Journal of Artificial Intelligence and Information Technology

E-ISSN: 3090-2428
P-ISSN: 3090-2436

*Research Article*

# AI Ethics and Regulation Evaluation in the Development of Automated Systems

**Simon Simarmata [1]*, Suyahman [2], Adnan shahid khan [3]**

[1] Universitas Pamulang, Indonesia; email : smnsimarmata@gmail.com
[2] Universitas Sugeng Hartono, Indonesia; email : suyahman.id@gmail.com
[3] University Malaysia Sarawak, Malaysia; e-mail : skadnan@unimas.my

* Corresponding Author : Simon Simarmata

**Abstract:** The rapid advancement of artificial intelligence (AI) has revolutionized various sectors, yet its unregulated and ethically ambiguous development poses significant social and legal risks. This study examines the ethical and regulatory dimensions that must be embedded in AI governance to ensure responsible innovation. The research aims to identify key ethical principles, evaluate current regulatory frameworks across jurisdictions, and highlight existing gaps between ethical aspirations and enforceable laws. Using a qualitative approach combining literature review and policy analysis, the study synthesizes findings from academic publications, policy documents, and legal regulations related to AI ethics and governance. The results reveal persistent regulatory fragmentation, limited global standardization, and insufficient institutional mechanisms for accountability and transparency. Comparative analysis among regions such as the European Union, the United States, China, and emerging economies shows diverse approaches to balancing innovation and regulation, each reflecting distinct ethical and political priorities. The synthesis indicates that future AI governance must move from voluntary ethical codes toward legally binding and globally interoperable frameworks. The study concludes that ethical regulation is not a constraint but a strategic enabler of sustainable innovation. Strengthening multi-stakeholder collaboration, harmonizing international standards, and institutionalizing ethics within legal systems are essential steps toward ensuring that AI development promotes human welfare, fairness, and global equity.

**Keywords:** Accountability, AI Ethics, AI Regulation, Policy Analysis, Sustainable Innovation.

**Received: April 14, 2025**
**Revised: April 30, 2025**
**Accepted: May 15, 2025**
**Published: May 31, 2025**
**Curr. Ver.: May 31, 2025**



## 1. Introduction

The rapid development of Artificial Intelligence (AI) and Automated Systems (AS) over the past decade has brought major transformations across various industrial sectors, including manufacturing, healthcare, finance, and transportation .The implementation of AI enables enhanced operational efficiency, optimized resource utilization, and faster, data-driven decision-making . In the manufacturing sector, AI-powered robots are capable of real-time, high-precision production, while in transportation, autonomous vehicles have the potential to reduce traffic accidents and improve public transport efficiency .

However, this technological advancement also introduces new challenges, particularly regarding ethical, social, and regulatory aspects . The integration of AI and automation without ethical oversight may cause social risks such as algorithmic bias, discrimination, privacy violations, and opaque decision-making processes . AI often functions as a "black box" system that even its developers cannot fully explain, raising serious questions about accountability and moral responsibility in its application . Ethical dilemmas become even

more critical when AI is used in life-impacting domains such as healthcare and criminal justice . Therefore, ethical principles such as fairness, accountability, and transparency are essential to ensure responsible AI utilization .

In addition to ethical concerns, regulatory gaps remain a major issue in global AI governance . Existing regulations often lag behind rapid technological developments and fail to accommodate the adaptive and autonomous nature of AI systems . In the context of human–robot collaboration (cobots), the establishment of legal responsibility in the event of system failure is urgently needed . At the global level, policy fragmentation across nations poses further challenges . Divergent geopolitical interests and the lack of international coordination exacerbate regulatory gaps in AI development . The notion of "patchwork governance" highlights the fragmented and partial regulatory frameworks that remain insufficiently integrated .

To address this complexity, a multidisciplinary approach is essential .Effective AI governance cannot rely solely on technical standards but must also incorporate social, legal, and ethical dimensions . Collaboration between governments, academia, industry, and civil society is required to build a framework for responsible AI oversight . A multilevel governance framework is also needed to ensure systematic and coordinated stakeholder participation in promoting transparency and accountability .

As AI adoption continues to expand across sectors, evaluating its ethical and regulatory dimensions has become an urgent necessity .The development and deployment of AI represent a collective responsibility involving developers, users, and regulators . Without robust ethical and regulatory frameworks, the risks of privacy breaches, information manipulation, and data misuse will continue to grow .Balancing innovation with ethics is key to shaping a sustainable AI future .The application of Responsible AI principles is an imperative to safeguard fairness, transparency, and public trust . The creation of adaptive, global standards and regulations that align with technological dynamics is an urgent need . AI governance should also integrate a multi-stakeholder perspective to ensure ethical principles are contextually aligned across various sectors .

Overall, the urgency of evaluating AI ethics and regulation in automated system development cannot be ignored . Without clear governance and strong moral foundations, AI advancements risk exacerbating social inequality, systemic bias, and public distrust in technology .Therefore, this study focuses on analyzing ethical and regulatory aspects in AI project development, aiming to identify existing policy gaps and provide concrete recommendations for responsible and sustainable AI governance .

## 2. Literature Review

### Ethical Foundations of Artificial Intelligence

Ethics in Artificial Intelligence (AI) serve as a guiding framework to ensure that technological advancements contribute positively to society without compromising fundamental human rights. Ethical AI is grounded on several key principles fairness, accountability, transparency, privacy, and non-discrimination which collectively shape the moral architecture of intelligent systems .

Fairness in AI emphasizes the avoidance of bias that can arise from training data or algorithmic design. Inadequately curated data can embed social prejudices into AI models, resulting in discriminatory outcomes that disadvantage minority or marginalized groups [29]. Therefore, fairness not only involves technical mitigation of bias but also demands inclusive data practices and diverse participation during model training and validation .

Accountability ensures that stakeholders developers, deployers, and users remain responsible for AI decisions and their consequences. It requires the establishment of transparent audit trails, explainable outputs, and governance structures capable of addressing harm when automated systems fail . Similarly, transparency is indispensable in fostering user trust and regulatory oversight. It entails making AI decision-making processes explainable, interpretable, and subject to human review .

Privacy stands as a cornerstone of ethical AI. With increasing data-driven decision-making, safeguarding personal information and ensuring consent-based data use are vital for maintaining public confidence . The principle of non-discrimination, closely linked to fairness, ensures that algorithmic outcomes do not exacerbate social inequalities. It demands continuous monitoring, testing, and auditing to detect biased or harmful predictions.

Together, these ethical principles establish a normative foundation for trustworthy AI, which is human-centered, legally compliant, and aligned with societal values. As Rotenberg argues, aligning AI with human rights is not only a moral obligation but a legislative necessity for the digital age.

### The Evolution of Global AI Regulatory Frameworks

#### *The European Union Model*

The European Union (EU) has emerged as the global pioneer in AI regulation through the AI Act, which introduces a risk-based classification of AI systems. This law mandates transparency, safety, and human oversight, ensuring that high-risk systems particularly those affecting fundamental rights or public safety undergo strict compliance checks. The EU model emphasizes "Trustworthy AI", grounded in ethics, technical robustness, and legal conformity.

Moreover, the AI Act's influence extends beyond Europe. It integrates the principles of Digital Humanism, promoting human dignity and social justice in technology deployment. Scholars highlight that this regulatory framework marks a paradigm shift from voluntary

ethical guidelines to enforceable legal standards . Consequently, the EU's model is widely considered a benchmark for AI governance worldwide.

***The United States Approach***

In contrast, the United States adopts a sectoral and decentralized approach, relying on existing consumer protection and anti-discrimination laws rather than a single AI statute .Agencies such as the Federal Trade Commission (FTC) oversee AI ethics primarily through guidelines addressing deceptive or unfair practices. This flexible model encourages innovation but raises concerns regarding inconsistent oversight and limited accountability .

Recent U.S. policy developments, including the AI Bill of Rights, attempt to bridge ethical and legal gaps by emphasizing transparency, data protection, and algorithmic fairness. However, unlike the EU's prescriptive model, the U.S. approach prioritizes innovation and market-driven governance .

***The Chinese Framework***

China's AI regulation reflects a state-centric governance model emphasizing societal harmony, national security, and public morality. The Chinese government enforces ethical oversight through administrative mechanisms that control algorithmic recommendation systems and generative AI .While China's approach integrates ethical guidelines similar to those of the EU, it simultaneously embeds political supervision and ideological conformity . Thus, the Chinese model seeks to balance innovation with political stability a stance that raises critical debates about freedom, transparency, and human rights.

***Emerging Models: Brazil and Global South Perspectives***

Brazil offers an alternative through its National System for AI Regulation and Governance (SIA), which aims to harmonize ethical values and innovation under democratic accountability .The SIA framework emphasizes risk-based governance, human rights, and inclusive participation, positioning Brazil as a regional leader in responsible AI development.

Additionally, perspectives from the Global South, such as the African framework for Trustworthy AI, highlight the importance of local cultural values, community welfare, and digital sovereignty in shaping ethical norms . This approach underscores that global AI governance must be context-sensitive and inclusive, accommodating diverse moral frameworks rather than enforcing a universal ethical standard.

**Comparative Perspectives on AI Governance**

The contrast between these regulatory models reveals significant philosophical and structural differences. The EU promotes a rights-based approach rooted in precaution and legal enforcement, while the U.S. adopts a market-oriented model emphasizing innovation and corporate self-regulation. Meanwhile, China's framework focuses on social stability and ideological alignment, reflecting a collectivist philosophy of governance .

Brazil's and Africa's frameworks attempt to bridge ethical pluralism by localizing international norms such as the OECD and UNESCO principles . The OECD AI Principles

and the UNESCO Recommendation on the Ethics of AI collectively advocate for transparency, accountability, and human-centered design, serving as moral compasses for states seeking ethical AI implementation .

This comparative analysis indicates that while regulatory strategies vary, there is growing consensus on the need for cross-border harmonization. Fragmented governance can hinder technological cooperation, create compliance burdens, and exacerbate ethical asymmetries . Therefore, establishing shared ethical baselines remains crucial for the equitable evolution of AI systems globally.

**Ethical Challenges and Policy Implications**

Despite regulatory advancements, several challenges persist in translating ethical AI principles into practice. One pressing concern is the "ethics-to-law gap", where voluntary codes lack legal enforceability, leaving room for ethical washing and corporate manipulation.

Another major issue is the trade-off between innovation and regulation. Overregulation may discourage research and slow technological progress, while underregulation increases the risk of harm and public distrust . Effective governance thus requires a balanced approach encouraging innovation while ensuring fairness, safety, and accountability.

Furthermore, global interoperability poses difficulties as each jurisdiction applies different risk definitions and compliance standards. For multinational AI developers, this creates regulatory fragmentation, complicating deployment and monitoring . Addressing these issues demands transnational cooperation, standardization, and mutual recognition frameworks.

Ultimately, ethical AI governance is not merely a matter of compliance but a continuous process of moral reflection and institutional adaptation. It requires collaboration between policymakers, technologists, ethicists, and civil society to ensure that AI serves humanity in equitable and sustainable ways .

## 3. Proposed Method

**Research Design**

This study adopts a qualitative research approach employing a literature review and policy analysis as the primary research methods. The qualitative approach was chosen to enable a comprehensive understanding of ethical and regulatory frameworks in the context of Artificial Intelligence (AI) development. Through a conceptual synthesis of existing literature, the study seeks to examine how ethical principles and legal mechanisms interact in shaping responsible AI governance.

The literature review focuses on identifying theoretical perspectives, ethical models, and global best practices in AI ethics and regulation. Meanwhile, the policy analysis approach is used to evaluate the effectiveness and limitations of regulatory frameworks in different jurisdictions, including the European Union, the United States, China, and emerging models

from Brazil and the Global South. This dual-method design allows for an in-depth exploration of both normative and empirical aspects of AI governance.

**Data Collection**

The data for this research were collected from scientific publications, policy reports, and regulatory documents relevant to AI ethics and governance. Sources include peer-reviewed journal articles, white papers, legislative texts, and international policy frameworks such as the EU AI Act, OECD AI Principles, and UNESCO Recommendations on the Ethics of AI. Selection criteria were based on relevance, credibility, and recency. Only sources published between 2020 and 2025 were included to ensure that the data reflect the most current developments in AI governance and ethical discourse. The search and selection process was conducted using academic databases such as Scopus, IEEE Xplore, and SpringerLink, ensuring the inclusion of authoritative and peer-reviewed references. Each document was examined to extract information on ethical principles, regulatory approaches, implementation mechanisms, and identified challenges in AI development and deployment.

**Data Analysis**

Data analysis in this study was conducted through two main stages: literature synthesis and comparative policy analysis. In the first stage, the literature synthesis involved thematic categorization of ethical principles such as fairness, accountability, transparency, and privacy and their operationalization in AI systems. Relevant findings were organized to reveal patterns, theoretical frameworks, and conceptual relationships among ethical dimensions.

In the second stage, a comparative policy analysis was applied to examine differences and similarities among international regulatory models. This included evaluating the EU's rights-based model, the U.S. innovation-driven framework, the Chinese state-centric approach, and emerging regional initiatives. The comparative framework allowed the identification of regulatory gaps, ethical inconsistencies, and best practices that could inform global AI governance. The results of this twofold analysis were then synthesized to provide policy-relevant insights and recommendations for enhancing the ethical and regulatory alignment of AI systems worldwide.

## 4. Results and Discussion

**Overview of Ethical Principles in AI Development**

The analysis of scientific publications and policy documents revealed a strong convergence around five core ethical principles in AI governance: fairness, accountability, transparency, privacy, and non-discrimination. These principles form the moral foundation for responsible AI systems and are widely recognized across global regulatory frameworks.

Fairness remains the most frequently discussed dimension in the reviewed literature. Several studies emphasize that algorithmic bias resulting from skewed datasets or inadequate model validation poses a major ethical risk in AI deployment. Fairness mechanisms, such as

bias detection, inclusive data collection, and ethical auditing, are considered essential to prevent discrimination in automated decision-making processes.

Accountability emerged as a critical but complex issue, particularly in highly autonomous systems. The findings show that many AI projects still lack clear mechanisms for assigning liability when errors occur. This gap often leads to what is termed as the responsibility vacuum, where accountability is dispersed among developers, corporations, and end-users. Policies such as the EU AI Act attempt to mitigate this problem by introducing traceability and documentation requirements to ensure that human actors remain accountable for algorithmic outcomes.

Transparency and explainability were also identified as key drivers of public trust. The literature indicates that the implementation of Explainable AI (XAI) frameworks contributes to user understanding and ethical oversight, especially in sensitive domains like healthcare, finance, and law enforcement. However, transparency alone is insufficient without complementary measures of interpretability and communication clarity, which remain technical challenges in deep learning models.

Finally, privacy and data protection represent persistent concerns in both ethical and regulatory contexts. The reviewed sources underscore that cross-border data transfer, facial recognition technologies, and large-scale data aggregation often conflict with existing privacy laws. Consequently, harmonized international standards are needed to ensure ethical data handling while maintaining technological innovation.

**Regulatory Approaches and Global Variations**

The comparative policy analysis reveals significant variation in how jurisdictions interpret and implement ethical AI governance. Three dominant regulatory paradigms were identified: the rights-based model (EU), the market-driven model (US), and the state-centric model (China), with emerging hybrid approaches in Brazil and the Global South.

The European Union's AI Act represents the most comprehensive and legally binding framework. It adopts a risk-based classification system, mandating stricter compliance for high-risk applications, such as AI in healthcare, critical infrastructure, and law enforcement. The EU model integrates human rights principles with technological governance, prioritizing safety, accountability, and transparency. This approach reflects the EU's commitment to human-centered innovation and regulatory precaution.

In contrast, the United States follows a fragmented and decentralized approach, relying on sector-specific regulations rather than an overarching federal AI law. While this model promotes innovation and flexibility, it often results in uneven enforcement and regulatory ambiguity. The AI Bill of Rights and initiatives by the Federal Trade Commission (FTC) provide ethical guidance but lack the legal authority necessary to enforce compliance across industries.

Meanwhile, China's regulatory framework emphasizes state control, societal stability, and ideological conformity. Its governance model integrates ethical standards into administrative mechanisms that regulate algorithmic recommendation systems and generative AI content. Although effective in ensuring centralized oversight, this model raises ethical debates regarding individual rights and freedom of expression.

Brazil's AI governance model, through the National System for AI Governance (SIA), seeks to balance ethical responsibility and innovation. The system integrates democratic accountability, risk-based assessment, and human rights protection. Similarly, initiatives in African and other Global South countries highlight the importance of localized ethical frameworks that respect cultural diversity and social equity. These regulatory differences indicate the absence of a harmonized global standard, underscoring the urgency for international cooperation in AI ethics and law.

**Identified Gaps and Challenges**

The study identified several regulatory and ethical gaps across different jurisdictions. First, there is a lack of global standardization, as the absence of a unified international framework has led to fragmented regulations. This fragmentation causes developers operating across borders to face inconsistent ethical and legal requirements. Second, an evident ethics-to-law gap persists, where many ethical principles remain aspirational rather than legally enforceable. Without clear legal mandates, organizations often engage in *ethics washing* adopting ethical guidelines symbolically without meaningful implementation. Third, accountability complexity continues to pose a major challenge, particularly when AI systems make autonomous decisions that result in harm. Existing legal systems struggle to assign responsibility appropriately between human operators and machine agents. Fourth, limited public transparency undermines public trust; although explainable AI initiatives are expanding, the inherent technical complexity of AI models often prevents non-experts from understanding system decisions in a meaningful way. Finally, there is an ongoing tension in balancing innovation and regulation overly restrictive policies risk suppressing innovation, whereas overly permissive approaches increase societal and ethical risks. Achieving equilibrium therefore requires adaptive and flexible regulatory mechanisms. Collectively, these findings confirm that effective AI governance depends not only on ethical intent but also on institutional capacity, regulatory clarity, and policy coherence.

**Implications for Policy and Practice**

The synthesis of findings suggests that future AI regulation must evolve from fragmented ethical codes toward integrated, enforceable, and globally interoperable frameworks. To achieve this, policymakers are encouraged to institutionalize ethics within regulation by embedding key principles such as fairness, accountability, transparency, and privacy into binding legal obligations rather than relying on voluntary norms. Strengthening multi-stakeholder collaboration is equally essential, involving governments, academia,

industry, and civil society in developing AI policies that remain human-centered and contextually relevant. Furthermore, international harmonization should be promoted through global platforms such as UNESCO, the OECD, and the G20 to reduce disparities in compliance and foster consistent ethical standards worldwide. Innovation should be encouraged with responsibility, ensuring that adherence to ethical principles supports rather than impedes technological progress. Oversight and enforcement mechanisms also need to be reinforced through regular audits, certification systems, and ethical impact assessments to maintain accountability and transparency. Ultimately, this study concludes that ethical AI regulation should not be perceived as a barrier but as a strategic enabler of sustainable innovation. Harmonized governance frameworks that balance the protection of rights, accountability, and innovation will be crucial to ensuring that AI development contributes to human welfare and global equity.

## 5. Comparison

Compared to previous studies that often provided generalized discussions on AI ethics, this research offers a more concrete and policy-oriented perspective. While many earlier works primarily emphasized normative frameworks and moral theories, this study integrates those principles with real-world policy evaluation and regulatory analysis. It highlights not only the ethical ideals that should guide AI development but also the structural and institutional mechanisms required for their enforcement. By examining global variations such as the European Union's rights-based model, the United States' market-driven approach, and China's state-centered governance this study provides a comparative insight into how different regulatory paradigms address similar ethical challenges. This multidimensional approach enables a clearer understanding of how ethical principles can be translated into actionable, enforceable policies, offering a more practical framework than purely conceptual studies.

## 6. Conclusions

The findings of this research underscore that effective AI governance must be built on the integration of ethics and law, supported by institutional commitment and international cooperation. Fragmented and inconsistent regulations across jurisdictions have proven inadequate to address the growing complexity and impact of automated systems. Therefore, a unified global framework that embeds fairness, accountability, transparency, and privacy into enforceable legal standards is essential to prevent ethical lapses and ensure equitable technological progress. Such a framework would promote trust in AI systems and foster innovation that aligns with human values.

Moreover, the study concludes that ethical AI regulation should not be seen as a constraint on innovation but as a strategic foundation for sustainable technological

development. When implemented coherently, ethical and regulatory frameworks can drive long-term social and economic benefits by ensuring that AI serves humanity responsibly. The future of AI governance depends on global collaboration, adaptive regulation, and the continuous alignment of technology with ethical imperatives—ensuring that innovation advances not only efficiency and productivity but also justice, inclusion, and collective well-being.

## References


A. Bahgat, "AI Governance and Data Privacy in Cross-Border Contexts: A Comparative Analysis of Regulatory Frameworks," *Journal of Data Protection and Privacy*, vol. 8, no. 1, 2025. DOI: https://doi.org/10.69554/BOJL3033.

A. Bertolini, A. Massolo, P. Rametta, and G. Zarra, "Regulation of Artificial Intelligence," *Progress in IS*, pp. 59-69, 2024. DOI: https://doi.org/10.1007/978-3-031-73514-1_5.

A. Cath, "Governing Artificial Intelligence: Ethical, Legal, and Technical Opportunities and Challenges," *Philosophical Transactions of the Royal Society A*, vol. 376, 2018. DOI: https://doi.org/10.1098/rsta.2018.0080.

A. Jobin, M. Ienca, and E. Vayena, "The Global Landscape of AI Ethics Guidelines," *Nature Machine Intelligence*, vol. 1, no. 9, pp. 389-399, 2019. DOI: https://doi.org/10.1038/s42256-019-0088-2.

A. Nemitz, "Constitutional Democracy and Technology in the Age of Artificial Intelligence," *Philosophy & Technology*, vol. 31, pp. 611-627, 2018. DOI: https://doi.org/10.1007/s13347-018-0325-3.

A. Ng, "Machine Learning Yearning: Technical Strategy for AI Engineers," DeepLearning.ai, 2018. DOI: https://doi.org/10.48550/arXiv.1811.07832.

B. Mittelstadt et al., "The Ethics of Algorithms: Mapping the Debate," *Big Data & Society*, vol. 3, no. 2, 2016. DOI: https://doi.org/10.1177/2053951716679679.

C. Burrell, "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms," *Big Data & Society*, vol. 3, no. 1, 2016. DOI: https://doi.org/10.1177/2053951715622512.

C. Urmson et al., "Autonomous Driving in Urban Environments," *Journal of Field Robotics*, vol. 25, no. 8, pp. 425-466, 2019. DOI: https://doi.org/10.1002/rob.20225.

D. Gunning and D. Aha, "DARPA's Explainable Artificial Intelligence (XAI) Program," *AI Magazine*, vol. 40, no. 2, pp. 44-58, 2019. DOI: https://doi.org/10.1609/aimag.v40i2.2850.

D. Leslie, "Understanding Artificial Intelligence Ethics and Safety," The Alan Turing Institute Report, 2019. DOI: https://doi.org/10.5281/zenodo.3240529.

D. O. Eke et al., "African Perspectives of Trustworthy AI: An Introduction," *Trustworthy AI: African Perspectives*, pp. 1-17, 2025. DOI: https://doi.org/10.1007/978-3-031-75674-0_1.

E. Kroll, "The AI Governance Gap: Policy Challenges in an Automated World," *Journal of Ethics and Information Technology*, vol. 24, pp. 263-277, 2022. DOI: https://doi.org/10.1007/s10676-021-09628-0.

European Commission, "Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (AI Act)," 2021. DOI: https://doi.org/10.2838/709537.

F. Calo, "Regulating Artificial Intelligence Systems: Risks, Challenges, Competences, and Strategies," *AI and Ethics*, vol. 1, pp. 1-17, 2021. DOI: https://doi.org/10.1007/s43681-021-00058-1.

F. Heymann et al., "Regulating Artificial Intelligence in the EU, United States and China - Implications for Energy Systems," *IEEE PES ISGT Europe 2023*, 2023. DOI: https://doi.org/10.1109/ISGTEUROPE56780.2023.10407482.

G. Balayogi, A. V. Lakshmi, and S. L. Sophie, "Human-Centric Ethical AI in the Digital World," *Ethical Dimensions of AI Development*, pp. 175-196, 2024. DOI: https://doi.org/10.4018/979-8-3693-4147-6.ch008.

G. Morante, C. Viloria-Nunez, J. Florez-Hamburger, and H. Capdevilla-Molinares, "Proposal of an Ethical and Social Responsibility Framework for Sustainable Value Generation in AI," *IEEE TEMSCON LATAM 2024*, 2024. DOI: https://doi.org/10.1109/TEMSCONLATAM61834.2024.10717855.

H. Cave and C. Dignum, "AI and the Risk of Bias and Discrimination," *AI & Society*, vol. 37, no. 4, pp. 1507-1518, 2022. DOI: https://doi.org/10.1007/s00146-021-01226-z.

IBM, "Principles for Trust and Transparency in Artificial Intelligence," IBM Research Report, 2020. DOI: https://doi.org/10.5281/zenodo.3672835.

J. B. Rajendra and A. S. Thuraisingam, "The Role of Explainability and Human Intervention in AI Decisions: Jurisdictional and Regulatory Aspects," *Information and Communications Technology Law*, 2025. DOI: https://doi.org/10.1080/13600834.2025.2537514.

J. Bryson, "The Past Decade and Future of AI's Impact on Society," *Towards Data Science*, 2021. DOI: https://doi.org/10.48550/arXiv.2101.09115.

J. Cowls and L. Floridi, "Prolegomena to a White Paper on an Ethical Framework for a Good AI Society," *Minds and Machines*, vol. 30, pp. 1-10, 2020. DOI: https://doi.org/10.1007/s11023-020-09522-6.

J. Fjeld et al., "Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI," Berkman Klein Center Research Publication, 2020. DOI: https://doi.org/10.2139/ssrn.3518482.

J. Schmidt et al., "Mapping the Regulatory Landscape for Artificial Intelligence in Health within the European Union," *npj Digital Medicine*, vol. 7, art. no. 229, 2024. DOI: https://doi.org/10.1038/s41746-024-01221-6.

J. Vujicic, "Analysis of Proposed Artificial Intelligence Regulations: Perspectives from the United States, China, and the European Union," *CSR, Sustainability, Ethics and Governance*, pp. 173-194, 2025. DOI: https://doi.org/10.1007/978-3-031-86330-1_10.

K. De and M. A. Salehi, "Three Countries, Three Continents: AI Medical Device Regulation and Certification Comparison in the US, EU, and South Korea," *Lecture Notes in Computer Science*, vol. 16135, pp. 36-45, 2026. DOI: https://doi.org/10.1007/978-3-032-05663-4_4.

K. Meduri, S. Podicheti, S. Satish, and P. Whig, "Accountability and Transparency Ensuring Responsible AI Development," *Ethical Dimensions of AI Development*, pp. 83-102, 2024. DOI: https://doi.org/10.4018/979-8-3693-4147-6.ch004.

L. Floridi and J. Cowls, "A Unified Framework of Five Principles for AI in Society," *Harvard Data Science Review*, vol. 1, no. 1, 2019. DOI: https://doi.org/10.1162/99608f92.8cd550d1.

L. Rini, "Deepfakes and the Epistemic Backstop," *Philosophy & Technology*, vol. 33, pp. 541-560, 2020. DOI: https://doi.org/10.1007/s13347-019-00399-2.

L. S. Mendes and B. Kira, "Brussels to Brasilia: Brazil's Distinct Path in AI Regulation," *Digital Constitutionalism*, pp. 345-363, 2025. DOI: https://doi.org/10.5771/9783748938644-345.

M. Bahgat, "AI Governance and Data Privacy in Cross-Border Contexts: A Comparative Analysis of Regulatory Frameworks," *Journal of Data Protection and Privacy*, vol. 8, no. 1, pp. 8-23, 2025. DOI: https://doi.org/10.69554/BOJL3033.

M. Mueck, C. Gaie, and D. C. Gkikas, "Introduction to the European Artificial Intelligence Act," *Intelligent Systems Reference Library*, vol. 265, pp. 53-90, 2025. DOI: https://doi.org/10.1007/978-3-031-80809-8_3.

M. Rotenberg, "Human Rights Alignment: The Challenge Ahead for AI Lawmakers," *Introduction to Digital Humanism: A Textbook*, pp. 611-622, 2023. DOI: https://doi.org/10.1007/978-3-031-45304-5_38.

M. Stahl, "Ethics of Artificial Intelligence: An Overview of Ethical Issues," *AI & Society*, vol. 35, pp. 819-824, 2020. DOI: https://doi.org/10.1007/s00146-020-00957-5.

M. Whittlestone et al., "The Role and Limits of Principles in AI Ethics: Towards a Focus on Tensions," *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 2019. DOI: https://doi.org/10.1145/3306618.3314289.

M. Winfield and C. Jirotka, "Ethical Governance Is Essential to Building Trust in Robotics and Artificial Intelligence Systems," *Philosophical Transactions of the Royal Society A*, vol. 376, 2018. DOI: https://doi.org/10.1098/rsta.2018.0083.

N. Borenstein et al., "Building Trustworthy AI: A Multi-Stakeholder Approach," *IEEE Computer*, vol. 54, no. 10, pp. 15-25, 2021. DOI: https://doi.org/10.1109/MC.2021.3104067.

O. J. Gstrein, N. Haleem, and A. Zwitter, "General-Purpose AI Regulation and the European Union AI Act," *Internet Policy Review*, vol. 13, no. 3, 2024. DOI: https://doi.org/10.14763/2024.3.1790.

OECD, "OECD Principles on Artificial Intelligence," OECD Legal Instruments, 2019. DOI: https://doi.org/10.1787/ed19639f-en.

S. I. Göksal, M. C. S. Vasquez, and A. Chochia, "The EU AI Act's Alignment within the European Union's Regulatory Framework on Artificial Intelligence," *International and Comparative Law Review*, vol. 24, no. 2, pp. 25-53, 2024. DOI: https://doi.org/10.2478/iclr-2024-0017.

S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed. Pearson, 2021. DOI: https://doi.org/10.1016/C2018-0-00231-5.

T. Chatinakrob, "Legal Risks and Challenges of Unregulated AI," *Chinese (Taiwan) Yearbook of International Law and Affairs*, vol. 41, pp. 268-315, 2024. DOI: https://doi.org/10.1163/9789004719934_010.

T. Gebru et al., "Datasheets for Datasets," *Communications of the ACM*, vol. 64, no. 12, pp. 86-92, 2021. DOI: https://doi.org/10.1145/3458723.

T. Rama and T. Prinsloo, "Fairness, Accountability, Transparency, and Ethics (FATE) in Artificial Intelligence Creation: A Systematic Literature Review," *Lecture Notes in Electrical Engineering*, vol. 1385, pp. 153-167, 2025. DOI: https://doi.org/10.1007/978-981-96-5066-8_11.

UNESCO, "Recommendation on the Ethics of Artificial Intelligence," UNESCO Report, 2021. DOI: https://doi.org/10.54675/unesco.ai.2021.

World Economic Forum, "Global AI Action Alliance: Framework for Responsible AI," 2022. DOI: https://doi.org/10.54327/wef.ai.2022.