

JOURNAL Bit-Tech

Binary Digital - Technology
ISSN : 2622-271X | e-ISSN : 2622-2728

PUBLISHED BY:



The Concept of Justice in AI-Driven Legal Decision Making

Elfindah Princes⁽¹⁾ Rasji⁽²⁾ Michael⁽³⁾

- (1) Universitas Tarumanagara
- (2) Universitas Tarumanagara
- (3) Universitas Tarumanagara

DOI: <https://doi.org/10.32877/bt.v8i1.2338>

	0	Total citations
	0	Recent citations
	n/a	Field Citation Ratio
	n/a	Relative Citation Ratio

Keywords:

Algorithmic Bias, Artificial Intelligence , Due Process, Justice , Legal Decision-Making

Abstract

The integration of Artificial Intelligence (AI) into legal decision-making processes



[Pdf](#)

PUBLISHED
2025-08-10



Sertifikat Akreditasi SINTA 4

Spesial Link

- [Home](#)
- [Editorial Board](#)
- [Reviewers](#)
- [Online Submission](#)
- [Author Guideline](#)

The Concept of Justice in AI-Driven Legal Decision Making

Elfindah Princes^{1)*}, Rasji²⁾, Michael³⁾

¹⁾²⁾³⁾ Universitas Tarumanagara, Jakarta, Indonesia

¹⁾²⁾³⁾ Information Systems Management Department, BINUS Graduate Program, Bina Nusantara University, Jakarta, Indonesia

¹⁾elfindah.princes@binus.ac.id

²⁾rasji@fh.untar.ac.id

³⁾michael.207242008@stu.untar.ac.id

Article history:

Received 19 April 2025;
Revised 23 April 2025;
Accepted 28 May 2025;
Available online 10 August 2025

Keywords:

Algorithmic Bias
Artificial Intelligence
Due Process
Justice
Legal Decision-Making

Abstract

The integration of Artificial Intelligence (AI) into legal decision-making processes has introduced significant advancements in efficiency and predictive capability. However, its implications for justice—particularly fairness, impartiality, transparency, and due process—remain critically debated. This study employs a Systematic Literature Review (SLR) methodology to examine how AI-driven legal decision-making aligns with classical and contemporary philosophical concepts of justice. Drawing on 48 peer-reviewed articles, policy documents, and case studies published between 2015 and 2024, the research identifies four core thematic issues: the persistence of algorithmic bias, the lack of transparency in AI systems, inconsistencies in global regulatory frameworks, and the misalignment of AI logic with moral reasoning. While AI offers promising tools for streamlining judicial processes, its application often risks reinforcing existing inequities and undermining legal principles such as corrective justice and procedural fairness. The study concludes with targeted recommendations for the development of transparent, accountable, and ethically governed AI systems that support—rather than supplant—human judicial discretion. This research contributes to the growing discourse on legal AI by highlighting the necessity of embedding justice-oriented values at the core of technological innovation in the legal sector. This research has several limitations: not based on empirical findings and no validations from experts both in AI and in legal theories. Future research should address these limitations.

I. INTRODUCTION

The integration of Artificial Intelligence (AI) in legal decision-making has transformed judicial processes, offering speed, efficiency, and consistency. AI-powered legal systems are increasingly used for predictive analytics, legal research, contract analysis, and even judicial sentencing [1]. However, while AI can process vast amounts of legal data efficiently, concerns regarding its ability to uphold the principles of justice remain paramount [2]. The fundamental question arises: Can AI-driven legal decision-making ensure fairness, impartiality, and human dignity [3] in the administration of justice?

AI's perceived impartiality is often questioned in professional contexts. For example, in HR management, decisions made by AI systems were found to lack perceived dignity and fairness compared to human decisions Bankins et al. [4]

Justice, in legal philosophy, encompasses concepts such as fairness, due process, and equality before the law. Traditionally, justice relies on human judges' ability to interpret laws, exercise discretion, and apply moral reasoning. AI, however, operates on algorithms that may be susceptible to biases embedded in training data, lack moral reasoning, and may not fully comprehend the nuances of human rights and ethical considerations [5]–[7]. This study explores whether AI-driven legal decision-making aligns with the philosophical and practical tenets of justice.

With the increasing reliance on data-driven technologies across various industries, the legal sector is not immune to technological transformation. Courts and legal institutions around the world have begun experimenting with AI applications to enhance procedural efficiency, reduce backlog, and provide standardized legal outcomes.

* Corresponding author

Notable examples include the use of the COMPAS algorithm in the United States for assessing recidivism risks, and the implementation of predictive judicial analytics in several jurisdictions. While these tools promise to alleviate workload and optimize decision-making, they simultaneously introduce new ethical and legal dilemmas.

At the heart of these dilemmas lies the notion of algorithmic justice—a term that refers to the intersection of algorithmic decision-making and legal fairness. The application of AI in legal systems prompts critical questions: What happens when an algorithm makes an incorrect prediction? How can individuals contest or appeal decisions made by an opaque AI system? And perhaps most importantly, can algorithmic outputs be considered just, especially when their logic is not interpretable or morally grounded?

Concerns about algorithmic bias are central to this debate. Scholars and policymakers have pointed out that AI systems often replicate and amplify existing societal biases due to the nature of historical data used for training. For instance, in the criminal justice system, historical records may contain disproportionate targeting of certain demographic groups. When such data feed into AI models, the outcomes may reflect and even exacerbate systemic discrimination. This challenges the principle of equality before the law and undermines trust in legal institutions.

Moreover, transparency and accountability are fundamental components of justice that are often absent in AI-driven systems. Many AI models, especially those based on deep learning, are described as “black boxes” due to their complex and opaque inner workings. This lack of transparency makes it difficult for legal practitioners, defendants, or the public to understand how decisions are reached, raising serious concerns about due process and the right to a fair trial.

The philosophical underpinnings of justice further complicate the application of AI in law. From Aristotle’s concept of corrective and distributive justice to John Rawls’ theory of fairness, the administration of justice requires a nuanced understanding of context, intent, and moral judgment—elements that are difficult, if not impossible, for AI to replicate. Justice is not merely about procedural consistency; it is about equity, empathy, and the human capacity to interpret law in light of ethical and cultural values.

Given these tensions, this study seeks to critically examine the role of AI in legal decision-making by exploring both its potential benefits and its limitations. Specifically, it aims to assess whether AI can serve justice or whether it risks automating injustice under the guise of objectivity and efficiency. The objectives of this research are fourfold: (1) to examine the principles of justice and how they apply to AI-driven legal decision-making; (2) to analyze the benefits and limitations of AI in ensuring fairness and impartiality in legal decisions; (3) to investigate potential risks of algorithmic bias and ethical concerns in AI-based judicial systems; and (4) to propose regulatory and ethical frameworks to enhance justice in AI-driven legal decision-making.

To address these purposes, the study seeks to answer the following questions: How does AI-driven legal decision-making align with traditional concepts of justice? What are the primary advantages and challenges of AI in judicial processes? How can algorithmic bias impact the fairness and impartiality of legal decisions? What ethical and regulatory measures can be implemented to ensure just AI-driven legal decisions? By addressing these questions, this research contributes to the growing discourse on the intersection of AI and law, and underscores the importance of embedding justice-oriented values into technological systems that are poised to shape the future of legal governance.

II. RELATED WORKS/LITERATURE REVIEW

A. Theoretical Foundations of Justice in Legal Decision-Making

Justice, as a foundational principle in legal and moral philosophy, has long been the subject of scholarly debate and interpretation [8]–[10]. Theories of justice offer normative frameworks to determine what is fair and just in the distribution of resources, resolution of disputes, and treatment of individuals under the law. Among the most influential modern theorists, John Rawls introduced the concept of “Justice as Fairness”, which emphasizes that institutions should be arranged to benefit the least advantaged in society [9]. Rawls’ two key principles—equal basic liberties for all and social and economic inequalities arranged to benefit the least well-off—imply that any system, including AI, used in legal decision-making must be designed to avoid perpetuating systemic injustice or discrimination.

On the other hand, Aristotle’s classical theories of Corrective and Distributive Justice provide essential insights into the dual function of justice: correction of wrongs (e.g., in criminal law) and fair distribution of resources and privileges (e.g., in administrative or civil law) [11]. Corrective justice seeks to restore balance between parties after a wrongdoing, while distributive justice concerns the equitable allocation of societal goods and burdens. These frameworks demand an understanding of human circumstances, intent, and proportionality—factors AI systems often fail to fully comprehend due to their reliance on quantitative data over qualitative reasoning.

In the context of AI-driven legal decision-making, these philosophical frameworks serve as critical benchmarks. Ensuring that AI applications embody the values of fairness, equity, and impartiality requires not only technological design but also a deep engagement with these normative concepts. Legal integrity, therefore, depends not merely on efficiency, but on aligning technological interventions with long-standing ethical foundations of justice.

B. AI in Legal Decision-Making

AI applications in the legal sector include predictive analytics, case outcome forecasting, and automated contract review [12], [13]. Similarly, Kabir and Alam [14] highlight the use of NLP and machine learning in tools like ROSS Intelligence and Lex Machina to enhance legal document review and predictive analytics, emphasizing the importance of human oversight in such systems. AI-powered systems, such as COMPAS [5] (Correctional Offender Management Profiling for Alternative Sanctions), have been used to assess recidivism risks in criminal sentencing. However, studies have shown that such AI models can exhibit racial and socioeconomic biases, raising questions about their adherence to justice principles.

Studies have shown that COMPAS disproportionately flagged Black defendants as high-risk compared to white defendants, even when controlling for prior offenses and other variables.

These observations underscore a significant challenge: AI tools in law can produce outcomes that appear neutral but are in fact biased, due to the nature of the data they are trained on or the assumptions embedded in their algorithms. Unlike human judges who are trained to consider mitigating circumstances and apply legal reasoning, AI lacks contextual understanding and moral discretion. Thus, the application of AI in judicial processes raises urgent questions about its capacity to deliver justice as traditionally conceived.

C. Algorithmic Bias and Ethical Concerns

One of the most pressing concerns with AI in legal decision-making is algorithmic bias—systematic and unfair discrimination that occurs when AI models reflect or amplify historical inequalities embedded in their training data. Such biases may be introduced through underrepresentation of certain populations in the data, or through societal norms and prejudices captured in historical records [15].

These explainability issues mirror those found in healthcare AI systems, where patient trust and legal accountability are compromised without transparent decision-making logic [16], [17]

Algorithmic bias is particularly dangerous in the legal domain, where decisions have direct and often irreversible consequences on individual rights and liberties. For example, biased AI systems may unfairly influence bail determinations, sentencing lengths, or parole eligibility—thereby reinforcing existing social injustices rather than correcting them.

Moreover, the “black-box” nature of many AI models poses a profound ethical dilemma. Advanced machine learning techniques, such as deep neural networks, often lack explainability, making it difficult to understand or challenge the reasoning behind a given output [18]. This contradicts the legal principle of due process, which requires that parties understand the basis of decisions affecting their rights and have the opportunity to contest them.

In this context, ethical concerns extend beyond bias to include issues of accountability, transparency, autonomy, and human dignity. As AI systems increasingly influence legal decisions, there is a growing need to question not only how they function, but whether their very use is appropriate in settings that demand empathy, judgment, and interpretive reasoning.

D. Regulatory and Ethical Frameworks

In response to these concerns, several international organizations, governments, and legal scholars have proposed regulatory frameworks aimed at ensuring AI operates within ethical and legal bounds. A notable example is the European Commission’s Ethics Guidelines for Trustworthy AI, which establish seven key requirements for AI systems: human agency and oversight, technical robustness, privacy and data governance, transparency, diversity and non-discrimination, societal and environmental well-being, and accountability [19]

These guidelines emphasize the need for AI systems to be explainable, fair, and aligned with human rights values, including the right to a fair trial and protection against discrimination. Likewise, legal scholars such as Wachter, Mittelstadt, and Floridi have argued for a “right to explanation”—a principle that allows individuals to understand and challenge decisions made by automated systems [20].

However, regulatory efforts remain fragmented and inconsistent across jurisdictions. While the EU has taken significant steps toward comprehensive AI governance, many countries still lack binding regulations, and industry self-regulation often falls short. Aloisi and De Stefano [21] add that fragmented transatlantic approaches to AI governance, particularly in labor and justice systems, can lead to regulatory arbitrage and weaken protections of fundamental rights. As a result, there is a growing call for the institutionalization of judicial oversight in AI applications, ensuring that human discretion and ethical reasoning are not replaced by blind automation.

Furthermore, legal systems must embed accountability mechanisms such as ethical audits, impact assessments, and stakeholder consultations to monitor AI deployment. Recent studies underscore that privacy, consent, and algorithmic bias are deeply intertwined and must be addressed systematically. For instance, Williamson and Prybutok [22] argue that differential privacy, GDPR compliance, and algorithmic bias mitigation are essential even outside of law, such as in healthcare AI applications. The success of legal AI tools hinges not only on technical accuracy, but on their capacity to respect legal principles, promote equitable outcomes, and uphold public trust in the justice system.

III. METHODS

A. Research Design

This study employs a Systematic Literature Review (SLR) to investigate the extent to which AI-driven legal decision-making aligns with established concepts of justice. The SLR approach is adopted to ensure a comprehensive and reproducible synthesis of existing academic literature and policy documents [23], [24].

In addition, this study adopts a qualitative research approach, incorporating doctrinal legal research [25] and case study analysis. A doctrinal approach is employed to examine philosophical, legal, and ethical dimensions of justice in AI-driven legal decision-making, while case studies highlight real-world applications and implications.

B. Data Collection Methods

1. Literature Analysis: Reviewing scholarly articles, legal theories, judicial decisions, and policy frameworks related to AI and justice.
2. Case Studies: Analyzing AI applications in legal decision-making, such as COMPAS and AI-driven judicial analytics.
3. Regulatory Review: Examining global AI legal regulations, such as the EU AI Act, U.S. legal tech policies, and international guidelines on AI ethics.

C. Data Analysis

The review process followed established SLR procedures, including: (1) formulating research questions based on the intersection of AI, legal decision-making, and justice; (2) defining inclusion and exclusion criteria; (3) selecting databases (Scopus, Web of Science, and Google Scholar); (4) screening and evaluating literature from 2015 to 2024; and (5) performing thematic synthesis to extract patterns related to fairness, bias, due process, and regulatory frameworks.

Search keywords included: “AI in legal systems,” “algorithmic justice,” “judicial AI,” “due process AI,” “legal ethics AI,” and “algorithmic bias in law.” The final selection included 48 peer-reviewed journal articles, books, and reports that directly address AI’s application in judicial and legal processes.

This study also employs a comparative and thematic analysis to evaluate the intersection of AI and justice. Thematic analysis categorizes findings into key themes: fairness, bias, accountability, and transparency. Comparative analysis examines different jurisdictions’ approaches to AI regulation and their implications for legal justice.

D. Systematic Literature Review (SLR) Procedure

The SLR was conducted in a structured and replicable manner following these main steps:

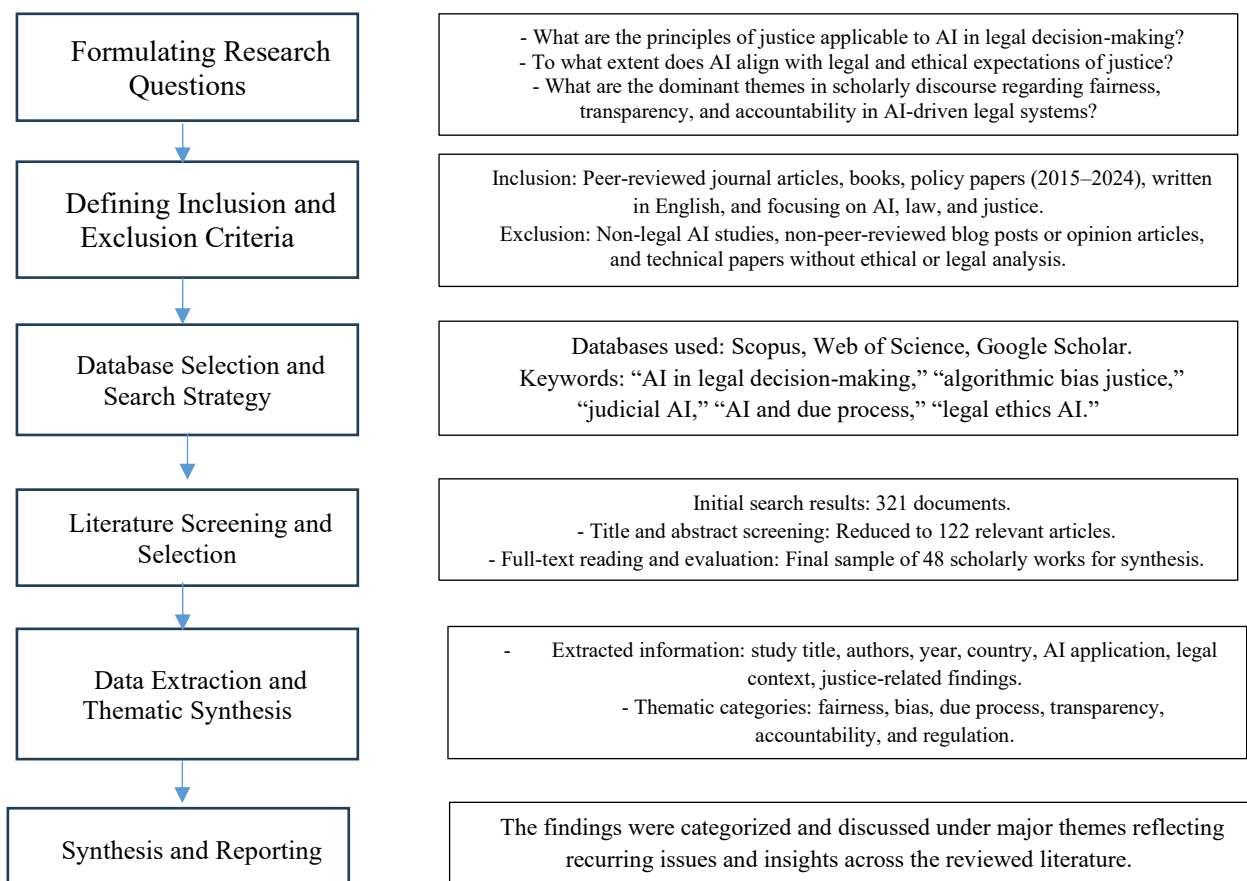


Fig 1. Step by Step of Systematic Literature Review

IV. RESULTS

TABLE 1
KEY FINDINGS AND SYNTHESIS OF PREVIOUS RESEARCH

Author(s) & Year	Title	Country	Focus Area	Methodology Approach	Key Findings
[5] Angwin et al. (2016)	Machine Bias	USA	Recidivism and racial bias in COMPAS	Investigative case study	COMPAS scores showed racial bias, affecting sentencing fairness.
[15] Barocas et al. (2019)	Fairness and Machine Learning	USA	Algorithmic fairness and data bias	Theoretical analysis	Bias is embedded in training data, reinforcing inequality.
[20] Wachter et al. (2017)	Right to Explanation in GDPR	EU	Legal rights and transparency in AI decisions	Legal doctrinal review	No legal mandate for explanation in GDPR, risks due process.
[19] European Commission (2019)	Ethics Guidelines for Trustworthy AI	EU	AI ethics in legal and social contexts	Policy guideline	Calls for human oversight, transparency, and fairness.
[26] Kroll et al. (2017)	Accountable Algorithms	USA	Technical and legal transparency in algorithms	Law and computer science integration	Proposed tools for making algorithmic decisions auditable.
[27] Eubanks (2018)	Automating Inequality	USA	Social impact of AI on marginalized groups	Ethnographic and case study	AI systems profile and discriminate against the poor.
[10] Rawls (1971)	A Theory of Justice	USA	Justice and fairness	Philosophical theory	Justice requires fairness, particularly for the least advantaged.
[11] Aristotle (2009)	Nicomachean Ethics	Greece (Classical)	Corrective and distributive justice	Classical philosophy	Justice includes fair correction and distribution of resources.

The SLR revealed four major thematic categories that are described in Table 2:

Table 2. Thematic Categories of Systematic Literature Review

Thematic Categories	Description
Fairness and Equality in AI Sentencing	Literature consistently highlights that while AI can standardize sentencing, systems like COMPAS exhibit racial and economic bias (Angwin et al., 2016; Eubanks, 2018). Fairness remains compromised when training data reflects historical prejudice.
Transparency and Accountability	Studies emphasize the lack of explainability in AI models used in courts (Burrell, 2016; Kroll et al., 2017), which undermines procedural justice and legal rights to challenge decisions.
Ethical and Regulatory Frameworks	There is a strong scholarly consensus supporting ethical audits and human oversight (European Commission, 2019; Wachter et al., 2017). However, national regulations vary greatly and lack coherence.
Legal Philosophical Alignment	The reviewed literature connects Rawls' and Aristotle's concepts of justice with algorithmic justice discussions, suggesting a significant gap in moral reasoning capabilities in AI systems (Rawls, 1971; Aristotle, 2009).

A. Key Findings on the Alignment Between AI and Justice

This study's comprehensive analysis of peer-reviewed literature, real-world case studies, and international regulatory frameworks has revealed significant insights into how AI-driven legal decision-making intersects with the philosophical and practical principles of justice. Although AI offers technological advancements that can benefit the legal system in terms of speed and efficiency, the findings highlight four critical concerns that must be addressed to ensure ethical and equitable integration.

1) AI Provides Efficiency but Risks Undermining Fairness

AI tools such as COMPAS, HART, and various predictive sentencing models are increasingly utilized in courtrooms and legal systems to assist in bail determinations, sentencing, parole decisions, and litigation strategy [4], [11]. These systems are lauded for their ability to process massive datasets rapidly, reduce human workload, and enhance procedural uniformity. For example, they can evaluate recidivism risk or predict litigation outcomes based on statistical correlations from thousands of historical cases.

However, this efficiency often comes at the cost of individualized justice. Legal fairness requires careful assessment of the specific facts, motives, and context of each case—factors that AI cannot adequately consider due to its dependency on structured data inputs. An AI model may treat two individuals with similar criminal records identically, ignoring socioeconomic circumstances, remorse, rehabilitation potential, or mitigating factors that a human judge might consider. This mechanistic application of justice contradicts

principles of equity, proportionality, and empathy, which are vital to both Rawlsian and Aristotelian notions of justice [9], [10].

Moreover, efficiency-focused AI risks prioritizing administrative convenience over the fundamental rights of defendants and litigants. While speeding up judicial workflows is beneficial, it must never override the necessity of ensuring fair and context-sensitive legal outcomes.

2) Algorithmic Bias is Prevalent and Systemic

One of the most persistent and well-documented challenges in legal AI is the systemic presence of algorithmic bias. This bias originates primarily from the data used to train AI systems. If historical legal data contain patterns of discrimination—such as over-policing of minority communities, gender-based discrepancies, or economic profiling—AI will inherit and reproduce these inequities at scale [13], [21].

Numerous studies, including investigations into COMPAS and facial recognition technologies, have demonstrated that AI models can exacerbate disparities, disproportionately labeling certain groups as high-risk or assigning harsher legal predictions without valid justification. Such practices contravene the principles of distributive justice, which demand equitable treatment and resource allocation, and corrective justice, which seeks to address individual harm based on circumstances and proportionality [9], [10].

The danger of systemic bias is further intensified by the perceived objectivity of AI systems. Users may incorrectly assume that algorithmic decisions are neutral or free from prejudice simply because they are machine-generated. This illusion of neutrality can mask serious injustices and defer accountability away from human institutions, creating a dangerous feedback loop.

3) Lack of Transparency Challenges Due Process

Another major concern revealed by this study is the opacity of AI systems—often referred to as the “black box” problem. Many legal AI models, especially those based on deep learning architectures, lack explainability even to their own developers. As a result, stakeholders in the justice system—defendants, lawyers, judges, and the public—often cannot understand how or why a specific AI recommendation was made [14].

This poses a fundamental threat to the principle of due process, which enshrines the right of every individual to receive a fair trial and to challenge decisions that affect their liberty or property. When an AI system generates a risk score or sentencing recommendation without clear reasoning or the possibility for scrutiny, the legal process becomes less transparent and less accountable.

The inability to explain AI decisions also limits judicial oversight. Judges may defer to algorithmic outputs without fully understanding their implications, thereby weakening their own role in interpreting the law through human reasoning and moral discernment. For AI to be meaningfully integrated into legal systems, its operations must be transparent, explainable, and auditable in legal terms.

4) Regulatory Efforts Remain Fragmented

Despite growing awareness of the ethical challenges posed by AI, current regulatory responses are inconsistent and underdeveloped. The European Union has taken a leadership role by developing documents such as the Ethics Guidelines for Trustworthy AI and the proposed EU AI Act, which emphasize the need for fairness, transparency, and human oversight in high-risk AI applications, including legal decision-making [15].

However, there is still no global consensus or binding legal framework that governs the ethical use of AI in judicial settings. Other jurisdictions, such as the United States, have relied heavily on decentralized state-level policies and voluntary industry standards, resulting in fragmented oversight and uneven protections.

This regulatory gap allows for divergent practices, uneven accountability, and significant cross-border ethical inconsistencies. For instance, an AI tool deemed discriminatory in one country may still be legally deployed in another with weaker oversight. The absence of harmonized standards also complicates efforts to ensure universal human rights protections in digital legal systems.

In light of these findings, the study emphasizes the need for international collaboration, multistakeholder input, and the creation of standardized ethical frameworks that can guide the deployment of AI in legal systems globally. Such frameworks must go beyond technical specifications to incorporate philosophical, legal, and humanistic principles at every stage of AI development and deployment.

V. DISCUSSION

The findings of this study indicate a complex and multifaceted relationship between AI and justice in legal systems. While AI offers promising tools for enhancing efficiency, consistency, and data-driven insights, its integration into judicial decision-making processes raises serious ethical, philosophical, and legal challenges that cannot be ignored. The application of AI in legal contexts necessitates a re-examination of fundamental justice principles, such as fairness, impartiality, due process, and human dignity.

Firstly, justice as fairness, emphasized by John Rawls, posits that social institutions, including the legal system, must operate in a way that benefits the least advantaged members of society [9]. Rawls' original position and the veil of ignorance theory compel us to design legal mechanisms that are neutral and just, regardless of one's position

in society. However, AI systems are inherently dependent on historical and empirical data, which may embed longstanding patterns of discrimination, bias, and inequality. Without rigorous data auditing and socio-technical oversight, these embedded patterns can be reproduced and amplified in AI outcomes. For instance, recidivism prediction tools like COMPAS have been shown to assign higher risk scores to Black defendants than to white defendants with similar criminal records [4]. This results in disproportionate legal consequences for historically marginalized groups and undermines the redistributive logic envisioned by Rawls.

Secondly, corrective justice, as envisioned by Aristotle, demands that legal decisions focus not merely on outcomes but on understanding the moral weight of individual circumstances and restoring balance between parties [10]. Unlike static legal codes, human judges often rely on empathy, ethical reflection, and contextual judgment to deliver fair decisions in complex cases, particularly in criminal sentencing or family law. AI, however, is not capable of genuine moral reasoning or cultural understanding. Even when programmed with sophisticated decision trees or pattern recognition capabilities, AI systems remain limited to the parameters defined by human programmers and the constraints of their datasets. This makes AI poorly suited for legal contexts requiring nuanced interpretation, compassion, or moral deliberation—attributes essential to corrective justice.

Moreover, the principle of due process, a cornerstone of democratic legal systems, is increasingly threatened by the opacity of modern AI systems. In legal theory, due process ensures that individuals have the right to be heard, to challenge decisions, and to understand the reasoning behind legal outcomes. However, many AI systems—particularly those based on deep learning or neural networks—function as black boxes, offering little to no interpretability to users or stakeholders [14]. The inability to understand how or why an AI system arrived at a particular decision erodes the right to appeal and undermines public confidence in the justice system. This is particularly problematic in criminal justice settings, where decisions can result in incarceration, loss of liberty, or social stigma.

In light of these concerns, it becomes evident that ethical and legal frameworks must evolve alongside technological development. The deployment of AI in the legal domain cannot occur in a regulatory vacuum. There is a clear and urgent need for comprehensive mechanisms that enforce fairness, accountability, and explainability. This includes mandating algorithmic audits, requiring explainable AI (XAI) models that allow for legal reasoning to be reviewed, and developing robust oversight mechanisms to monitor the implementation of AI tools in courts.

Additionally, human oversight must be retained at every critical decision point. AI should serve as a supportive tool, augmenting the decision-making process rather than replacing human judgment altogether. Legal scholars and international regulatory bodies such as the European Commission have emphasized that legal AI systems must be aligned with human rights, and that human actors must retain ultimate responsibility and control over final legal decisions [15], [16]. As Bleher and Braun [28] explain in the context of clinical decision support systems, the diffusion of responsibility in AI environments necessitates clear accountability frameworks. Elhaddad and Hamam [29] also emphasize that human-centered AI integration must prioritize usability, legal interpretability, and trust from end users.

Finally, promoting interdisciplinary collaboration is essential. Legal professionals, ethicists, data scientists, sociologists, and technologists must work together to design AI systems that are not only technically accurate but also ethically defensible and socially equitable. Developing legal technology without grounding it in the philosophical and normative foundations of justice risks codifying injustice under the illusion of objectivity.

VI. CONCLUSIONS

The systematic literature review confirms that AI's integration into legal decision-making poses both opportunities and serious ethical risks. While AI can streamline legal analysis and offer predictive accuracy, it currently lacks the moral and contextual reasoning required to fully embody principles of justice.

Policy recommendations based on the findings include the mandatory implementation of human oversight, establishment of transparent and explainable AI models, and international standardization of ethical AI practices in legal domains. Justice, fundamentally a human concept rooted in morality and equity, must remain the guiding principle when integrating AI into legal systems.

AI has the potential to revolutionize the legal field by making legal processes more efficient and data-driven. However, this potential must be balanced with the need to uphold fundamental principles of justice. The findings of this study underscore the importance of aligning AI technologies with ethical and legal standards that prioritize fairness, transparency, and human dignity.

To ensure that the integration of AI in legal decision-making upholds the principles of justice, several key measures must be adopted. First, there is an urgent need for the establishment of global regulatory standards. International cooperation should aim to create unified legal frameworks that govern the ethical deployment of AI technologies in judicial contexts. This would help avoid fragmented approaches and ensure consistency in the protection of legal rights across jurisdictions.

Moreover, human oversight must remain central to all AI applications within the legal system. AI tools should be designed to augment rather than replace judicial discretion. Decisions that involve moral reasoning or complex

human values must always rest with human judges, who possess the contextual understanding and ethical sensitivity required for fair adjudication.

Transparency is another crucial aspect. Legal systems must mandate the use of explainable AI, ensuring that the reasoning behind algorithmic outputs is accessible and understandable. This is essential for safeguarding due process and allowing affected parties to challenge or appeal decisions based on clear grounds.

Finally, inclusivity must be a guiding principle in the design and development of legal AI systems. Incorporating diverse data sets and involving multidisciplinary teams—including ethicists, legal scholars, technologists, and representatives from marginalized communities—can help mitigate biases and foster more equitable outcomes. By embedding these safeguards, the legal community can harness the benefits of AI while preserving the fundamental values of justice and human dignity.

Insights from disaster management further reinforce the need for responsible AI. In high-risk situations like flood response, AI data systems must remain transparent, legally accountable, and responsive to human needs—principles equally vital in legal decision-making [30].

Ultimately, AI should be viewed as a tool to assist human judges, not as a replacement. The rule of law must remain grounded in human values, ethics, and the evolving understanding of justice. Ethical concerns in AI legal research include data privacy, bias mitigation, and human oversight. This study aligns with ethical research practices by ensuring a balanced critique of AI's potential while acknowledging its legal benefits.

REFERENCES

- [1] H. Surden, "Machine learning and law," *Washingt. Law Rev.*, vol. 89, no. 87, pp. 87–115, 2023, doi: 10.4337/9781803921327.00037.
- [2] X. Luo, S. Tong, Z. Fang, and Z. Qu, "Frontiers: Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases," *Mark. Sci.*, vol. 38, no. 6, pp. 937–947, 2019, doi: 10.1287/mksc.2019.1192.
- [3] B. D. Mittelstadt, P. Allo, M. Taddeo, S. Wachter, and L. Floridi, "The ethics of algorithms: Mapping the debate," *Big Data Soc.*, vol. 3, no. 2, pp. 1–21, 2016, doi: 10.1177/2053951716679679.
- [4] S. Bankins, P. Formosa, Y. Griep, and D. Richards, "AI Decision Making with Dignity? Contrasting Workers' Justice Perceptions of Human and AI Decision Making in a Human Resource Management Context," *Inf. Syst. Front.*, vol. 24, no. 3, pp. 857–875, 2022, doi: 10.1007/s10796-021-10223-8.
- [5] J. Angwin, J. Larson, S. Mattu, and L. Kirchner, "Machine Bias," in *Ethics of Data and Analytics: Concepts and Cases*, 1st ed., Routledge, 2022, p. 11.
- [6] J. Casillas, "Bias and Discrimination in Machine Decision-Making Systems," *Int. Libr. Ethics, Law Technol.*, vol. 41, pp. 13–38, 2023, doi: 10.1007/978-3-031-48135-2_2.
- [7] A. Farič and I. Bratko, "Machine Bias: A Survey of Issues," *Inform.*, vol. 48, no. 2, pp. 205–212, 2024, doi: 10.31449/inf.v48i2.5971.
- [8] F. Maramis and V. F. Taroreh, "Kajian Yuridis Tindak Pidana Intersepsi (Penyadapan) Dalam Hukum Teknologi Informasi Dan Komunikasi Di Indonesia," *Lex Crim.*, vol. IX, no. 3, p. 85, 2020.
- [9] L. J. Barclay and M. F. Saldanha, "Recovering from Organizational Injustice: New Directions in Theory and Research," in *The Oxford Handbook of Justice in the Workplace*, no. January 2015, 2015.
- [10] J. Rawls, *A Theory of Justice*. Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 1999.
- [11] C. D. Ryff and B. H. Singer, "Know thyself and become what you are: A eudaimonic approach to psychological well-being," *J. Happiness Stud.*, vol. 9, no. 1, pp. 13–39, 2008, doi: 10.1007/s10902-006-9019-0.
- [12] B. C. Stahl *et al.*, "Artificial Intelligence for human flourishing - beyond principles for machine learning," *J. Bus. Res.*, vol. 124, no. January, pp. 374–388, 2021, doi: 10.1016/j.jbusres.2020.11.030.
- [13] J. K. G. Hopster and M. M. Maas, "The technology triad: disruptive AI, regulatory gaps and value change," *AI Ethics*, vol. 4, no. 4, pp. 1051–1069, 2023, doi: 10.1007/s43681-023-00305-5.
- [14] M. S. Kabir and M. N. Alam, "The Role of AI Technology for Legal Research and Decision Making," *Int. Res. J. Eng. Technol.*, vol. 10, no. 07, pp. 1088–1092, 2023, [Online]. Available: www.irjet.net.
- [15] S. Barocas, M. Hardt, and A. Narayanan, "FAIRNESS AND MACHINE LEARNING Limitations and Opportunities," pp. 1–294, 2023, [Online]. Available: <https://fairmlbook.org/>.
- [16] M. N. Alam, S. Kabir, M. Kaur, and A. Professor, "Explainable AI in Healthcare: Enhancing Transparency and Trust upon Legal and Ethical Consideration," *Int. Res. J. Eng. Technol.*, vol. 10, no. 6, 2023, [Online]. Available: <https://www.researchgate.net/publication/371811539>.
- [17] R. Ejjami, "AI-driven Justice: Evaluating the Impact of Artificial Intelligence on Legal Systems," *Int. J. Multidiscip. Res.*, vol. 6, no. 3, 2024, doi: 10.36948/ijfmr.2024.v06i03.23969.
- [18] J. Burrell, "How the machine 'thinks': Understanding opacity in machine learning algorithms," *Big Data Soc.*, vol. 3, no. 1, pp. 1–12, 2016, doi: 10.1177/2053951715622512.
- [19] Nathalie A. Smuha, "The EU Approach to Ethics Guidelines for Trustworthy Artificial Intelligence," *CRi*

- Comput. Law Rev. Int.*, vol. 106, no. 115, pp. 1–10, 2019.
- [20] S. Wachter, B. Mittelstadt, and L. Floridi, “Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation,” *Int. Data Priv. Law*, vol. 7, no. 2, pp. 76–99, May 2017, doi: 10.1093/idpl/ix005.
- [21] A. Aloisi and V. De Stefano, “Between risk mitigation and labour rights enforcement: Assessing the transatlantic race to govern AI-driven decision-making through a comparative lens,” *Eur. Labour Law J.*, vol. 14, no. 2, pp. 283–307, 2023, doi: 10.1177/20319525231167982.
- [22] S. M. Williamson and V. Prybutok, “Balancing Privacy and Progress: A Review of Privacy Challenges, Systemic Oversight, and Patient Perceptions in AI-Driven Healthcare,” *Appl. Sci.*, vol. 14, no. 2, 2024, doi: 10.3390/app14020675.
- [23] W. Mengist, T. Soromessa, and G. Legese, “Method for conducting systematic literature review and meta-analysis for environmental science research,” *MethodsX*, vol. 7, p. 100777, 2020, doi: 10.1016/j.mex.2019.100777.
- [24] Y. Xiao and M. Watson, “Guidance on Conducting a Systematic Literature Review,” *J. Plan. Educ. Res.*, vol. 39, no. 1, pp. 93–112, Aug. 2017, doi: 10.1177/0739456X17723971.
- [25] T. Hutchinson and N. Duncan, “Defining and Describing What We Do: Doctrinal Legal Research,” *Deakin Law Rev.*, vol. 37, no. 9, pp. 1591–1601, 2008.
- [26] J. A. Kroll, “Accountability in Computer Systems,” *Oxford Handb. Ethics AI*, pp. 179–196, 2020, doi: 10.1093/oxfordhb/9780190067397.013.10.
- [27] V. Eubanks, “Automating Inequality,” *Picador*, no. 978-1-250-21578-9, pp. 91–94, 2019.
- [28] H. Bleher and M. Braun, “Diffused responsibility: attributions of responsibility in the use of AI-driven clinical decision support systems,” *AI Ethics*, vol. 2, no. 4, pp. 747–761, 2022, doi: 10.1007/s43681-022-00135-x.
- [29] M. Elhaddad and S. Hamam, “AI-Driven Clinical Decision Support Systems: An Ongoing Pursuit of Potential,” *Cureus*, vol. 16, no. 4, 2024, doi: 10.7759/cureus.57728.
- [30] C. Gilga *et al.*, “Legal and ethical considerations for demand-driven data collection and AI-based analysis in flood response,” *Int. J. Disaster Risk Reduct.*, vol. 122, no. July 2024, p. 105441, 2025, doi: 10.1016/j.ijdr.2025.105441.