International Journal of Humanities, Social Sciences and Business (INJOSS) e-ISSN: 2962-6781
Vol. 5 No. 1 January 2026, page 11-25

AI-GOVERNANCE AND ALGORITHMIC ACCOUNTABILITY: RETHINKING LEGAL STANDARDS IN THE AGE OF AUTONOMOUS DECISION-MAKING

Loso Judijanto*[1]
IPOSS Jakarta, Indonesia
Email: losojudijantobumn@gmail.com

Rustiyana Rustiyana
Universitas Bale Bandung, Indonesia
Email: rustiyana.unibba@gmail.com

Rusdi
Universitas Muhammadiyah Sorong, Indonesia
Email: rusdi@um-sorong.ac.id

Abstract

The rapid development of artificial intelligence has driven the use of autonomous decision-making systems in various sectors, including public administration, finance, healthcare, and law enforcement. While AI offers greater efficiency and objectivity, its application also poses serious challenges related to algorithmic governance and accountability, particularly when the resulting decisions significantly impact individual rights and obligations. This research aims to critically examine the existing legal framework and regulatory standards addressing the phenomenon of AI-based autonomous decision-making, and to evaluate the extent to which the principles of accountability, transparency, fairness, and legal responsibility can be applied to algorithmic systems. The research method used is a literature review, examining academic sources, international regulations, public policies, and reports from global institutions relevant to AI governance and algorithmic accountability. The results show that traditional legal standards still face limitations in accommodating the complex, adaptive, and often opaque (black box) characteristics of AI. Therefore, this research emphasizes the need for a new regulatory approach that is adaptive, risk-based, and ethically oriented to ensure that the use of AI remains aligned with human rights protection, legal certainty, and public trust in the era of autonomous decision-making.

Keywords: AI governance, algorithmic accountability, autonomous decision-making, legal standards

[1] Correspondence author

INTRODUCTION

The development of artificial intelligence (AI) has brought about fundamental changes in decision-making across various sectors of social, economic, and governmental life. AI systems no longer merely function as analytical tools but have evolved into entities capable of autonomous decision-making through complex algorithms, machine learning, and large-scale data processing. The use of AI is now widespread in finance, healthcare, law enforcement, public administration, employment, and even citizen data management. This phenomenon marks a significant shift from human-based decision-making to autonomous decision-making systems, which in turn raises increasingly complex legal, ethical, and governance challenges (Sannerholm, 2022).

In the modern legal context, the principle of accountability is a key pillar in the exercise of power and decision-making, both by the state and private entities. Accountability requires clarity regarding who is responsible for a decision, how the decision was made, and what mechanisms are available to assess, correct, or challenge the decision if it causes harm. However, the emergence of autonomous, black-box, and probabilistic AI systems has challenged the concept of legal accountability, which has traditionally been built on the assumption of rational human actors who can be held directly accountable. When decisions are made by algorithms that are not fully explainable, the question of who is responsible—the developer, the system provider, the user, or even the state—becomes increasingly blurred (Gladwin, 2024).

This issue is further complicated because AI algorithms are often trained using historical data loaded with social, economic, and political biases. As a result, decisions made by AI systems have the potential to reproduce or even reinforce structural injustice, discrimination, and social inequality. Various international studies and cases have shown that the use of AI in employment recruitment, criminal risk assessment, credit granting, and determining access to public services can result in decisions that systematically disadvantage certain groups (Bullock et al., 2024). In such situations, existing legal frameworks are often inadequate to guarantee human rights protection, as applicable standards of proof, transparency, and accountability have not been designed to address the technical characteristics of AI systems.

Globally, the discourse on AI governance and algorithmic accountability has become a major concern for policymakers, academics, and international organizations. Various regulatory initiatives have emerged, such as the

European Union's Artificial Intelligence Act, the AI principles developed by the OECD, and UNESCO's AI ethics guidelines. However, these approaches still exhibit significant variation in terms of risk definitions, oversight mechanisms, and accountability standards. This reflects the lack of global consensus on how the law should comprehensively and effectively address the challenges of autonomous decision-making (Areo, 2025).

Furthermore, many national legal systems still rely on traditional regulatory frameworks that focus on individual responsibility, negligence, and clearly identifiable errors. This framework is difficult to apply to AI systems, which are adaptive, dynamic, and often operate through complex networks of actors. When an algorithmic decision results in adverse legal or social impacts, the process of tracing cause and effect becomes extremely complex, making legal accountability mechanisms weak or even ineffective. This condition has the potential to create an accountability gap, namely a situation where no party can be held adequately accountable for the impact of AI decisions (Cheong, 2024a).

These issues highlight the tension between technological innovation and fundamental legal principles, such as legal certainty, justice, and the protection of individual rights. On the one hand, AI offers efficiency, consistency, and analytical capabilities that surpass human capacity. On the other hand, excessive reliance on algorithmic systems without a robust governance framework can erode public trust in legal and government institutions (Cheong, 2024a). Therefore, AI governance is not only concerned with regulating the technical use of technology, but also concerns how legal and ethical values can be effectively integrated into the design, implementation, and oversight of AI systems.

In this context, algorithmic accountability is a key concept that demands transparency, clarity of responsibility, and adequate evaluation and oversight mechanisms for AI systems. However, this concept still faces conceptual and practical challenges, particularly related to the limited explainability of algorithms, the protection of trade secrets, and technical complexity that is difficult for law enforcement officials and the general public to understand. As a result, existing legal standards often lag behind the pace of AI technological development, creating a gap between legal norms and practical reality (Rico, 2024).

Based on these conditions, efforts are needed to review and reformulate relevant legal standards in the era of autonomous decision-making. Rethinking legal standards is crucial so that the law is not merely reactive to the negative

impacts of AI, but also proactive in shaping responsible, inclusive governance oriented toward protecting human rights. This review encompasses fundamental questions regarding the definition of legal responsibility, the role of the state in AI oversight, and the balance between innovation and regulation. This research is relevant because it seeks to critically examine the relationship between AI governance and algorithmic accountability in the context of a shifting decision-making paradigm. Using a literature review approach, this research seeks to identify the weaknesses and limitations of existing legal standards, while also exploring various normative and conceptual approaches developed in the legal and public policy literature. Through this analysis, this research is expected to provide theoretical contributions to enriching the discourse on AI governance, as well as practical contributions to policymakers in designing a regulatory framework that is more adaptive and responsive to the challenges of the artificial intelligence era.

Ultimately, this background confirms that AI governance and algorithmic accountability are not merely technical or sectoral policy issues, but rather fundamental issues that touch on the very core of legal legitimacy and justice in modern society. Without adequate updates to legal standards, the use of AI has the potential to create legal uncertainty and new, difficult-to-control injustices. Therefore, research on reviewing legal standards in the era of autonomous decision-making is crucial and urgent to ensure that the development of AI technology remains aligned with legal principles and human values.

RESEARCH METHOD

This research uses a literature review method with a normative qualitative approach to analyze the concepts of AI governance and algorithmic accountability in the context of autonomous decision-making. The literature review was conducted by systematically reviewing various relevant scientific sources, including reputable international journal articles, academic books, reports from international institutions, and policy documents and legal regulations related to artificial intelligence. These sources were selected based on their thematic relevance to the issues of AI governance, the legal responsibility of algorithms, the transparency of automated systems, and the development of legal standards in various jurisdictions. The literature search was conducted through academic databases such as Scopus, Web of Science, and Google Scholar using keywords related to AI governance, algorithmic accountability, autonomous decision-making, and legal standards.

Data analysis was conducted using content analysis and conceptual analysis techniques to identify patterns of thought, theoretical frameworks, and key debates within the existing literature. Each source was analyzed to understand how traditional legal standards are questioned, adapted, or reconstructed in the face of complex, autonomous AI-based decision-making systems. The results of the analysis are then critically synthesized to formulate a comprehensive understanding of the challenges and opportunities for strengthening algorithmic accountability, while also offering conceptual reflections on the need to update the legal framework to align with the development of artificial intelligence technology in the digital era.

RESULT AND DISCUSSION

Definition and Scope of Algorithmic Accountability in AI-Based Systems

Algorithmic accountability is an increasingly central concept in the discourse on artificial intelligence governance, as AI-based systems are increasingly used in decision-making that significantly impact individuals and society. Conceptually, algorithmic accountability can be defined as the principles and mechanisms that require algorithmic systems, particularly autonomous or semi-autonomous ones, to be held ethically, legally, and socially accountable for their decisions and the resulting impacts (Atoum, 2025). Accountability in this context relates not only to the ability to explain how a decision was made but also to establishing responsibility when an AI system causes errors, bias, discrimination, or other harm. Thus, algorithmic accountability positions technology not as a neutral entity, but as a social product designed, trained, and operated within a framework of specific values, interests, and power relations.

In AI-based systems, algorithms no longer function simply as simple mathematical instructions but as complex structures that learn from vast amounts of data and dynamically adjust their behavior. These characteristics pose fundamental challenges to traditional concepts of accountability, which generally rely on a clear causal relationship between human actions and their outcomes. When decisions are generated by "black box" machine learning models, the system's internal reasoning processes are often difficult to understand, even for the developers themselves (Hassen, 2025). Therefore, the definition of algorithmic accountability has evolved beyond demands for purely technical transparency to a more holistic approach encompassing clarity of the system's objectives, the legitimacy of data use, and the compliance of decision outcomes with applicable legal and ethical norms.

The scope of algorithmic accountability in AI-based systems encompasses the entire technology lifecycle, from design and development to implementation to post-deployment evaluation. At the design stage, accountability relates to how the system's objectives are formulated and what values are integrated into the algorithm's design. Initial decisions regarding the variables used, prioritized performance indicators, and the underlying assumptions of the model have long-term implications for how the system operates and its impacts (Boch et al., 2022). Therefore, the scope of accountability cannot be limited solely to algorithmic outputs but must encompass the normative processes that precede the creation of the AI system itself.

During the model development and training stages, algorithmic accountability encompasses the management of the data used as the basis for the system's learning. Data is not an objective representation of reality, but rather a reflection of existing social structures, historical biases, and inequalities. When biased data is used to train AI systems, algorithms have the potential to reproduce and even reinforce existing discrimination (Cheong, 2024b). In this context, accountability demands responsibility for the selection of data sources, data cleaning methods, and the evaluation of potential biases and discriminatory impacts. Thus, the scope of algorithmic accountability extends to issues of data fairness and epistemic responsibility in AI-based knowledge production.

The implementation of AI systems in real-world contexts expands the scope of accountability into the institutional and legal realms. When algorithms are used in the public and private sectors, such as in credit assessments, employment selection, law enforcement, or healthcare, the resulting decisions can impact individuals' fundamental rights. In this context, algorithmic accountability is closely linked to oversight mechanisms, algorithmic audits, and the rights of users or citizens to obtain explanations and raise objections to decisions made by AI systems. Accountability is not only internal to the organizations developing or using the technology, but also external through the roles of regulators, oversight bodies, and civil society (Nuredin, 2024).

Furthermore, the scope of algorithmic accountability also encompasses the relational dimension between humans and AI systems. Although decisions are generated by algorithms, moral and legal responsibility cannot be entirely transferred to machines. Therefore, the concept of accountability demands clarity regarding the distribution of responsibility among the various actors involved, including software developers, data providers, user organizations,

and policymakers. In complex and distributed AI-based systems, the assignment of responsibility is often blurred, necessitating an accountability framework capable of identifying points of human intervention and ensuring corrective mechanisms when the system fails to function fairly and accurately.

In a global context, algorithmic accountability also has a transnational scope because many AI systems are developed and operated across national borders. This raises challenges related to differing legal, ethical, and cultural standards in interpreting responsibility and fairness. Therefore, algorithmic accountability is not merely a technical or juridical issue, but also a global governance issue that demands harmonization of international principles and standards (Kumar et al., 2025). Thus, its scope includes collaborative efforts between states, international organizations, and non-state actors to formulate common norms for the responsible development and use of AI.

The Challenge of Algorithmic Transparency in Black-Box Systems

Algorithmic transparency is one of the most crucial issues in the development of modern artificial intelligence systems, especially when the algorithms used are black-box. Black-box systems refer to algorithmic systems whose decision-making processes cannot be directly understood by humans, either due to model complexity, proprietary nature, or limited access to the algorithm's internal structure (Chaudhary, 2024). In the context of the increasingly widespread use of AI in various strategic sectors such as finance, healthcare, education, law, and public administration, this closedness poses serious challenges to the principles of accountability, fairness, and public trust. Algorithmic transparency is both a normative and technical requirement so that system-generated decisions can be accounted for, monitored, and critically evaluated.

One of the main challenges to algorithmic transparency in black-box systems lies in the technical complexity of the AI models themselves. Advanced machine learning models such as deep neural networks operate through millions to billions of parameters that are nonlinearly interconnected (von Eschenbach, 2021). This structure makes the causal relationship between inputs and outputs extremely difficult to trace, even for their developers. When a system provides credit recommendations, medical diagnoses, or criminal risk predictions, the explanation of how those decisions are made often cannot be simplified without losing substantive meaning. This complexity creates a gap between the need for users and regulators to understand the system and the

technical realities of increasingly sophisticated and abstract algorithmic designs.

Beyond technical complexity, transparency challenges also arise from the conflict between openness and commercial interests. Many black-box algorithms are developed by technology companies with significant research investments and are protected by intellectual property rights. Full disclosure of the source code or internal logic of algorithms is often perceived as risky to competitive advantage and business security. As a result, companies tend to limit the information accessible to the public or regulators, even when the algorithms are used for decision-making that directly impacts the rights and interests of individuals. This situation creates a dilemma between protecting innovation and the need for transparency in the public interest (Thalpage, 2023).

The challenges of algorithmic transparency are also closely related to the limitations of the concepts and explanatory standards used. Transparency does not necessarily mean disclosing all code or mathematical models, but rather providing explanations that are understandable to different stakeholders (Gryz & Rojszczak, 2021). However, there is still no global consensus on the form and level of explanation considered adequate. Overly in-depth technical explanations may only be understandable to experts, while overly simplistic explanations risk misleading or misrepresenting the true decision mechanisms. The tension between explainability and accuracy is a fundamental issue in achieving meaningful transparency.

Furthermore, algorithmic transparency in black-box systems faces contextual and social challenges. Algorithms do not operate in a vacuum but are built on historical data rife with structural biases and social inequalities. When algorithms are opaque, biases embedded in the data and models become increasingly difficult to identify and correct. This can reinforce systemic discrimination, for example in credit assessments, job recruitment, or public oversight systems. Algorithmic opacity makes it difficult for disadvantaged groups to challenge or challenge decisions that disadvantage them, thereby weakening the principle of procedural fairness.

From a legal and governance perspective, algorithmic transparency in black-box systems also poses complex regulatory challenges. Many traditional legal frameworks are based on the assumption that decision-making can be traced to a distinct human actor. In autonomous algorithmic systems, responsibility becomes distributed among developers, data providers, system operators, and end users. Without adequate transparency, determining legal

responsibility for errors or losses becomes unclear. Regulations that require explainability of decision-making processes often encounter technical limitations in requiring black-box systems to be fully explainable.

Another challenge relates to institutional capacity and technological literacy. Algorithmic transparency depends not only on the availability of information but also on the ability of those receiving that information to understand it. Regulators, auditors, and oversight bodies often lack the resources and technical expertise to evaluate complex AI systems. Consequently, even if some information is disclosed, the resulting transparency is superficial and ineffective in oversight practices. This capacity gap undermines the primary purpose of transparency as a tool of control and accountability (Vorras & Mitrou, 2021).

Ultimately, the challenge of algorithmic transparency in black-box systems reflects a fundamental tension between technological innovation and normative values in democratic societies. On the one hand, black-box systems offer efficiency, accuracy, and analytical capabilities that surpass human capacity. On the other hand, the closed nature of these systems threatens the principles of openness, fairness, and the protection of individual rights. Achieving algorithmic transparency cannot be achieved solely through technical approaches, but requires the integration of responsible technology design, adaptive legal frameworks, and increased public literacy and participation. Thus, algorithmic transparency is not only a technological challenge but also an ethical, social, and institutional one that will shape the future direction of AI governance.

Algorithmic Bias and Its Impact on Human Rights and Substantive Justice

The rapid development of artificial intelligence and algorithm-based decision-making systems has brought fundamental changes to various aspects of social, economic, and legal life. Algorithms are now used to determine creditworthiness, in the workforce recruitment process, in criminal risk assessments, in the distribution of social assistance, and in the prioritization of public services. Although these technologies are often promoted as objective, efficient, and free from human subjectivity, the reality is that algorithms are never completely neutral. Algorithms are developed, trained, and implemented within specific social contexts that are fraught with values, assumptions, and structural inequalities (Green, 2022). It is in this context that algorithmic bias becomes a crucial issue, particularly when its impacts directly impact human rights and principles of substantive justice.

Algorithmic bias refers to the tendency of algorithm-based systems to produce decisions that systematically disadvantage certain individuals or groups. This bias can arise from various sources, such as unrepresentative training data, model designs containing discriminatory assumptions, or implementation contexts that ignore social complexity. The historical data used to train algorithms often reflects past discriminatory practices, potentially replicating and even reinforcing existing inequities. When such bias is internalized in automated systems, discrimination is no longer visible but hidden behind technical processes that are difficult for the public to understand (Land & Aronson, 2020).

The impact of algorithmic bias on human rights becomes even more significant when algorithms are used in decision-making that affects fundamental rights. The right to equality before the law, the right to non-discrimination, the right to employment, education, and social welfare can be threatened when individuals are assessed solely on statistical patterns that are insensitive to personal and social circumstances. For example, in automated recruitment systems, algorithms trained on historical company data can indirectly exclude candidates from certain groups due to biased past patterns. As a result, the right to a fair opportunity is undermined by the logic of algorithmic efficiency (Clarissa Hennig Leal & Soares Crestane, 2023).

Furthermore, the use of algorithms in the criminal justice system raises serious human rights concerns. Risk assessment systems used to determine bail, sentencing, or parole often exhibit bias against minority groups or marginalized communities. When algorithms associate certain social factors with a higher risk of crime, individuals from those groups may be treated more harshly without considering their individual context. This not only violates the principle of the presumption of innocence but also threatens the rights to liberty and fair treatment (Balakrishnan, 2024; Binns, 2022).

In the context of substantive justice, algorithmic bias challenges traditional understandings of justice, which rely solely on formal equal treatment. Substantive justice emphasizes fair outcomes by considering real conditions, structural vulnerabilities, and social inequalities. However, algorithms tend to operate based on statistical generalizations that ignore the complexity of human experience. Thus, the application of biased algorithms has the potential to widen social inequalities and produce injustices legitimized by technology.

Substantive justice also demands corrective mechanisms when decisions are proven to be detrimental to certain groups. In algorithm-based systems,

these mechanisms are often difficult to implement due to the closed or "black box" nature of algorithms. This lack of transparency makes it difficult for individuals to understand the basis of decisions that affect their lives, let alone challenge them legally. This situation potentially violates the right to information, the right to an effective legal remedy, and the principle of accountability in a state of law. Furthermore, algorithmic bias also raises issues of moral and legal responsibility. When discriminatory decisions are made by automated systems, it becomes unclear who should be held accountable—the algorithm developer, the data provider, the user institution, or the regulator. This lack of clarity risks creating a vacuum of accountability that ultimately harms victims of algorithmic bias. From a human rights perspective, this situation contradicts the obligations of states and non-state actors to respect, protect, and fulfill the fundamental rights of citizens.

Efforts to address algorithmic bias require a multidisciplinary approach that integrates technological, legal, ethical, and social science perspectives. From a technical perspective, developing fairer algorithms requires inclusive data, systematic bias testing methods, and model design that considers social impacts. However, a purely technical approach is insufficient without a robust legal and ethical framework. Regulations oriented toward protecting human rights need to ensure that the use of algorithms adheres to the principles of transparency, accountability, and non-discrimination (Balakrishnan, 2024).

Within a substantive justice framework, it is crucial to place humans at the center of AI-based decision-making systems. Algorithms should function as tools, not as final arbiters that ignore contextual assessments and human values. Meaningful human oversight is a key element in ensuring that algorithmic decisions can be corrected when they conflict with principles of justice and human rights. Without such oversight, the risk of dehumanization in the decision-making process will increase.

Ultimately, algorithmic bias is not simply a technical issue, but a reflection of the social structures and values inherent in society. Its impact on human rights and substantive justice demands a comprehensive and sustained response. If not addressed seriously, algorithms have the potential to become new instruments that reinforce old injustices in more subtle and difficult-to-question forms. Conversely, with proper governance, algorithms can also be directed to support substantive justice and the protection of human rights in an era of autonomous decision-making.

Legal Liability for Autonomous Decisions: Who Is Responsible?

The development of artificial intelligence and autonomous systems has brought about fundamental changes in the way decisions are made and executed in various sectors of life. Autonomous systems now function not only as human aids but also as capable of making decisions independently based on data processing, machine learning, and complex algorithms. This phenomenon has serious legal implications, particularly when these autonomous decisions result in harm, rights violations, or significant social impacts (Yazdanpanah et al., 2023). In this context, a fundamental question arises in modern law: who should be held responsible for decisions made by autonomous systems in the absence of direct human intervention.

Traditionally, the concept of legal liability is based on the assumption that the legal subject acting is a human or a legal entity controlled by a human. The elements of fault, negligence, and intent are the primary basis for determining liability (Yazdanpanah et al., 2023). However, autonomous systems based on artificial intelligence operate through processes that are not always fully predictable by their developers or users. Machine learning algorithms, for example, can produce decisions that differ from the initial designed scenario, as the system dynamically learns from new data. This situation challenges the conventional legal paradigm that associates direct actions with personally liable subjects.

In the face of this complexity, legal approaches often attempt to trace the chain of responsibility back to the human actors involved in the lifecycle of autonomous systems. Software developers and algorithm designers are often viewed as potentially liable parties, particularly if harm arises from design flaws, algorithmic bias, or negligence in system testing (Yazdanpanah et al., 2021). Within this framework, legal liability can be constructed as a form of product liability, where autonomous systems are treated as products that must meet certain safety and precautionary standards. If the system fails to function as intended and causes harm, the developer can be held liable.

Conversely, users or operators of autonomous systems are also often held liable, particularly when they have operational control or the authority to activate, deactivate, or supervise the system's use (Gunkel, 2020). In this context, the law tends to examine whether the user has acted negligently, for example by using the system beyond its intended purpose, ignoring usage guidelines, or failing to exercise reasonable oversight. However, this approach becomes problematic when the system operates truly autonomously without

significant human intervention, making it difficult to prove negligence or fault on the part of the user.

In addition to developers and users, corporate entities as owners or providers of autonomous systems are also often subject to legal liability. In many cases, artificial intelligence systems are developed and operated within an organization or company, so legal responsibility can be attributed to these legal entities. This approach aligns with the principle of vicarious liability or corporate liability, where companies are responsible for the actions and decisions resulting from the technology they manage and utilize. This approach is considered more realistic in the context of law enforcement, as companies have the financial and structural capacity to assume risks and implement internal control mechanisms (Jedličková, 2025).

However, the greater the degree of autonomy of a system, the greater the gap between its actions and human control. This has given rise to discourse about the possibility of recognizing artificial intelligence systems as limited legal subjects, or electronic personhood. This idea aims to create a framework of responsibility more in line with technological realities, in which autonomous systems can be treated as entities with specific legal obligations. However, this idea remains widely debated, raising ethical and philosophical issues related to the attribution of fault, moral capacity, and the legitimacy of granting legal status to non-human entities.

In practice, the lack of clarity regarding the subject of legal responsibility for autonomous decisions has the potential to create a legal vacuum and weaken protection for victims. Therefore, many contemporary legal approaches emphasize the importance of the principles of accountability and due care in the development and implementation of autonomous systems. This principle encourages a proportional division of responsibility among the various actors involved, from designers and developers to data providers and end users (Buiten et al., 2023). Thus, legal responsibility is not concentrated on a single party but is distributed according to the role and level of control of each actor. Ultimately, the issue of legal liability for autonomous decisions reflects the significant challenge of adapting legal systems to rapid technological developments. The law is required to maintain certainty, justice, and rights protection, without hindering technological innovation. Therefore, an adaptive and responsive legal framework is needed that accommodates the unique characteristics of autonomous systems while simultaneously emphasizing that every technological decision must always be accompanied by a clear accountability mechanism. With this approach, the law can act as a regulatory

instrument, ensuring that the use of artificial intelligence and autonomous systems is carried out responsibly and in accordance with the values of social justice.

CONCLUSION

This study concludes that the development of increasingly autonomous artificial intelligence systems has challenged traditional legal foundations, which have centered on human actors as the primary subject of accountability. AI-governance and algorithmic accountability have emerged as an urgent need to bridge the gap between technology's autonomous decision-making capabilities and a less-than-adaptive legal framework. Conventional legal standards such as fault, intent, and causality become difficult to apply when decisions are generated by complex, adaptive, and often opaque systems. Therefore, this study emphasizes the importance of shifting the legal paradigm toward a more functional and preventative approach, emphasizing the principles of transparency, explainability, fairness, and multi-layered oversight mechanisms throughout the algorithm's lifecycle.

Furthermore, the literature review indicates that strengthening AI-governance serves not only as a risk control instrument but also as a means to maintain legal legitimacy and public trust in the use of AI in both public and private spheres. Rethinking legal standards in the era of autonomous decision-making requires the integration of risk-based regulations, a clear division of responsibilities between AI developers, providers, and users, and cross-jurisdictional harmonization to address the global nature of this technology. Thus, this study concludes that the future of algorithmic accountability cannot rely on a single legal approach, but rather requires a legal framework that is dynamic, interdisciplinary, and responsive to the evolution of artificial intelligence technology.

REFERENCES

Areo, G. (2025). *AI Governance: The Implications of Autonomous Decision- Makers in Government*.

Atoum, I. (2025). Revolutionizing AI Governance: Addressing Bias and Ensuring Accountability Through the Holistic AI Governance Framework. *International Journal of Advanced Computer Science and Applications*, *16*. https://doi.org/10.14569/IJACSA.2025.0160283

Balakrishnan, A. (2024). *ETHICAL AND LEGAL IMPLICATIONS OF AI JUDGES: BALANCING EFFICIENCY AND THE RIGHT TO FAIR TRIAL* [Master Thesis]. https://studenttheses.uu.nl/handle/20.500.12932/48242

Binns, R. (2022). Human Judgment in algorithmic loops: Individual justice and automated decision-making. *Regulation & Governance*, *16*(1), 197–211. https://doi.org/10.1111/rego.12358

Boch, A., Hohma, E., & Trauth, R. (2022). *Towards an Accountability Framework for AI: Ethical and Legal Considerations*. https://doi.org/10.13140/RG.2.2.10231.50086

Buiten, M., de Streel, A., & Peitz, M. (2023). The law and economics of AI liability. *Computer Law & Security Review*, *48*, 105794. https://doi.org/10.1016/j.clsr.2023.105794

Bullock, J. B., Chen, Y.-C., Himmelreich, J., Hudson, V. M., Korinek, A., Young, M. M., & Zhang, B. (2024). *The Oxford Handbook of AI Governance*. Oxford University Press.

Chaudhary, G. (2024). Unveiling the Black Box: Bringing Algorithmic Transparency to AI. *Masaryk University Journal of Law and Technology*, *18*(1), 93–122.

Cheong, B. C. (2024a). Transparency and accountability in AI systems: Safeguarding wellbeing in the age of algorithmic decision-making. *Frontiers in Human Dynamics*, *6*. https://doi.org/10.3389/fhumd.2024.1421273

Cheong, B. C. (2024b). Transparency and accountability in AI systems: Safeguarding wellbeing in the age of algorithmic decision-making. *Frontiers in Human Dynamics*, *6*. https://doi.org/10.3389/fhumd.2024.1421273

Clarissa Hennig Leal, M., & Soares Crestane, D. (2023). Algorithmic discrimination as a form of structural discrimination: Standards of the Inter-American Court of Human Rights related to vulnerable groups and the challenges to judicial review related to structural injunctions. *UNIO – EU Law Journal*, *9*(1), 29–44. https://doi.org/10.21814/unio.9.1.5271

Gladwin, O. (2024). *Human Rights in the Age of AI: Establishing Fairness and Accountability through Algorithmic Governance*. https://doi.org/10.13140/RG.2.2.13279.06567

Green, B. (2022). Escaping the Impossibility of Fairness: From Formal to Substantive Algorithmic Fairness. *Philosophy & Technology*, *35*(4), 90. https://doi.org/10.1007/s13347-022-00584-6

Gryz, J., & Rojszczak, M. (2021). Black box algorithms and the rights of individuals: No easy solution to the "explainability" problem. *Internet Policy Review*, *10*(2), 1–24. https://doi.org/10.14763/2021.2.1564

Gunkel, D. J. (2020). Mind the gap: Responsible robotics and the problem of responsibility. *Ethics and Information Technology*, *22*(4), 307–320. https://doi.org/10.1007/s10676-017-9428-2

Hassen, M. Z. (2025). Beyond the algorithm: Applying critical lenses to AI governance and societal change. *AI and Ethics*, *5*(6), 5759–5765. https://doi.org/10.1007/s43681-025-00781-x

Jedličková, A. (2025). Ethical approaches in designing autonomous and intelligent systems: A comprehensive survey towards responsible development. *AI & SOCIETY*, *40*(4), 2703–2716. https://doi.org/10.1007/s00146-024-02040-9

Kumar, A. B., Sanjaya, K., & Saleem, M. (2025). Justice in the Age of Algorithms: Ensuring Transparency, Accountability, and Fairness in AI-Driven Legal Systems. In *Artificial Intelligence in Peace, Justice, and Strong Institutions* (pp. 191–210). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3693-9395-6.ch009

Land, M. K., & Aronson, J. D. (2020). Human Rights and Technology: New Challenges for Justice and Accountability. *Annual Review of Law and Social Science*, *16*(Volume 16, 2020), 223–240. https://doi.org/10.1146/annurev-lawsocsci-060220-081955

Nuredin, A. (2024). *ALGORITHMIC BIAS IN LAW: THE DISCRIMINATORY POTENTIAL AND LEGAL LIABILITY OF AI-BASED DECISION SUPPORT SYSTEMS* (p. 125). https://doi.org/10.55843/ISC2024conf105n

Rico, P. (2024). *AI and Data Governance: A Legal Framework for Algorithmic Accountability and Human Rights*.

Sannerholm, R. (2022). Responsibility and Accountability: AI, Governance, and the Rule of Law. *The Swedish Law and Informatics Research Institute*, 223–246. https://doi.org/10.53292/208f5901.40d940a1

Thalpage, N. (2023). Unlocking the Black Box: Explainable Artificial Intelligence (XAI) for Trust and Transparency in AI Systems. *Journal of Digital Art & Humanities*, *4*(1), 31–36. https://doi.org/10.33847/2712-8148.4.1_4

von Eschenbach, W. J. (2021). Transparency and the Black Box Problem: Why We Do Not Trust AI. *Philosophy & Technology*, *34*(4), 1607–1622. https://doi.org/10.1007/s13347-021-00477-0

Vorras, A., & Mitrou, L. (2021). Unboxing the Black Box of Artificial Intelligence: Algorithmic Transparency and/or a Right to Functional Explainability. In T.-E. Synodinou, P. Jougleux, C. Markou, & T. Prastitou-Merdi (Eds.), *EU Internet Law in the Digital Single Market* (pp. 247–264). Springer International Publishing. https://doi.org/10.1007/978-3-030-69583-5_10

Yazdanpanah, V., Gerding, E. H., Stein, S., Dastani, M., Jonker, C. M., Norman, T. J., & Ramchurn, S. D. (2023). Reasoning about responsibility in autonomous systems: Challenges and opportunities. *AI & SOCIETY*, *38*(4), 1453–1464. https://doi.org/10.1007/s00146-022-01607-8

Yazdanpanah, V., Gerding, E., Stein, S., Dastani, M., Jonker, C. M., & Norman, T. (2021). *Responsibility Research for Trustworthy Autonomous Systems*. 57–62. https://doi.org/10.48448/w5y9-qk13