JSH
Jurnal Smart Hukum
https://ejournal.ipinternasional.com/index.php/jsh
Open Access Journal
by Inovasi Pratama Internasional. Ltd
Member of Crossref
ISSN. 2961-841X

Algorithmic Bias in Artificial Intelligence as A Threat to The Right to Equality

Bagus Gede Ari Rama[1], Kadek Julia Mahadewi[2], Dewa Krisna Prasada[3]
[123] Faculty Of Law, Universitas Pendidikan Nasional, Indonesia
E-mail: arirama@undiknas.ac.id[1] , juliamahadewi@undiknas.ac.id[2] , krisnaprasada@undiknas.ac.id[3]
Article history: *Received May 05, 2026: Revised May 25, 2026: Accepted June 18, 2026*

ABSTRACT

This study examines the implications of algorithmic bias in artificial intelligence (AI) systems, particularly its potential to generate digital discrimination that undermines the principle of equality before the law in Indonesia. While AI is widely adopted to enhance data-driven decision-making across various sectors, its outputs are not inherently neutral, as they are shaped by training data and algorithmic design. This condition raises concerns regarding unfair treatment of individuals or groups, which may conflict with constitutional guarantees of equality and non-discrimination under the 1945 Constitution of the Republic of Indonesia and Law No. 39 of 1999 on Human Rights. Using a normative legal research method with statutory and conceptual approaches, this study analyzes relevant legal frameworks governing AI and human rights protection in Indonesia. The findings show that existing regulations, including provisions on electronic system governance and personal data protection, remain general and insufficient to address the complexity of algorithmic bias. Accordingly, this study argues for strengthening the legal framework through the adoption of algorithmic transparency requirements, mandatory bias testing in AI development, and the establishment of independent auditing mechanisms. These measures are necessary to ensure that AI systems operate fairly, accountably, and in alignment with human rights principles, particularly the protection of equality before the law.

Keywords: Algorithmic Bias, Digital Discrimination, Artificial Intelligence, Legal Equality, Technology Regulation.

INTRODUCTON

The development of digital technology over the past decade has been marked by a substantial increase in the deployment of Artificial Intelligence (AI) across a wide range of sectors. AI systems have evolved beyond basic computational and automation functions into sophisticated tools capable of performing complex data analyses, generating predictive insights, and supporting or even replacing human decision-making processes. These systems are now widely utilized in strategic domains, including employment recruitment, financial credit scoring, healthcare diagnostics, public security, and digital public service deliveries. Recent reports indicate that AI-driven decision-making systems are increasingly shaping individuals' access to economic and social opportunities, raising critical governance and regulatory concerns.(OECD, 2024)

The adoption of AI in decision-making is often justified based on efficiency, consistency, and perceived objectivity. Algorithmic systems are assumed to reduce human subjectivity by relying on data-driven processes. However, a growing body of interdisciplinary research demonstrates that these systems are not inherently neutral. Instead, AI systems may embed and reproduce the structural inequalities present

in the data and institutional environments in which they operate. This phenomenon, commonly referred to as algorithmic bias, describes systematic distortions in automated decision-making that result in unequal or unfair outcomes across social groups.(Ferrara, 2023)

Algorithmic bias may arise from multiple sources, including biased or non-representative training data, imbalanced data distributions, flawed feature selection, and assumptions embedded in the model design. Where AI systems are trained on historical datasets reflecting pre-existing social inequalities, they may replicate and even amplify these inequalities in automated decision-making processes. Contemporary research in data science and AI ethics demonstrates that such biases can persist even in the absence of explicit discriminatory intent because of the presence of proxy variables and latent correlations within datasets.(Friedler et al., 2021)

Empirical evidence has highlighted the manifestation of algorithmic bias in various AI applications. Studies on facial recognition technologies have revealed significant disparities in accuracy rates across demographic groups, particularly in terms of race and gender. These findings illustrate how AI systems trained on unbalanced datasets may produce systematically biased outcomes that disproportionately affect marginalized communities.(Raji & Buolamwini, 2022) Similar concerns have been identified in AI-assisted recruitment systems, where automated evaluation tools may disadvantage candidates based on characteristics unrelated to professional competence, such as linguistic patterns or socio-cultural background.(Kordzadeh & Ghasemaghaei, 2022)

The deployment of AI in public security and law enforcement contexts raises concerns regarding misidentification and its implications for individual rights. Empirical studies and institutional assessments indicate that differential error rates in algorithmic systems may result in wrongful identification, thereby posing risks to due process, privacy, and equality before the law.(Nations, 2024) These developments demonstrate that the impact of AI extends beyond technical performance to fundamental questions of justice and rights protection.

From a human rights perspective, the principle of equality constitutes a foundational norm that guarantees that all individuals are entitled to equal treatment before the law without discrimination. However, the emergence of algorithmic decision-making has introduced new and complex forms of discrimination that differ from traditional legal paradigms. Unlike direct discrimination, algorithmic bias often operates indirectly through opaque computational processes, making it difficult to detect, explain, and challenge such biases. Contemporary legal scholarship increasingly conceptualizes this as a form of systemic or indirect discrimination, requiring adapted legal and regulatory responses.(Xenidis & Senden, 2023)

At the international level, regulatory attention to the human rights implications of AI has intensified. The adoption of the *Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law* by the Council of Europe in 2024 reflects a growing consensus that AI development and deployment must be aligned with human rights standards, democratic values, and the rule of law. This development underscores the recognition that AI technologies can profoundly affect the protection of fundamental rights, including the right to equality in a digital society.(Europe, 2024)

Nevertheless, existing legal frameworks are largely designed to address discrimination arising from identifiable human actions or institutional policies. Traditional legal doctrines of discrimination are premised on the existence of intent or conduct that can be clearly attribute to discrimination. In contrast, algorithmic discrimination may emerge from complex data-driven processes without explicit intent from developers or users. This creates a structural gap between technological advancement and existing legal frameworks, particularly in ensuring the effective protection of the right to equality in the context of automated decision-making.

Based on the foregoing, this research addresses two principal questions:
(1) What are the characteristics of algorithmic bias in Artificial Intelligence systems that have the potential to give rise to discrimination against the right to equality in a digital society?
(2) How can the principles of the protection of the right to equality within the framework of human rights law be applied to mitigate the risks of discrimination generated by Artificial Intelligence systems?

METHOD

This study adopts a normative legal research method aimed at examining legal issues concerning algorithmic bias in Artificial Intelligence (AI) systems and its implications for the protection of the right to equality from a human rights perspective. Normative legal research conceptualizes law as an autonomous system of prescriptive norms analyzed through doctrinal inquiry based on statutory provisions, legal principles, and legal doctrines.(Taekema, 2021) This approach emphasizes the internal coherence of the legal system and focuses on *das sollen* rather than empirical realities, thereby positioning legal reasoning within a structured normative framework.(Ansari & Negara, 2023) Accordingly, this research focuses on the analysis of legal norms, principles, and conceptual frameworks governing the use of artificial intelligence and the protection of human rights.

The approaches employed in this study include the statute, conceptual, and case approaches. The statute approach is used to examine relevant legal instruments governing digital technologies, human rights protection, and the principles of equality and non-discrimination within both national and international legal systems, with particular emphasis on statutory interpretation as the primary source of normative legal authority.(Dyzenhaus et al., 2023) The conceptual approach is applied to analyze legal doctrines and theoretical developments in academic literature, particularly those relating to algorithmic bias, digital discrimination, and the principle of equality in human rights law, drawing on contemporary interdisciplinary scholarship at the intersection of law, ethics, and artificial intelligence. A case approach is used to assess empirical instances of AI deployment that give rise to algorithmic bias, as reflected in real-world sociotechnical systems, public policy practices, and corporate implementations, thereby enabling an evaluation of how algorithmic systems operate within complex social and institutional contexts.

The legal materials used in this study consist of primary, secondary, and tertiary sources. Primary legal materials include international human rights instruments such as the *Universal Declaration of Human Rights* (1948) and the *International Covenant on Civil and Political Rights* (1966), as well as emerging regulatory frameworks on artificial intelligence, including the *Framework Convention on Artificial Intelligence* adopted by the Council of Europe (2024). These instruments form the normative foundation for assessing equality and non-discrimination in digital governance regimes.

Secondary legal materials comprise scholarly books, peer-reviewed journal articles, research reports, and policy documents addressing Artificial Intelligence, algorithmic bias, and human rights protection. Contemporary research highlights that algorithmic systems may reproduce structural inequalities embedded in training data and sociotechnical infrastructures, necessitating interdisciplinary legal analysis.

Tertiary legal materials include legal dictionaries, encyclopedias, and authoritative reference sources that provide clarification of legal terminology and conceptual frameworks relevant to this study, particularly in interpreting evolving legal concepts within technology regulation and human rights discourse.(Garner, 2019)

Legal materials were collected through library research, involving a systematic review of academic literature, institutional reports issued by international organizations such as the Organisation for Economic Co-operation and Development, scientific publications, and reputable media sources. This method enables the identification and synthesis of normative and analytical insights derived from authoritative legal and policy-oriented sources, particularly in the rapidly evolving domain of artificial intelligence governance. These materials were subsequently compiled, classified, and analyzed to obtain a comprehensive understanding of the relationship between algorithmic bias in AI systems and the protection of the right to equality within a human rights framework.

The analysis of legal materials was conducted qualitatively using a descriptive-analytical method. This method involves the systematic interpretation of legal norms, doctrines, and conceptual frameworks, followed by their correlation with empirical developments in the deployment of AI systems that may lead to digital discrimination. Qualitative legal analysis emphasizes interpretative reasoning and doctrinal coherence when examining how legal principles operate within contemporary technological contexts.(Flick, 2022) Through this analytical process, the study aims to develop a structured understanding of the characteristics of algorithmic bias in Artificial Intelligence systems and to formulate a conceptual framework for the protection of the right to equality within the broader framework of human rights law.

RESULTS AND DISCUSSION

Characteristics of Algorithmic Bias in Artificial Intelligence Potentially Giving Rise to Discrimination Against the Right to Equality

The increasing deployment of Artificial Intelligence (AI) systems in decision-making processes has generated substantial scholarly attention, particularly regarding their implications for equality and non-discrimination. Recent interdisciplinary research demonstrates that algorithmic systems, despite being framed as objective and data-driven, frequently reproduce systemic inequalities embedded in social and institutional structures.

Algorithmic bias is broadly defined as a systematic distortion in automated decision-making processes that leads to unequal or unfair outcomes across different groups. Contemporary research highlights that such bias may originate not only from data but also from model architecture, optimization strategies, and broader sociotechnical environments in which AI systems operate.(Suresh & Guttag, 2021) In high-impact domains such as employment, financial services, and public administration, these biases can significantly affect individuals' access to socio-economic opportunities (Suresh & Guttag, 2021). From a human rights perspective, the right to equality is firmly established in Article 2 of the Universal Declaration of Human Rights (1948) and Article 26 of the International Covenant on Civil and Political Rights (1966). Recent legal scholarship emphasizes that algorithmic bias may constitute a form of indirect discrimination, in which ostensibly neutral systems produce disproportionate adverse effects on protected groups. (Xenidis & Senden, 2023) The United Nations has further stressed that AI systems must be evaluated in light of international human rights standards, particularly concerning equality and non-discrimination in digital governance.

Within the Indonesian legal framework, the constitutional guarantees of equality under Articles 27(1) and 28D(1) of the 1945 Constitution, as well as Law No. 39 of 1999 on Human Rights, provide a normative foundation for addressing discriminatory outcomes arising from digital technologies. Accordingly, the deployment of AI systems does not diminish the obligation to uphold equal treatment before the law but rather reinforces the need for regulatory vigilance in technologically mediated decision-making.

A central characteristic of algorithmic bias lies in the reliance of AI systems on historical data. Empirical studies show that machine learning models trained on biased or unrepresentative datasets tend to replicate and amplify existing inequalities. Recent research in data science has demonstrated that bias can persist even when sensitive attributes are excluded because of the presence of proxy variables and latent correlations within datasets.(Corbett-Davies & Goel, 2018)

Another defining feature of algorithmic bias relates to model design and feature selection. Contemporary research highlights that technical decisions—such as variable selection, model optimization, and evaluation metrics—can significantly influence the distribution of outcomes across demographic groups. Studies in algorithmic fairness show that inherent trade-offs between predictive accuracy and fairness may lead to structurally unequal outcomes, requiring explicit normative choices in system design(Chouldechova & Roth, 2020)

Algorithmic opacity is an additional structural concern. Advanced AI systems, particularly those based on deep learning architectures, often operate as "black boxes," making their decision-making processes difficult to interpret and understand. Recent research on AI governance underscores that limited explainability undermines accountability, due process, and the ability of affected individuals to challenge adverse decisions. (Ananny & Crawford, 2018) International standards, including recent AI governance frameworks, emphasize explainability and transparency as essential components of trustworthy AI systems (Ananny & Crawford, 2018).(ISO/IEC 2023)

These characteristics demonstrate that algorithmic discrimination fundamentally differs from traditional forms of discrimination. Rather than arising from explicit intent, algorithmic bias often manifests as complex and indirect processes embedded within sociotechnical systems. Contemporary legal scholarship increasingly conceptualizes this as a form of systemic or structural discrimination, requiring adaptive legal frameworks and interdisciplinary regulatory approaches.(Zarsky, 2022)

From a theoretical perspective, the implications of algorithmic bias can be analyzed using modern theories of justice. Recent applications of capability theory in digital contexts emphasize that equality should

be assessed in terms of individuals' real opportunities to access resources and participate in society rather than formal equality alone.(Robeyns, 2017) Similarly, critical studies on digital governance highlight how automated systems may reinforce existing power asymmetries, particularly affecting marginalized groups, thereby necessitating a substantive rather than formal approach to equality in the age of AI.(Couldry & Mejias, 2019)

Legal Regulation in Preventing Algorithmic Bias in Artificial Intelligence Systems in Indonesia

The rapid proliferation of Artificial Intelligence (AI) technologies has intensified scholarly and regulatory attention toward governance frameworks aimed at mitigating the risks of algorithmic bias. Contemporary comparative legal research demonstrates a growing shift toward rights- and risk-based regulatory models that seek to align technological innovation with the protection of fundamental rights.(Hacker et al., 2023)

In the European context, the Artificial Intelligence Act (2024) establishes a comprehensive regulatory regime grounded in a risk-based classification of AI systems. Under this framework, high-risk AI applications—particularly those deployed in employment, education, and law enforcement—are subject to stringent obligations, including requirements related to data quality, transparency, human oversight, and ex ante conformity assessments. Recent legal analyses have emphasized that such regulatory mechanisms are essential to ensure compliance with fundamental rights, particularly equality and non-discrimination, in automated decision-making systems.(Madiega, 2024)

Empirical research further underscores the importance of governance mechanisms, such as algorithmic auditing and evaluation. Recent studies on AI accountability have demonstrated that structured auditing processes can effectively identify sources of bias and assess the fairness of machine learning systems prior to deployment, thereby reducing discriminatory risks in practice.(Mökander, 2023) In addition, emerging research highlights the critical role of documentation practices, including model cards and dataset transparency frameworks, in enhancing accountability and enabling external scrutiny of algorithmic systems.(Kapoor & Narayanan, 2023)

In Indonesia, the regulatory framework governing digital technologies is largely principle-based and sectorally fragmented. Law No. 11 of 2008 on Electronic Information and Transactions, as amended by Law No. 1 of 2024, establishes general obligations for electronic system operators to ensure system reliability, security, and accountability. Law No. 27 of 2022 on Personal Data Protection introduces the foundational principles of lawful, fair, and transparent data processing. While these legal instruments provide an important normative baseline, they do not explicitly address the specific risks associated with algorithmic bias or automated decision-making systems.

Recent scholarship on AI governance in developing and Southeast Asian contexts indicates that reliance on general legal principles alone is insufficient to address the technical and institutional complexities of AI systems in healthcare. Studies emphasize the need for context-specific regulatory approaches that incorporate technical standards, risk assessment tools, and enforceable accountability mechanisms, particularly in jurisdictions undergoing rapid digital transformation.(ASEAN, 2024; Gurumurthy & Bharthur, 2022)

To address these regulatory gaps, contemporary research proposes several key governance instruments. First, algorithmic impact assessments are increasingly recognized as an ex ante regulatory tool to evaluate potential harms, including discriminatory effects, prior to system deployment.(Institute, 2023) Second, independent auditing frameworks are essential to ensure ongoing compliance with fairness and accountability standards in operational systems. Third, transparency mechanisms—such as standardized documentation of datasets, model design, and system limitations—are critical for enabling oversight and public accountability. Finally, the establishment of specialized regulatory bodies equipped with technical expertise is necessary to bridge the gap between legal norms and technological implementation.

In conclusion, research-based evidence demonstrates that algorithmic bias presents a significant challenge to the realization of equality in digital society. Addressing this challenge requires not only the interpretation and application of existing legal principles but also the development of specialized regulatory frameworks grounded in interdisciplinary research. In the Indonesian context, the integration of human

rights norms, data governance principles, and technical accountability mechanisms is essential to ensure that AI systems operate consistently with the principles of equality and non-discrimination.

CONCLUSION

Artificial intelligence systems in digital decision-making processes have significant potential to reproduce social inequalities, as they are developed using datasets containing historical bias or imbalanced representation. Machine learning mechanisms that rely on training data may generate algorithmic decisions that indirectly result in differential treatment of certain individuals or groups. This condition demonstrates that algorithmic bias may give rise to forms of digital discrimination that are incompatible with the principle of equality before the law, as guaranteed under Article 27(1) of the 1945 Constitution of the Republic of Indonesia, as well as the principle of non-discrimination enshrined in Law No. 39 of 1999 on Human Rights.

Simultaneously, the increasing deployment of artificial intelligence necessitates the strengthening of regulatory frameworks capable of anticipating the risks of algorithmic discrimination. Although the principle of responsibility for electronic system operators is regulated under the Electronic Information and Transactions Law and data processing accountability is governed under the Personal Data Protection Law, the complexity of artificial intelligence technologies indicates the need for more specific regulatory provisions. In particular, legal frameworks must address algorithmic transparency, bias testing, and independent auditing mechanisms for technological systems that may affect fundamental human rights. Through such regulatory developments, the utilization of artificial intelligence can remain aligned with the principles of justice, equality, and human rights protection within the Indonesian legal system.

ACKNOWLEDGEMENT

The author expresses sincere gratitude to Universitas Pendidikan Nasional as the affiliated institution for providing academic support and facilities throughout the research and manuscript preparation. Appreciation is also extended to all parties who have contributed through assistance, constructive input, and other forms of support, which enabled the completion of this study.

REFERENCES

Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal. *New Media & Society*, *20*(3), 973–989. https://doi.org/10.1177/1461444816676645

Ansari, T., & Negara, S. (2023). Normative Legal Research in Indonesia: Its Origin and Approaches. *ACLJ*, *4*, 1–9.

ASEAN. (2024). *ASEAN guide on AI governance and ethics*. https://asean.org

Chouldechova, A., & Roth, A. (2020). *The frontiers of fairness in machine learning*. https://arxiv.org/abs/1810.08810

Corbett-Davies, S., & Goel, S. (2018). *The measure and mismeasure of fairness*. https://arxiv.org/abs/1808.00023

Couldry, N., & Mejias, U. A. (2019). *The costs of connection: How data is colonizing human life*. Stanford University Press.

Dyzenhaus, D., Hunt, M., and Taggart, M. (2023). The unity of public law and the role of statutory interpretation. *Oxford Journal of Legal Studies*. https://doi.org/10.1093/ojls/gqad012

Europe, C. of. (2024). *Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*. https://www.coe.int

Ferrara, E. (2023). Fairness and bias in artificial intelligence: A survey of sources, impacts and mitigation strategies. In *ACM Transactions on Intelligent Systems and Technology* .

Flick, U. (2022). *An introduction to qualitative research*.

Friedler, S. A., Scheidegger, C., & Venkatasubramanian, S. (2021). The impossibility of fairness. In *Communications of the ACM* (vol. 64, Issue 4).

Garner, B. A. (2019). *Black's law dictionary*.

Gurumurthy, A., & Bharthur, D. (2022). *AI in developing countries: Governance challenges and opportunities*.

Hacker, P., Engel, A., & Mauer, M. (2023). Regulating AI systems: Risk, fairness and fundamental rights. *European Journal of Risk Regulation*. https://doi.org/10.1017/err.2023

Institute A. L. (2023). *Algorithmic impact assessments: A practical framework*. Ada Lovelace Institute. https://www.adalovelaceinstitute.org

ISO/IEC. (2023). *Artificial intelligence—Risk management (ISO/IEC 23894:2023)*.

Kapoor, S., & Narayanan, A. (2023). Leakage and reproducibility crisis in ML-based science. *Patterns*. https://doi.org/10.1016/j.patter.2023

Kordzadeh, N., & Ghasemaghaei, M. (2022). Algorithmic bias: A review and research agenda. *Information Systems Frontiers*.

Madiega, T. (2024). *EU Artificial Intelligence Act: Overview and implications*.

Mökander, J. (2023). Auditing AI: A systematic review of algorithmic auditing. In *AI and Ethics*. https://doi.org/10.1007/s43681-023

United Nations (2024). *Governing AI for humanity: Final report*.

OECD. (2024). *OECD AI policy observatory*. https://oecd.ai

Raji, I. D., & Buolamwini, J. (2022). *Actionable auditing: Investigating the impact of public audits* (Issue ACM Conference on Fairness, Accountability and Transparency).

Robeyns, I. (2017). *Wellbeing, freedom and social justice: The capability approach re-examined*.

Suresh, H., & Guttag, J. V. (2021). *A framework for understanding sources of harm throughout the machine learning lifecycle* (Issue ACM Conference on Equity and Access in Algorithms).

Taekema, S. (2021). Theoretical and normative frameworks for legal research. *Law and Method*, *2021*, 1–18. https://doi.org/10.5553/REM/.000065

Xenidis, R., & Senden, L. (2023). EU Non-Discrimination Law in the Era of Artificial Intelligence. *Common Market Law Review* .

Zarsky, T. (2022). Privacy and data protection in algorithmic decision-making In *Columbia Law Review*.