, M. (2023). What does the public think about artificial intelligence?—A criticality map to understand bias in the public perception of AI. *Frontiers in Computer Science*, *5*, 1113903.

[24] Bedeley, R. T., Hao, H., & Ghoshal, T. (2025). Cognitive Biases in Online Opinion Platforms: A Review and Mapping. *SAGE Open*, *15*(1), 21582440251315564.

[25] Mangus, C. W., & Mahajan, P. (2022). Decision making: healthy heuristics and betraying biases. *Critical Care Clinics*, *38*(1), 37-49.

[26] Purushothaman, M. B., Jessica, P., & Rotimi, F. E. (2025). Analysis of Cognitive Biases in Construction Health and Safety in New Zealand. *Buildings*, *15*(7), 1033.

[27] Alvarez, J. M., Colmenarejo, A. B., Elobaid, A., Fabbrizzi, S., Fahimi, M., Ferrara, A., ... & Ruggieri, S. (2024). Policy advice and best practices on bias and fairness in AI. *Ethics and Information Technology*, *26*(2), 31.

[28] Schirrmeister, E., Göhring, A. L., & Warnke, P. (2020). Psychological biases and heuristics in the context of foresight and scenario processes. *Futures & Foresight Science*, *2*(2), e31.

[29] Lewis Jr, N. A. (2023). Cultivating equal minds: Laws and policies as (de) biasing social interventions. *Annual Review of Law and Social Science*, *19*(1), 37-52.

[30] Goyal, A., & Bengio, Y. (2022). Inductive biases for deep learning of higher-level cognition. *Proceedings of the Royal Society A*, *478*(2266), 20210068.

[31] Todorović, D. (2020). What are visual illusions?. *Perception*, *49*(11), 1128-1199.

As described earlier, cognitive biases are consistent and predictable errors that happen to all of us when we process and interpret information.[32] Additionally, there is a consensus in the literature that cognitive biases reflect adaptive processes which are typically referred to as heuristics. Heuristics are mental shortcuts that enable individuals to take speedy and effortless judgments and decisions in situations of uncertainty, and they generally lead to efficiency, adaptive, and required.[33] But when circumstances become different, individuals might continue applying the same shortcut, and under such circumstances, the outcome may be maladaptive, flawed, and troublesome. In such instances, they are referred to as biases. That is, heuristics and biases are the two faces of the same coin. Heuristics are adaptive strategies for handling uncertain situations, but they at times lead to biases, that is, systematic errors that arise in most individuals under specific conditions.[34]

The Impacts of Human Cognitive Biases on AI

One of the most significant potential sources of bias in AI systems is biases or imbalances in the data that the algorithms are trained on.[35] AI algorithms learn from patterns and make predictions of historical data. High-quality data sets to feed into the algorithms are hard to come by, especially when they relate to clinical data.[36] Therefore, if a data set is imbalanced or biased, the AI systems trained on that data will learn and even replicate the biases. In medicine, the data to train the algorithms typically is the result of previous human choices, such as medical choices about what type of patient needs a specific test or what type of patient would be better served by a specific treatment. If the history of the decisions made in the past is revealed to have extended systematic flaws like, for example, misclassification of a particular pathology when some characteristics are present, then the AI system that is trained in this history will simply learn this prejudice. Consequently, this systematic flaw might be exemplified by repeated misclassification of a particular pattern of colored pixels in an image. But this bias would bring catastrophic effects if this type of model flaw isn't understood, and the AI system is implemented in decision-making in high-risk situations. One of the very strong arguments for the implementation of cognitive biases in humans is their reduction of effort and speed.

Human beings require biases since they lack sufficient computation power to deal with all the information they possess regarding their world.[37] Naturally, human beings limited computational capacity is absent from computers. But apart from considerations on evolutionary pressure concerning intelligence formation mentioned at the beginning of the

---

[32] Saposnik, G., Redelmeier, D., Ruff, C. C., & Tobler, P. N. (2016). Cognitive biases associated with medical decisions: a systematic review. *BMC medical informatics and decision making*, *16*, 1-14.

[33] Raue, M., & Scholl, S. G. (2018). The use of heuristics in decision making under risk and uncertainty. *Psychological perspectives on risk and risk analysis: Theory, models, and applications*, 153-179.

[34] Raue, M., & Scholl, S. G. (2018). The use of heuristics in decision making under risk and uncertainty. *Psychological perspectives on risk and risk analysis: Theory, models, and applications*, 153-179.

[35] Ferrara, E. (2023). Fairness and bias in artificial intelligence: A brief survey of sources, impacts, and mitigation strategies. *Sci*, *6*(1), 3.

[36] Ahmed, Z., Mohamed, K., Zeeshan, S., & Dong, X. (2020). Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine. *Database*, *2020*, baaa010.

[37] Soprano, M., Roitero, K., La Barbera, D., Ceolin, D., Spina, D., Demartini, G., & Mizzaro, S. (2024). Cognitive biases in fact-checking and their countermeasures: a review. *Information Processing & Management*, *61*(3), 103672.

chapter and considerations regarding 'Green AI', meaning resource efficiency and less computational requirements in machine learning systems.[38] Besides effort reduction and higher accuracy, other benefits of human cognitive biases in psychology are transparency and accessibility.

Biased human judgment is transparent, it does not need extensive training, and even laymen can make good decisions, for which normally experts are needed, if there is a good heuristic, for instance, in the form of a fast-and-frugal tree, which is a small decision tree for binary classification problems. This benefit can be adapted to the area of machine learning since algorithms will gain from the receipt of sparser training, particularly in case of small training data sets.

The Concept of Fairness in AI Decision-Making

Fairness in general means to act equally and fairly, impartially, and without discrimination. Fairness in AI means making sure that the systems act in that manner – fairly.[39] That is, because AI and ML handle a lot of data, fairness means handling this data responsibly and making decisions that do not have unjustified negative impacts on any specific individual or group. AI fairness is all about designing, developing, and deploying AI models that do not discriminate, or benefit individuals and groups based on factors such as race, gender, or economic status.[40]

In addition to the legal implications of unfair AI, fairness in AI is also necessary if we wish to have the mass acceptance and adoption of AI systems.[41] This is because we want to be assured that AI can make the correct and fair decisions, particularly if we intend to introduce AI into business and more complicated operations.

Legal Frameworks and Regulatory Challenges

The regulatory and legal systems have an important role to play in entheansparency and effective accountability in information systems that employ Artificial Intelligence.[42] Privacy legislation improves equitable processing by guaranteeing that companies communicate to the public details of their use of data and by enabling users to contribute and have a say in how their personal data is being utilized.[43] This legislation also encourages accountability because the GDPR provides a subject right to pursue an action against an organization for an illegal

---

[38] Bolón-Canedo, V., Morán-Fernández, L., Cancela, B., & Alonso-Betanzos, A. (2024). A review of green artificial intelligence: Towards a more sustainable future. *Neurocomputing*, 128096.

[39] Pfeiffer, J., Gutschow, J., Haas, C., Möslein, F., Maspfuhl, O., Borgers, F., & Alpsancar, S. (2023). Algorithmic fairness in AI: an interdisciplinary view. *Business & Information Systems Engineering*, *65*(2), 209-222.

[40] Pulivarthy, P., & Whig, P. (2025). Bias and Fairness Addressing Discrimination in AI Systems. In *Ethical Dimensions of AI Development* (pp. 103-126). IGI Global.

[41] Kelly, S., Kaye, S. A., & Oviedo-Trespalacios, O. (2023). What factors contribute to the acceptance of artificial intelligence? A systematic review. *Telematics and Informatics*, *77*, 101925.

[42] Díaz-Rodríguez, N., Del Ser, J., Coeckelbergh, M., de Prado, M. L., Herrera-Viedma, E., & Herrera, F. (2023). Connecting the dots in trustworthy Artificial Intelligence: From AI principles, ethics, and key requirements to responsible AI systems and regulation. *Information Fusion*, *99*, 101896.

[43] McGraw, D., & Mandl, K. D. (2021). Privacy protections to encourage use of health-relevant digital data in a learning health system. *NPJ digital medicine*, *4*(1), 2.

automated decision. Similarly, accountability is enabled by anti-discrimination legislations that also prohibit the creation of AI systems with discriminatory outcomes to civil societies while simultaneously providing individuals with avenues to sue for discrimination.[44] There have been controversies, debates, issues and more controversies with the practical application of this right as to how much, and in what way, it can be enforced. There is still some restriction in the existing data protection legislation in relation to coverage and enforcement of AI systems.

In the case against Clearview AI[45], American facial recognition startup was charged by the American Civil Liberties Union (ACLU) with violating Illinois' Biometric Information Privacy Act (BIPA). The company harvested billions of photos from Fb and other web sites in a bid to build a facial recognition database, which they snapped against the will of the customers. This case raises relevant questions about the exploitation of AI systems that rely on individual data.

Those issues and voids are doubtlessly best served by some gradual tinkering with data protection law and some additional AI governance approaches. This would possibly involve widening the definition of 'personal data' to account for inferred/inferenced data, introducing privacy impact assessment and algorithmic audit to high-risk AI, and requiring full model reporting to make way for breadth.

Ethical Considerations in AI-Driven Legal Systems

Of the problems enumerated below, some are about the nature and design of AI itself, and some are about implementation and use of AI problems (though frequently the nature of AI design allows for or generates implementation and use problems). Some of the problems will be boundary-spanning (i.e., they will occur in one or more areas/applications of use). These issues mainly relate to all technology (i.e., data protection/privacy), some of which are linked to each other (i.e., fairness, transparency, accountability), and perhaps might not be in existence individually. There is no need, however, to downplay AI's ability to augment, and or generate, harmful implications related to the generation and/or use of these problems.

The opaqueness of algorithms is one of the high-profile topics that squarely fall under the purview of legal debate on AI.[46] As the mass adoption of AI is increasingly coming to light in areas classified as high-risk, pressure mounts to develop and regulate AI for responsibility, equity, and transparency.[47] The deployment of AI weapons outside of human control; AI vulnerability to ctargetingy issues; the targeting of surveillance or cyber security through AI for national security providing a new way in for attack based on 'data diet vulnerability'; the deployment of network intervention techniques by means of foreign-located AI; an even more

---

[44] Cheong, B. C. (2024). *Transparency and Accountability in AI Systems: Safeguarding Wellbeing in the Age of Algorithmic Decision-Making. Frontiers in Human Dynamics, 6, 1421273.*

[45] European Data Protection Board. The French SA fines Clearview AI EUR 20 million. October 20, 2022 (accessed via https://www.edpb.europa.eu/news/national-news/2022/french-sa-fines-clearview-ai-eur-20-million_en)

[46] Al-Dulaimi, A. O. M., & Mohammed, M. A. A. W. (2025). Legal responsibility for errors caused by artificial intelligence (AI) in the public sector. *International Journal of Law and Management*.

[47] Lim, D. (2025). Determinants of Socially Responsible AI Governance. *Duke Law & Technology Review*, 23(1), 183-232.

pervasive and advanced variant of existing high-level targeting of political content through social media. They are critical problems because they put critical infrastructures at risk of damage with devastating impacts on society and human beings, posing to destroy human security, as well as resource access.

Security threats in cyberspace are also serious threats because they are usually stealthy and made visible too late.[48] Unfairness, discrimination, and bias appear to be iterative and significant regarding the problems posed by the application of automated and algorithmic decision-making systems (e.g. to render a decision in employment, health, criminal justice, and insurance).

Some rights for individuals are offered by European Union data protection law for challenging and requiring a review of automated decision-making with significant effect on the rights or legitimate interests of an individual (GDPR 2016/679). Data subjects are also free, at any moment, to object, on grounds pertaining to the case of the individual, to personal data processing concerning the individual with respect to the individual for reasons performed in public interest or in legitimate interest.[49] In addition, pursuant to Article 22(3) GDPR and other provisions that are applicable to data controllers, where the processing of personal data leads to decisions that are taken solely on the basis of automated processing, data controllers shall take adequate measures to give due consideration to a data subject's rights, interests, and legitimate rights, at least to access human intervention by the controller, to express views and to be able to challenge the decision. There is disagreement as to whether AI and/or robotics systems "fit into existing legal categories or whether a new category must be created based on these specific features and implications." (European Parliament Resolution 16 February 2017).

It is not just a matter of law; it is politically controversial. The High-Level Expert Group on Artificial Intelligence (AI HLEG) has appealed to "policymakers not to grant a legal personality to AI systems or robots"[50] arguing that it is "inherently inconsistent with the principle of human agency, accountability and responsibility" and entails the "moral hazard." Intellectual property rights are covered in the Universal Declaration of Human Rights (UDHR, Article 27), the International Covenant on Economic, Social and Cultural Rights (ICESCR, Article 15), the International Covenant on Civil and Political Rights (ICCPR, Article 19) and Vienna Declaration and Programme of Action (VDPA) 1993. Intellectual property rights have a "human rights character" and "have increasingly been contextualized in a range of policy areas" World Intellectual Property Organization (WIPO).

---

[48] Aslan, Ö., Aktuğ, S. S., Ozkan-Okay, M., Yilmaz, A. A., & Akin, E. (2023). A comprehensive review of cyber security vulnerabilities, threats, attacks, and solutions. *Electronics*, *12*(6), 1333.

[49] Vogiatzoglou, P., & Valcke, P. (2022). Two decades of Article 8 CFR: A critical exploration of the fundamental right to personal data protection in EU law. In *Research handbook on EU data protection law* (pp. 11-49). Edward Elgar Publishing.

[50] European Parliament, Artificial Intelligence ante portas: Legal & ethical reflections (https://www.europarl.europa.eu/at-your-service/files/be-heard/religious-and-non-confessional-dialogue/events/en-20190319-artificial-intelligence-ante-portas.pdf)

The IBA Global Employment Institute report (2017) addresses the workplace implications of AI and robotics.[51] Some of the consequences include: adjustments to the types of employment they will need in the future; erosion of free hours; shifting in employment relations; new patterns of work and new employment forms; laying off of employees; inequities in the "new" employment setting; potential inclusion of individuals in the "new" employment setting without the critical training; employment relations (and the how it could affect union activity and collective bargaining aspects and range of changes that could hinder employee representatives); workplace health and safety effects; effects on working time; effects on pay (wage effects and pensions); and social security effects. Both data protection enforcing authorities and academics are of the opinion that AI poses significant privacy and data protection rights risks (among others).[52]

These risks are monitoring and informed consent; and violation of data protection rights of individuals, such as the right to receive their own personal data, the right to halt any processing taking place that has the potential to cause damage or distress, and the right not to be subjected to a decision based on purely automated processing. Application and deployment of AI technologies can cause damage to human life and property; consider some examples - maiming pedestrians when the steering component of an autonomous vehicle is faulty, crash and destruction caused by a half or part-autonomous drone, and incorrect treatment diagnosis through a defective AI program software report. As highlighted in the Assessment List for Trustworthy AI (ALTAI), accountability requires that there are means to hold responsible the development, deployment and/or usage of AI systems in a manner that can be susceptible to risk management, discovering and addressing risk in a manner that is transparent and explainable to third party audits. The accountability philosophy within the context of AI requires the (i) action-guiding function (through action-guiding beliefs and decision-making) and (ii) action-explaining function (through situating actions in a broader context or classification under moral values).

International human rights conventions impose obligations on States Parties to comply with and uphold: States must avoid interfering with rights and take positive action to bring them into operation. None of them now specifically mention artificial intelligence/AI or machine learning', but their general wording would embrace most of the issues and challenges outlined above.

Mitigating Bias and Enhancing Fairness in AI Legal Systems

Relative to the problem of a lack of transparency in algorithms, it might be addressed by (i) awareness raising: education, watchdogs and whistle-blowers; (ii) accountability on the algorithms used by public actors; and (iii) regulation and legal liability; and (iv) global

---

[51] Lubinga, S. N., Maramura, T. C., & Masiya, T. (2023). Adoption of Fourth Industrial Revolution challenges in South African higher education institutions.

[52] Yanamala, A. K. Y., & Suryadevara, S. (2023). Advances in Data Protection and Artificial Intelligence: Trends and Challenges. *International Journal of Advanced Engineering Technologies and Innovations*, *1*(01), 294-319.

coordination of algorithmic governance. Some more specific solutions advanced to provide algorithmic transparency include algorithmic impact assessments.[53]

Regarding cybersecurity vulnerabilities, numerous various methods and tools exist or are in the process of being proposed for solving this problem.[54] For instance, having sound protection and recovery procedures, including vulnerability awareness within the design procedure, involving human analysts in decisions, using risk management programs, and applying patches to software.

On contestability, the idea of defining original contestability has been advanced to better protect the rights of decisions by solely automated processing, as a condition precedent to each stage of an artificial intelligence system's lifecycle.

Regarding the topic of whether Artificial Intelligence should have legal personality, the model of keeping human accountability for AI and its behavior as a tool could be the most preferred, and this would mean the responsibility would fall on the developers, users, or owners of AI, not on the AI. When we use AI producing value, we should treat it like any physical product, where manufacturers and operators are responsible for the damage and harm caused by it.[55] If AI systems became completely autonomous (imagine errors in autonomous shipping businesses), it may be appropriate for the system to be classified as a corporate entity to define its legal responsibilities as an AI business.[56] AI could then have limited rights and responsibility like the concept of corporate personhood but primarily related to the enforcement of contracts.[57] The positive aspect to the model of legal personhood could be for AI to be recognized as an agent of its owner, creating a clear pathway of legal responsibility back to a human.[58]

Policy Recommendations for Ethical AI Governance

Regulatory frameworks and guidelines need to be developed and implemented to promote the ethical and responsible application of AI in legal decision-making. These guidelines should respond to the ethical issues raised above and make practical suggestions for the application of AI in the judiciary.

---

[53] Stankovich, M. I. R. I. A. M., Behrens, E. R. I. C. A., & Burchell, J. U. L. I. A. (2023). Toward Meaningful Transparency and Accountability of AI Algorithms in Public Service Delivery. *DAI Shaping a more livable world*, 1-33.

[54] Mazhar, T., Irfan, H. M., Khan, S., Haq, I., Ullah, I., Iqbal, M., & Hamam, H. (2023). Analysis of cyber security attacks and its solutions for the smart grid using machine learning and blockchain methods. *Future Internet*, *15*(2), 83.

[55] Schütte, B., Majewski, L., & Havu, K. (2021). Damages liability for harm caused by Artificial Intelligence–EU law in flux. *Helsinki Legal Studies Research Paper*, (69).

[56] Copp, C. J., Cabell, J. J., & Kemmelmeier, M. (2023). Plenty of blame to go around: Attributions of responsibility in a fatal autonomous vehicle accident. *Current Psychology*, *42*(8), 6752-6767.

[57] Lovell, J. (2023). Legal Aspects of Artificial Intelligence Personhood: Exploring the Possibility of Granting Legal Personhood to Advanced AI Systems and the Implications for Liability, Rights and Responsibilities. *Rights and Responsibilities (May 10, 2023)*.

[58] Avila Negri, S. M. (2021). Robot as legal person: Electronic personhood in robotics and artificial intelligence. *Frontiers in Robotics and AI*, *8*, 789327.

Stakeholder collaboration, such as legal practitioners, AI professionals, ethicists, and policymakers, is important in developing comprehensive and effective guidelines. Professional associations, legal bodies, and governments can be important in setting and implementing these guidelines.

Ethical implications and compliance with guidelines in the areas of explainability, avoidance of bias, data security and privacy, responsibility, and proper utilization of AI are essential in making the utilization of AI in legal decision-making responsible and efficient. Through emphasis on transparency, equity, and safeguarding personal information, legal experts and developers can harness the potential of AI technologies while sustaining ethical positions in the legal profession. It is with a careful and responsible mindset that AI can enhance the judicial process, enhancing efficiency, accuracy, and accessibility of justice.

There is a necessity to continue the interdisciplinary tendencies and to refine the above models in accordance with the new technologies and social environments. Above all, present and future attempts to make AI 'responsible' need to deal with the essence of power relations between AI creators and AI users. At the central level, this includes involving the target AI victims who are vulnerable and at greatest risk of AI harm in AI governance and regulation. Therefore, it is also essential to maintain participatory methods in the policymaking of AI independently and to their agenda to democratize AI responsibility.

Institutional governance mechanisms also improve accountability in all AI systems. This also includes controllable supra-organizational AI regulatory bodies coming under the technical category and equipped with the investigative powers to examine AI systems and enforce compliance with the transparency provisions. Establishing AI ombudsperson and public advocate may enable the concerned communities to bring the issues and seek remedies for AI-inflicted cruelties. Other reforms which are anticipated to motivate more ethical behavior from within AI firms are whistleblower protections and ethical AI oaths.

Conclusion

There are apparent advantages to incorporating AI in the legal decision-making process, such as a faster turnaround time for decisions, more comprehensive analysis of pertinent legal factors, and more access to justice. However, it is necessary to recognize the ethical implications attached to the implementation of AI within the judicial process. The proper and effective application of AI in the legal arena requires thorough examination of concerns regarding explainability, bias, data privacy, and security.

For the sake of explainability and transparency, AI systems must offer understandable results, so legal experts and the public can comprehend and assess the rationale for AI-made decisions. To reduce bias and ensure fairness, there is a need for rigorous curation of training data, frequent auditing of AI systems, and cross-disciplinary collaboration to question biases and ensure fairness.

Although AI is currently the topic of conversation, it is believed that with the addition of robotics and IoT, all new developments in technologies, including AI, will turn the discussed fresh set of unique problems into laws and societal values that will require new robust discussions. This article suggests a systematic review and analysis of ethical concerns and issues related to AI and ethical standards and principles published by a variety of organizations, frameworks and methodologies for addressing ethical issues of AI or implementing the ethical principles of AI, and their degree of ethicality (or morality in AI). In addition, a few concerns in the practice of AI ethics are presented and suggestions for future research.