Open Access Library Journal
2025, Volume 12, e13155
ISSN Online: 2333-9721
ISSN Print: 2333-9705

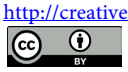
The Indispensable Role of Law in Governing AI: Why Legal Ethics and Regulation Are Essential for Responsible AI Development

Emmanuel A. M. Sam

Faculty of Law, Fourah Bay College, University of Sierra Leone, Freetown, Sierra Leone
Email: emmanuelmacpherson2010@gmail.com

How to cite this paper: Sam, E.A.M. (2025) The Indispensable Role of Law in Governing AI: Why Legal Ethics and Regulation Are Essential for Responsible AI Development. *Open Access Library Journal*, **12**: e13155.
https://doi.org/10.4236/oalib.1113155

Received: February 25, 2025
Accepted: November 25, 2025
Published: November 28, 2025



Abstract

Artificial Intelligence (AI) has emerged as a transformative force in various industries, including law, healthcare, finance, and governance. However, AI's potential benefits come with significant ethical and legal challenges. AI can pose risks related to bias, accountability, transparency, and human rights violations without a robust legal framework to regulate its operations. This paper argues that AI is insufficient without laws to guide its conduct and ethical application. Legal ethics are crucial in ensuring that AI operates within the boundaries of fairness, justice, and accountability. The absence of legal oversight could lead to unchecked AI systems that perpetuate discrimination, invade privacy, and disrupt societal order. This study underscores the necessity of legal principles governing AI's development and deployment by examining the intersection of AI and law. It advocates for a structured regulatory approach to ensure that AI remains a tool for positive innovation while upholding fundamental ethical and legal standards.

Subject Areas

Sociology

Keywords

Artificial Intelligence, Law, Regulation, Data Protection, Ethics, Justice, Oversight, Human, Machines

1. Introduction

Artificial Intelligence (AI) has become integral to modern society, influencing healthcare, finance, governance, and legal industries. While AI offers significant

advancements in efficiency, automation, and decision-making, its rapid development raises critical ethical and legal concerns [1]. The question remains: Can AI function effectively and ethically without legal oversight? This paper argues that AI is insufficient without laws and moral guidelines to regulate its operations and impact on society.

Without a structured legal framework, AI systems risk perpetuating biases, infringing on individual rights, and functioning without accountability. Issues like data privacy, algorithmic discrimination, and decision-making transparency emphasize the need for legal intervention to guarantee responsible AI deployment [2]. Legal ethics safeguard against AI's potential harms, ensuring its application aligns with fairness, justice, and human rights principles [2]. This study examines the essential role of law in directing AI behavior and preventing unethical practices. By exploring key areas such as accountability, transparency, and fairness, the paper emphasizes the critical need for legal frameworks that govern AI's development and use [3]. Ultimately, it contends that AI alone is inadequate—without legal oversight, its advantages may be eclipsed by ethical challenges and societal risks.

Artificial Intelligence (AI) has transformed industries by improving efficiency, automation, and decision-making. However, its rapid advancement raises critical ethical and legal concerns, including bias, accountability, privacy, and transparency. Without a well-defined legal framework, AI may operate in ways that conflict with fundamental human rights and societal values [4]. This research examines why AI is insufficient without laws to regulate its operations and why legal ethics are essential in ensuring responsible AI development.

As AI advances, establishing a regulatory body or enforceable guidelines is critical to aligning with ethical, legal, and human rights standards. The Ruggie Principles provide a strong precedent for AI governance, emphasizing accountability, due diligence, and human rights protections [5]. AI could exacerbate societal inequalities without proper oversight, leading to discrimination, privacy violations, and security risks [4]. **A** global, legally binding AI governance framework must be developed to prevent these harms and ensure that AI serves humanity responsibly and ethically.

2. AI and Its Challenges to Global World Order

2.1. Why Is AI Insufficient without Laws to Regulate Its Operations?

Artificial Intelligence (AI) has revolutionized various industries, including healthcare, finance, law, and governance, by improving efficiency, automating decision-making, and analyzing vast amounts of data. However, despite its numerous advantages, AI is insufficient without legal regulations to govern its operations. AI systems can pose serious ethical, legal, and societal risks without appropriate legal oversight, including bias, discrimination, privacy violations, and accountability issues.

AI systems are trained on large datasets that often reflect historical biases. If

unregulated, these biases can be amplified, leading to discriminatory outcomes, particularly in hiring processes, law enforcement, and lending decisions [1]. For example, studies have shown that AI-powered recruitment tools can unintentionally discriminate against women and minority groups due to biased training data [2]. Legal frameworks are essential to ensure that AI algorithms are regularly audited for fairness and do not reinforce systemic biases.

2.2. AI Lack of Transparency and Accountability

One of the most significant concerns surrounding Artificial Intelligence (AI) systems is their lack of transparency and accountability. Many AI algorithms, especially in machine learning and deep learning, function as "black boxes," meaning that their decision-making processes are not easily understandable or interpretable by humans [6]. This opacity raises significant ethical and legal concerns, particularly in law, healthcare, and finance, where AI decisions can profoundly impact individuals' lives. The absence of clear understanding about how an AI system arrives at its conclusions or recommendations makes it difficult to assign responsibility for errors or harmful outcomes.

The lack of transparency in AI systems presents a challenge when those systems are deployed in decision-making processes that affect individuals' rights and freedoms. For example, predictive algorithms that assess the likelihood of re-offending in criminal justice can impact sentencing and parole decisions. Yet, the inner workings of these algorithms are often proprietary and not disclosed to the public or even to the individuals being judged [7]. This lack of insight prevents affected parties from contesting decisions and seeking accountability for potential harm caused by biased or incorrect AI predictions.

Moreover, AI's opacity exacerbates issues related to accountability. Suppose an AI system makes a harmful or discriminatory decision. In that case, it can be challenging to pinpoint who should be held liable—the developers who created the algorithm, the companies that deployed it, or the end-users who relied on its recommendations [7]. Without proper accountability frameworks, AI systems could be used in ways that violate human rights or legal standards without any clear party being held responsible for the consequences. This makes the establishment of regulatory frameworks that mandate transparency, explainability, and accountability for AI systems essential to ensure justice and fairness in their deployment.

Legal scholars argue that regulations must be established to ensure that AI's decision-making processes are transparent, and stakeholders can understand, challenge, and seek remedies for harmful outcomes. The European Union's General Data Protection Regulation (GDPR) includes provisions that give individuals the right to an explanation when subjected to automated decision-making, recognizing the need for transparency in AI systems [8]. This regulatory approach serves as a potential model for ensuring that AI operates within ethical and legal boundaries, providing accountability and fairness in its use.

2.3. The Risk of Bias and Discrimination

Artificial Intelligence (AI) systems, particularly those relying on machine learning algorithms, are increasingly used in critical decision-making areas such as hiring, law enforcement, lending, and healthcare. However, these AI systems pose significant risks of bias and discrimination due to the nature of the data they are trained on and the design choices made by developers [9]. When AI models are trained on biased or unrepresentative datasets, the systems may perpetuate or exacerbate societal inequalities. These biases can lead to discriminatory outcomes that disproportionately affect marginalized groups, including racial minorities, women, and people with disabilities [2].

One of the primary concerns regarding AI bias is the data used to train algorithms. Machine learning models learn from historical data, and if that data reflects prejudiced patterns or outcomes, AI systems can inherit and amplify these biases [2]. For example, suppose an AI system used in recruitment is trained on data from past hiring decisions that reflect gender or racial biases. In that case, the system may favor male candidates or individuals from particular racial groups, even if no explicit bias is intended [10]. Such biases can profoundly impact opportunities for underrepresented groups, further entrenching inequalities.

In the criminal justice system, AI algorithms are increasingly used to assess the risk of recidivism and inform parole decisions. COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) is a widely discussed case, a risk assessment tool used to predict the likelihood of a criminal re-offending. Research by ProPublica found that COMPAS was biased against Black defendants, predicting higher recidivism rates for them than for White defendants, even when controlling for the nature of their crimes [7]. This kind of biased decision-making is dangerous because it can disproportionately affect already marginalized groups and lead to unfair legal outcomes.

The design of AI systems plays a crucial role in determining their potential for bias. AI developers may inadvertently introduce biases during the design and testing phases due to a lack of diverse perspectives in the development teams or insufficient attention to fairness and equity [2]. This lack of diversity in AI development has been identified as contributing to biased AI outcomes, as systems may not account for the experiences and needs of all demographic groups.

Legal and ethical frameworks must be established to mitigate these risks that require AI systems to be transparent, explainable, and regularly audited for fairness. Legal regulations should mandate that AI developers ensure their systems are free from biases that could lead to discrimination and are designed to operate fairly across all demographic groups. Furthermore, policymakers must push for the inclusion of diverse voices in AI development to ensure that the technology benefits all members of society equitably [11].

2.4. Privacy and Data Security Concerns

As Artificial Intelligence (AI) continues to permeate various sectors, including

healthcare, finance, and social media, privacy and data security concerns have become more pressing. AI systems rely on vast amounts of data, often including sensitive personal information, to function effectively. While AI can provide valuable insights and streamline decision-making processes, collecting, processing, and storing such data introduces significant risks related to privacy violations, unauthorized access, and misuse.

One of the most significant concerns is the potential for AI systems to infringe on individual privacy. AI systems often process data that individuals may not even be aware is being collected, including information from online behaviors, medical records, financial transactions, and social media activities [12]. The aggregation of such personal data can lead to privacy breaches, especially when it is collected without informed consent or when data is used for purposes beyond what was initially agreed upon. The use of AI to track personal activities and predict behavior, often without transparency, undermines individuals' autonomy and raises questions about consent in a digital age.

For instance, law enforcement agencies use AI-powered facial recognition technology to identify and track individuals in various public spaces. This practice has been met with widespread criticism, as it could lead to mass surveillance, disproportionately affecting marginalized communities [13]. Furthermore, concerns over AI-driven profiling, where individuals are categorized based on their behavior or online presence, can result in privacy violations and unwanted manipulation, such as targeting vulnerable individuals for advertisements or political messaging [12].

In addition to privacy concerns, AI systems present significant data security risks. As AI systems process and store large amounts of data, they become attractive cyberattack targets. AI data storage and processing breaches can expose sensitive personal information, including medical histories, financial data, and private communications. For example, violating an AI-driven healthcare system could compromise patient privacy and expose individuals to identity theft or fraud. The use of AI in cybersecurity itself, though designed to protect data, can be weaponized by malicious actors, leading to AI-enabled cyberattacks on other AI systems or infrastructure [10].

Moreover, AI's dependence on massive datasets often raises the issue of data ownership and control. Individuals may not be able to fully understand or control how their data is used by AI systems, which complicates the enforcement of privacy rights. The European Union's General Data Protection Regulation (GDPR) aims to address these concerns by giving individuals more control over their data and requiring that organizations ensure data protection by design [8]. However, such regulatory frameworks are not universally adopted, and different regions may have varying data privacy and security standards, leaving gaps in individuals' protection.

Given these concerns, governments and regulatory bodies must establish robust legal frameworks to protect individual privacy and secure data in AI systems. Such

regulations should ensure that AI systems are transparent, that individuals have control over their data, and that data security protocols are implemented at every stage of AI deployment. Additionally, AI developers must be required to design systems that prioritize data protection, ensuring that privacy is respected and that security risks are minimized.

One of the significant challenges of AI is the "black box" nature of many machine learning models, where decision-making processes are opaque and complex to interpret [6]. This lack of transparency raises concerns, particularly in legal and medical settings, where AI-driven decisions can significantly impact individuals' rights and well-being. AI developers may not be held accountable for errors or unethical behavior without regulatory oversight. Laws mandating explainability and accountability can ensure that AI systems provide justifiable and transparent decisions [14].

AI relies on massive amounts of data, raising concerns about privacy and security. Unregulated AI systems can exploit personal data without user consent, violating fundamental privacy rights [12]. The European Union's General Data Protection Regulation (GDPR) is an example of a legal framework that sets clear guidelines on data protection, ensuring that AI applications handle personal data responsibly. Without such laws, AI could be misused for mass surveillance, identity theft, and unauthorized data processing, posing significant threats to individuals' rights.

3. The Importance of Guiding the Conduct of AI

3.1. Ethical and Legal Responsibility in AI

The rise of Artificial Intelligence (AI) has prompted crucial discussions regarding ethical and legal responsibility, especially when AI systems make decisions that affect individuals' lives and rights. From autonomous vehicles to predictive algorithms in criminal justice, AI is increasingly involved in decision-making processes traditionally handled by humans. However, deploying AI brings significant challenges regarding who should be held responsible when these systems cause harm or make flawed decisions.

One of the primary ethical concerns surrounding AI is the question of accountability. When AI systems operate autonomously and make decisions without human intervention, it becomes difficult to pinpoint who is responsible for any adverse outcomes [15]. For example, in the case of an autonomous vehicle involved in a traffic accident, should the responsibility lie with the vehicle's manufacturer, the software developer, or the vehicle owner? This dilemma is incredibly complex given that AI systems often operate as "black boxes," with their decision-making processes opaque and complex to interpret [6]. Victims may find it difficult to seek redress or compensation in such cases, as the party responsible for the AI's actions is unclear.

From a legal perspective, the question of liability becomes even more complicated. Legal systems are generally structured around the assumption that human

actors are responsible for their actions. However, AI complicates this model by introducing systems that can act independently and make decisions based on data-driven patterns rather than human judgment. For instance, if an AI system wrongly denies someone access to a loan or incorrectly predicts their likelihood to commit a crime, determining who should be held liable is not straightforward. This raises the issue of whether AI should be granted legal personhood or if existing legal structures are sufficient to address AI-related harms [3]. Some legal scholars advocate for "AI as a tool" liability, where the developers or users of AI systems would be held accountable for any harm caused by the technology [16]. Others propose creating a new legal framework specific to AI to address the unique challenges it poses to liability.

Ethically, AI systems must be designed with fairness, transparency, and accountability principles. Ethical guidelines for AI development, such as those proposed by the European Commission, advocate for the responsible use of AI in ways that respect human dignity and ensure that AI systems are aligned with societal values [17]. This includes ensuring that AI does not perpetuate discrimination or bias, and that users' privacy and rights are safeguarded. Developers must also be held to ethical standards to ensure that their systems are used for positive social good and not for harmful purposes, such as surveillance or manipulation.

To address these ethical and legal challenges, there is a growing call for comprehensive regulatory frameworks to govern AI development and deployment. The European Union has taken steps toward this with its proposal for an AI Act, which outlines specific requirements for AI systems, including transparency, accountability, and risk management [18]. Such frameworks seek to balance innovation and ethical responsibility, ensuring that AI technologies contribute positively to society while minimizing the risks associated with their misuse.

3.2. Science Cannot Do It Alone

Artificial Intelligence (AI) has made remarkable strides in recent years, revolutionizing fields such as healthcare, finance, law, and transportation. AI's capacity to process vast amounts of data, identify patterns, and make decisions is unparalleled. Despite its advantages, however, the rapid development and deployment of AI technologies highlight a crucial limitation: while science and technology offer incredible advancements, they are insufficient on their own to answer all the questions surrounding the ethical, legal, and societal implications of AI without a legal framework to guide its operation.

AI systems, by their nature, lack the moral and ethical reasoning capabilities that human decision-makers possess. While AI can optimize processes and make data-driven decisions, it does not have an inherent understanding of the ethical consequences of its actions [16]. For instance, an autonomous vehicle may be capable of navigating a complex road system and making split-second decisions in the event of an accident. Still, it cannot inherently understand the ethical implications of its choices, such as whom to harm in a life-or-death situation [16]. These

decisions raise moral dilemmas that AI alone cannot resolve without a legal and ethical framework to guide them.

Moreover, the development and deployment of AI raise critical questions about privacy, accountability, and discrimination. AI systems often rely on vast datasets, including sensitive personal information. Without laws to govern data collection, storage, and use, these systems risk infringing on individuals' privacy and exposing them to harm [12]. The absence of clear legal boundaries can lead to biased AI decisions, as these systems can perpetuate or even exacerbate existing societal inequalities [2]. Thus, AI cannot function socially responsibly without laws that mandate transparency, fairness, and accountability.

Science alone also cannot address the broader social implications of AI. While AI can drive efficiency and innovation, it can also disrupt labor markets, exacerbate social inequalities, and create new forms of surveillance and control. For example, the automation of jobs through AI systems may lead to mass unemployment if not adequately regulated. Without legal frameworks to manage these transitions and ensure equitable distribution of benefits, the promise of AI may be overshadowed by societal harm [19].

The rapid pace at which AI is evolving necessitates the creation of legal structures that keep pace with technological advancements. Legal systems can uniquely address issues such as accountability, liability, and governance in a way that science alone cannot. Laws can provide the necessary oversight to ensure that AI systems are used in ways that align with societal values and human rights [3]. For instance, the European Union's General Data Protection Regulation (GDPR) sets guidelines for the ethical use of data and AI to protect individuals' privacy and data rights while fostering innovation [8].

3.3. Technological Advancement Can Benefit People If It Is Subject to Democratic Principles

Technological advancements, particularly in fields like Artificial Intelligence (AI), biotechnology, and digital communication, hold immense potential to transform society for the better. These innovations can lead to greater efficiency, improved healthcare, and increased access to information. However, the benefits of technological progress are not automatic. They can only be fully realized when technological developments are subject to democratic principles that ensure fairness, accountability, and the protection of individual rights. Without democratic oversight, technological advancements may exacerbate inequalities, violate privacy, and undermine public trust, ultimately limiting their potential benefits.

One of the fundamental tenets of democratic principles is the protection of individual rights and freedoms. In a democratic society, technology should serve the public good and not merely the interests of a select few. For instance, AI technologies used for surveillance or data collection must be governed by laws that ensure transparency and protect individuals' privacy [12]. In this regard, democratic principles such as transparency, participation, and accountability can ensure that

technology is deployed in ways that respect the rights of citizens. By involving the public in discussions about how technology should be used and regulating its deployment, governments can mitigate risks such as surveillance capitalism, where companies exploit personal data for profit [12].

Furthermore, democracy necessitates that technological development is guided by the people's needs and preferences, rather than being driven solely by powerful corporations or governments. The democratic process facilitates public discourse, debate, and input on significant issues, ensuring technological innovation aligns with societal values [11]. For instance, advancements such as telemedicine and genetic editing in healthcare can offer life-saving treatments and enhance access to care [19]. However, suppose these technologies are developed without democratic oversight. In that case, they may be utilized in ways that prioritize profit over patient welfare or lead to disparities in healthcare access [19]—ensuring that the public has a voice in developing and using these technologies can promote ethical and equitable outcomes.

Furthermore, democratic principles advocate for accountability in the use of technology. Governments, regulators, and corporations must be held accountable for how technology is deployed and its societal impact. This includes ensuring that AI systems used in hiring, criminal justice, and lending are fair and non-discriminatory [2]. Without robust accountability frameworks, AI and other technologies could perpetuate existing social inequalities, making it harder for marginalized groups to access opportunities. In a democratic society, laws and regulations can ensure that individuals and institutions are held responsible for the negative impacts of technology, such as biased algorithms or data breaches [7].

The benefits of technological advancement are maximized when democratic ideals inform the development and application of fairness, transparency, and accountability. A commitment to democratic principles ensures that technology serves all people equitably and its risks are managed responsibly. Moreover, when technology is subject to democratic oversight, it is more likely to be developed in a way that promotes the common good, addresses societal challenges, and respects human rights.

3.4. AI Can Only Be Ethical If It Aligns with Human Values

As Artificial Intelligence (AI) systems continue to advance, questions surrounding their ethical use have become increasingly significant. AI systems can process vast amounts of data and make decisions that can profoundly impact individuals and society. However, for AI to be considered ethical, it must be designed and operated in a manner that aligns with human values. These values—such as fairness, transparency, accountability, and respect for privacy—are essential in ensuring that AI systems contribute positively to society and do not cause harm.

One of the core challenges of creating ethical AI is aligning machine decisions with human values. AI systems are designed to optimize for specific outcomes based on the given data. However, these systems can sometimes produce results

that contradict or undermine societal values without guidance [1]. For instance, if an AI system is used in hiring decisions, it may prioritize efficiency or productivity over inclusivity and fairness, leading to biased outcomes that disproportionately affect certain demographic groups [2]. Therefore, ensuring that AI aligns with human values requires the developers to embed ethical guidelines and principles into the algorithms from the outset. This process is crucial in avoiding the risk of perpetuating harmful biases or creating systems that act in ways inconsistent with societal well-being.

Moreover, human values are essential in determining the transparency and accountability AI systems should uphold. Transparency in AI refers to the ability of humans to understand how an AI system makes decisions [7]. Without transparency, the decisions made by AI are often seen as "black boxes" that are difficult to interpret, even for the developers who designed the systems. This lack of clarity can erode public trust in AI and lead to skepticism about its fairness and reliability [7]. For AI to be ethical, it must be accountable to the people it impacts. This means ensuring that mechanisms for human oversight and AI systems can be audited and explained to prevent harmful or discriminatory outcomes [16].

Moreover, respecting human values is essential in privacy and individual rights. AI systems often rely on vast datasets containing sensitive personal information. Implementing AI can result in privacy breaches, discrimination, or exploitation without proper safeguards. For instance, AI-driven surveillance systems could monitor individuals without consent, compromising personal freedoms [12]. To prevent such ethical issues, AI systems should honor individuals' privacy and autonomy, following the principle that people have the right to control their data and how it is utilized [8].

Ethical AI must also reflect values like social equity and inclusivity. AI can significantly impact marginalized communities by exacerbating existing disparities or providing new opportunities for inclusion and empowerment. For instance, AI-powered healthcare systems must ensure equitable access to medical services, especially for vulnerable populations, and should not discriminate based on socioeconomic status or race [19]. Aligning AI with human values ensures that these technologies reduce inequalities rather than reinforce them.

4. The Need for a Regulatory Body or Guidelines to Govern AI Conduct

As Artificial Intelligence (AI) becomes increasingly integrated into society, concerns over its ethical, legal, and societal implications continue to grow. AI can perpetuate biases, infringe on privacy, and pose security risks without proper governance [7]. Just as corporations are regulated by frameworks such as the United Nations Guiding Principles on Business and Human Rights (UNGPs)—commonly known as the Ruggie Principles—there is an urgent need for a regulatory body or global guidelines to oversee AI development and deployment [7]. The Ruggie Principles provide a model for ensuring corporate accountability by

emphasizing human rights due diligence, corporate responsibility, and state obligations [5]. Similarly, AI governance must establish clear ethical and legal frameworks to prevent harm, promote transparency, and ensure AI aligns with human rights values.

4.1. The Ruggie Principles and Corporate Conduct

The Ruggie Principles, developed by John Ruggie under the United Nations mandate, establish a three-pillar framework: 1) the state duty to protect human rights, 2) corporate responsibility to respect human rights, and 3) access to remedies for victims of corporate abuses [5]. These principles hold corporations accountable for their impact on society, requiring them to implement due diligence processes to identify, prevent, and mitigate harm [5]. AI, much like corporations, has the potential to affect fundamental rights, and thus requires a similar governance approach to ensure its responsible use [20].

4.2. AI Governance: The Need for Global Regulatory Standards

AI systems influence critical areas such as law enforcement, healthcare, employment, and finance. Without a regulatory body or enforceable guidelines, AI may be used irresponsibly, exacerbating discrimination, privacy violations, and security threats [21]. Many experts have called for global AI regulations modeled after frameworks like the Ruggie Principles, emphasizing transparency, accountability, and ethical responsibility in AI development and deployment [14].

Several governments and international organizations have recognized this need. The European Union's AI Act and the OECD AI Principles aim to introduce regulatory oversight for AI, ensuring compliance with human rights and ethical standards [18]. However, these efforts remain fragmented. A unified global regulatory body, similar to those governing human rights and corporate conduct, would provide comprehensive oversight, preventing AI misuse and ensuring ethical alignment with human values.

4.3. AI's Impact on Human Rights and the Role of Governance

Just as the Ruggie Principles protect individuals from corporate abuses, AI governance must safeguard human rights from potential AI harms. Automated decision-making in hiring, policing, and social welfare has already demonstrated risks of bias, exclusion, and unfair treatment [22]. AI developers and corporations deploying AI must be held accountable for the consequences of their technologies [20]. Without a structured governance model, AI could contribute to systemic discrimination and reinforce social inequalities [10].

A global regulatory body or set of principles for AI governance should ensure:

- Transparency: AI systems should be explainable and auditable.
- Accountability: Developers and companies must be responsible for AI decisions.
- Human Rights Compliance: AI must align with fundamental rights and ethical

considerations.

- Remediation Mechanisms: Individuals impacted by AI-related harms should have avenues for redress, similar to corporate accountability under the Ruggie Principles.

5. Artificial Intelligence in the Legal Algorithm

5.1. AI Cannot Replace Judges in the Courtroom.

Integrating Artificial Intelligence (AI) into the legal system has revolutionized various aspects of legal practice, including legal research, contract analysis, and predictive analytics for case outcomes. However, despite AI's capabilities in processing vast amounts of data, identifying patterns, and making legal recommendations, it cannot replace judges in the courtroom. Judicial decision-making requires more than just the application of legal principles; it involves human judgment, ethical considerations, discretion, and an understanding of the broader societal impact of decisions. AI lacks the moral reasoning, empathy, and interpretive skills essential for judicial decision-making, making it unsuitable as a replacement for human judges.

One of the primary reasons AI cannot replace judges is the necessity of judicial discretion and moral reasoning in legal decision-making. Judges do not merely apply the law mechanically; they interpret statutes, assess the credibility of witnesses, and consider mitigating and aggravating factors unique to each case [14]. Legal reasoning often involves weighing competing interests and exercising moral and ethical judgment, which AI cannot do [3]. AI operates based on algorithms and pre-programmed decision-making models. Still, it cannot exercise the moral and ethical considerations that human judges apply when sentencing or resolving complex legal disputes [23].

A crucial aspect of judicial decision-making is the ability to empathize with individuals involved in legal proceedings. Judges consider cases' emotional, psychological, and social contexts, particularly in family law, criminal sentencing, and asylum cases [10]. AI, by contrast, lacks emotional intelligence and the ability to understand the human condition [16]. A justice system that is purely algorithmic risks making decisions that are devoid of compassion, which could lead to unfair or disproportionate legal outcomes.

AI systems are trained on existing legal data, meaning they can inherit biases in historical court decisions. Studies have shown that AI algorithms used in predictive policing and sentencing can reinforce racial and socioeconomic biases, disproportionately affecting marginalized communities [2]. Unlike human judges, AI cannot assess and challenge biased precedents critically. Additionally, AI-driven judicial decisions raise concerns about accountability. If an AI system makes an incorrect or unjust ruling, who is responsible? Judicial accountability is a cornerstone of legal systems, and replacing human judges with AI would create a vacuum in accountability [24].

Judges do not simply apply laws—they interpret them based on constitutional

principles, legal precedents, and evolving societal values. AI cannot understand legal philosophy, historical context, or changing jurisprudence [9]. Legal interpretation often involves subjective reasoning, as judges may interpret statutes differently based on their legal philosophy and societal considerations [25]. AI cannot engage in such interpretive rationale, making it unsuitable for judicial roles that require complex legal analysis.

The judiciary's legitimacy is based on public confidence in the fairness and impartiality of legal decisions. People are more likely to accept court rulings when they believe a human judge has carefully considered their case, rather than an impersonal algorithm [10]. Replacing judges with AI could erode trust in the judicial system and lead to public resistance against AI-driven justice [26]. The perception of justice is just as crucial as its administration, and human judges assure that legal decisions are made with human oversight and fairness.

5.2. AI Has no Common Sense

Artificial Intelligence (AI) has made significant strides in automating complex decision-making processes. Still, it fundamentally lacks common sense, moral reasoning, and the ability to understand human values like humans do. Because of these limitations, AI should be strictly regulated and kept away from control over nuclear weapons. The absence of ethical judgment, emotional intelligence, and the ability to assess the broader consequences of actions makes AI unsuitable for handling life-and-death decisions, particularly in warfare. Without robust legal and ethical safeguards, the use of AI in nuclear weapons systems poses an existential risk to humanity.

One of the most significant limitations of AI is its inability to apply common sense to decision-making. Unlike humans, AI does not possess an intrinsic understanding of right and wrong, nor can it comprehend the moral weight of its decisions [27]. AI functions based on pattern recognition and statistical probabilities, which may work well in structured environments but fail in unpredictable, high-stakes situations like nuclear warfare. A simple error in algorithmic calculations or a misinterpretation of incoming data could lead to catastrophic consequences [28].

Allowing AI to control or significantly influence nuclear weapons systems introduces unacceptable risks. AI lacks the capacity for strategic thinking, diplomacy, or the consideration of ethical implications in warfare. In the case of false alarms or ambiguous situations, human judgment is essential to prevent unnecessary escalation. Historical incidents, such as the 1983 Soviet nuclear false alarm, demonstrate how human discretion prevented atomic war when an automated system mistakenly detected incoming missiles [29]. If AI controlled such decisions, there would be no room for human intuition, skepticism, or last-minute intervention.

Strict laws and ethical guidelines must be established to prevent AI from being integrated into nuclear weapons systems. International agreements, such as the

Treaty on the Non-Proliferation of nuclear weapons (NPT), should be expanded to include explicit prohibitions on developing and deploying AI-driven nuclear decision-making systems [30]. Ethical frameworks should ensure that humans remain in control of all critical decisions related to nuclear weapons, as any reliance on AI for such matters could increase the risk of accidental or unauthorized launches.

AI-driven systems are susceptible to errors, adversarial manipulation, and cyberattacks. If hostile entities were to hack an AI-controlled nuclear command system, they could manipulate data to trigger unintended actions, leading to devastating consequences [31]. AI cannot question whether its information has been corrupted, making it an unreliable safeguard for something as destructive as nuclear weapons.

6. Conclusions

Artificial Intelligence (AI) has emerged as a transformative force across various sectors, enhancing efficiency, automating decision-making, and analyzing vast data. However, AI remains inherently insufficient despite its advancements, and it lacks legal and ethical frameworks to regulate its operations. The lack of transparency and accountability in AI systems and risks of bias, discrimination, privacy violations, and security concerns highlight the necessity for stringent legal oversight. Furthermore, AI cannot replace human judgment in critical areas such as the judiciary, where moral reasoning, empathy, and interpretive skills are indispensable.

AI's lack of common sense and ethical awareness further underscores the dangers of granting it control over high-stakes decision-making processes, such as nuclear weapons management. The potential for catastrophic errors, adversarial manipulation, and unintended escalations necessitates global legal restrictions to ensure AI remains a tool for human enhancement rather than a decision-maker in life-and-death situations. Additionally, technological advancements must be subject to democratic principles to ensure they benefit society, rather than reinforcing existing inequalities or posing new threats.

AI can only be ethical and beneficial if it aligns with human values and operates within a well-defined legal and regulatory framework. Science alone is not equipped to address the moral dilemmas posed by AI; law and governance must play a fundamental role in shaping AI's future. As AI evolves, international cooperation, interdisciplinary collaboration, and proactive legal reforms will ensure that AI serves humanity responsibly and ethically.

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] Binns, R. (2018) Fairness in Machine Learning: Lessons from Political Philosophy.

Proceedings of the 2018 *Conference on Fairness, Accountability, and Transparency*, New York, 23-24 February 2018, 149-159.

[2] O'Neil, C. (2016) Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. Crown Publishing.

[3] Scherer, M.U. (2016) Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies. *Harvard Journal of Law & Technology*, **29**, 353-400.

[4] Bowen, S.A. (2024) "If It Can Be Done, It Will Be Done:" AI Ethical Standards and a Dual Role for Public Relations. *Public Relations Review*, **50**, Article 102513. https://doi.org/10.1016/j.pubrev.2024.102513

[5] Ruggie, J.G. (2011) Guiding Principles on Business and Human Rights: Implementing the United Nations "Protect, Respect and Remedy" Framework. United Nations Human Rights Council (UNHRC).

[6] Burrell, J. (2016) How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms. *Big Data & Society*, **3**, 1-12. https://doi.org/10.1177/2053951715622512

[7] Angwin, J., Larson, J., Mattu, S. and Kirchner, L. (2016) Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased against Blacks. ProPublica. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

[8] Voigt, P. and Von dem Bussche, A. (2017) The EU General Data Protection Regulation (GDPR): A Practical Guide. Springer.

[9] Barocas, S., Hardt, M. and Narayanan, A. (2019) Fairness and Machine Learning: Limitations and Opportunities. MIT Press.

[10] Binns, R. (2018) Algorithmic Accountability and Public Reason. *Philosophy & Technology*, **31**, 543-556. https://doi.org/10.1007/s13347-017-0263-5

[11] Crawford, K. (2021) Atlas of AI: Power, Politics, and the Planet-Sized Implications of Artificial Intelligence. Yale University Press.

[12] Zuboff, S. (2019) The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. Public Affairs.

[13] Garvie, C., Bedoya, A. and Frankle, E. (2016) The Perpetual Line-Up: Unregulated Police Face Recognition in America. Georgetown Law Center on Privacy & Technology. https://www.perpetuallineup.org/

[14] Fjeld, J., Achten, N., Hilligoss, H., Nagy, A. and Srikumar, M. (2020) Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI. Berkman Klein Center for Internet & Society. https://doi.org/10.2139/ssrn.3518482

[15] Guszcza, J., Mahoney, S. and van der Pol, L. (2018) AI and the Responsibility Gap: Who's Liable for Autonomous Systems? Deloitte Insights. https://www2.deloitte.com/us/en/insights/industry/public-sector/ai-and-responsibility-gap.html

[16] Bryson, J.J., Diamantis, M.E. and Grant, T.D. (2017) Of, for, and by the People: The Legal and Ethical Implications of Autonomous Vehicles. *Proceedings of the* 2017 *International Conference on Robotics and Automation*, Singapore, 29 May-3 June 2017, 4199-4204.

[17] European Commission (2019) Ethics Guidelines for Trustworthy AI. European Commission. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

[18] European Commission (2021) Proposal for a Regulation on a European Approach for Artificial Intelligence (AI Act). European Union.

[19] Brynjolfsson, E. and McAfee, A. (2014) The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies. W. W. Norton & Company.

[20] Taylor, L., Floridi, L. and van der Sloot, B. (2021) Group Privacy: New Challenges of Data Technologies. Springer.

[21] Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M. and Floridi, L. (2017) Artificial Intelligence and the 'Good Society': The US, EU, and UK Approach. *Science and Engineering Ethics*, **24**, 505-528. https://doi.org/10.1007/s11948-017-9901-7

[22] Eubanks, V. (2018) Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. St. Martin's Press.

[23] Pasquale, F. (2020) New Laws of Robotics. Harvard University Press. https://doi.org/10.2307/j.ctv3405w6p

[24] Sartor, G. (2020) Artificial Intelligence and Judicial Decisions: A Critical Perspective on the Future of Adjudication. *Artificial Intelligence and Law*, **28**, 311-332.

[25] Solum, L.B. (2020) Legal Reasoning and Artificial Intelligence. Cambridge University Press.

[26] Sourdin, T. (2018) Judge V Robot? Artificial Intelligence and Judicial Decision-Making. *University of New South Wales Law Journal*, **41**, 1114-1133. https://doi.org/10.53637/zgux2213

[27] Marcus, G. and Davis, E. (2019) Rebooting AI: Building Artificial Intelligence We Can Trust. Pantheon Books.

[28] Bostrom, N. (2014) Superintelligence: Paths, Dangers, Strategies. Oxford University Press.

[29] Lewis, P. (2018) The Doomsday Machine: Confessions of a Nuclear War Planner. Penguin Books.

[30] Scharre, P. (2018) Army of None: Autonomous Weapons and the Future of War. W. W. Norton & Company.

[31] Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., *et al.* (2018) The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation. arXiv:1802.07228.