ARTIFICIAL INTELLIGENCE, ETHICS, AND HUMAN RIGHTS: AN ANALYTICAL PERSPECTIVE

Ankita Shukla[1], Dr. Shameem Ahmed Khan[2]

Research Scholar, Department of Law, ISBM University[1]

Professor, Department of Law, ISBM University[2]

Abstract

The emergence of artificial intelligence has fundamentally transformed contemporary society, presenting unprecedented challenges to established human rights frameworks and ethical principles. This paper examines the intricate relationship between artificial intelligence technologies, ethical considerations, and fundamental human rights through an analytical lens grounded in international legal instruments and emerging regulatory frameworks. The research analyzes key legal developments including the European Union's Artificial Intelligence Act (2024), UNESCO's Recommendation on the Ethics of Artificial Intelligence (2021), and the Council of Europe's Framework Convention on AI and Human Rights (2024), alongside foundational human rights instruments such as the Universal Declaration of Human Rights. The analysis reveals critical tensions between technological innovation and human rights protection, particularly concerning privacy, non-discrimination, freedom of expression, and algorithmic accountability. This paper argues that effective governance of AI requires a human-centric approach anchored in international human rights law, supported by robust regulatory mechanisms, and informed by ethical principles that prioritize human dignity, transparency, and accountability in AI system lifecycles.

Keywords: *Artificial Intelligence[1], Human Rights[2], Ethics[3], Data Protection[4], Algorithmic Accountability[5]*

1. Introduction

The rapid advancement of artificial intelligence technologies represents one of the most transformative developments of the twenty-first century, fundamentally reshaping social, economic, and legal landscapes globally. From automated decision-making systems in employment and credit allocation to facial recognition technologies in law enforcement, AI applications increasingly mediate critical aspects of human life and governance. While these technologies offer tremendous potential for societal benefit, they simultaneously pose significant risks to fundamental human rights and ethical values that underpin democratic societies. The intersection of artificial intelligence, ethics, and human rights has emerged as a critical area of legal scholarship and policy development. International organizations, regional bodies, and national governments have begun constructing regulatory frameworks to govern AI development and deployment. The European Union adopted the Artificial Intelligence Act in June 2024, establishing the world's first comprehensive legal framework for AI regulation.[1] UNESCO's member states unanimously adopted the Recommendation on the Ethics of Artificial Intelligence in November 2021, providing global ethical guidance.[2] Most recently, the Council of Europe

---

[1] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), Official Journal of the European Union, L 2024/1689, entered into force on 1 August 2024.

[2] UNESCO, Recommendation on the Ethics of Artificial Intelligence, adopted by the General Conference at its 41st session (23 November 2021), UNESCO Doc. SHS/BIO/REC-AIETHICS/2021, 2021.

opened for signature the Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law in September 2024, creating the first legally binding international treaty specifically addressing AI and human rights.[3]

These developments reflect growing international consensus that AI governance must be anchored in human rights principles and ethical considerations. The Universal Declaration of Human Rights (UDHR), adopted in 1948, established foundational rights including privacy, equality, non-discrimination, and freedom of expression.[4] These rights, developed in a pre-digital era, now face novel challenges from AI systems capable of mass surveillance, automated profiling, and decision-making that can significantly affect individuals' life opportunities without meaningful human oversight. The ethical dimensions of AI encompass multiple considerations including fairness, transparency, accountability, safety, and human dignity. These ethical principles increasingly inform legal frameworks, creating a symbiotic relationship between ethics and law in AI governance. However, significant questions remain regarding the adequacy of existing human rights frameworks to address AI-specific challenges, the effectiveness of emerging regulatory approaches, and the balance between innovation and rights protection. This paper undertakes a comprehensive analytical examination of the relationship between artificial intelligence, ethics, and human rights, focusing on international legal developments and their implications for rights protection in the AI age.

2. Objectives of The Study

The primary objectives of this research are:

1. To analyze the impact of artificial intelligence on key human rights, including privacy, equality, freedom of expression, and due process.
2. To assess the adequacy of international, regional, and national AI regulatory frameworks in safeguarding human rights and ensuring ethical AI development.

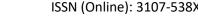## 3. Artificial Intelligence: Conceptual Framework and Technological Context

Artificial intelligence refers to systems with the capability to process data and information in ways that resemble intelligent behavior, typically involving aspects of reasoning, learning, perception, prediction, planning, or control.[5] The technological landscape of AI encompasses various approaches including machine learning, deep learning, natural language processing, computer vision, and expert systems. Modern AI systems increasingly rely on large-scale data processing, algorithmic decision-making, and automated pattern recognition. The capabilities of contemporary AI systems have expanded dramatically, enabling applications across virtually every sector of society. In healthcare, AI assists in diagnostic imaging and treatment recommendations. In finance, algorithmic systems make credit decisions and detect fraudulent transactions. Law enforcement agencies deploy facial recognition and predictive policing tools. Educational institutions use AI for admissions decisions and personalized learning. Employment processes increasingly incorporate AI-powered screening and assessment tools. This pervasive deployment of AI systems creates multiple touchpoints where technology intersects with human rights. The automated, scalable, and often opaque nature of AI decision-making raises fundamental questions about human agency, dignity, and the protection of rights in increasingly automated societies. The concentration of AI capabilities in a limited number of organizations and jurisdictions further amplifies concerns about power imbalances and equitable access to AI benefits.

---

[3] Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, opened for signature 5 September 2024, Council of Europe Treaty Series No. 225 (adopted 17 May 2024).

[4] Universal Declaration of Human Rights, G.A. Res. 217A (III), U.N. Doc A/810 at 71 (1948).

[5] Article 3(1), Council of Europe Framework Convention on AI defines AI system as "a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations or decisions that may influence physical or virtual environments."

237 | P a g e
www.mjapjournal.com
2025/MJAP/09/0018

4. Human Rights Framework: Foundational Principles

The international human rights framework provides the normative foundation for evaluating AI technologies' impact on individuals and societies. The Universal Declaration of Human Rights establishes core rights that remain applicable in the digital age, though their interpretation and application require adaptation to technological contexts.

Right to Privacy and Data Protection

Article 12 of the UDHR declares: "No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation."[6] This fundamental right faces unprecedented challenges from AI systems that process vast quantities of personal data, often without individuals' meaningful knowledge or consent. AI-powered surveillance systems, data mining operations, and profiling technologies can reveal intimate details about individuals' lives, associations, and behaviors. The European Union's General Data Protection Regulation (GDPR), adopted in 2016, provides the most comprehensive legal framework for data protection globally.[7] Article 22 of the GDPR specifically addresses automated decision-making, establishing that individuals have "the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her."[8] This provision reflects recognition that automated decision-making poses distinct risks to human dignity and autonomy. India's Digital Personal Data Protection Act, 2023, represents a significant development in data protection law in South Asia.[9] Following the landmark Puttaswamy judgment by the Supreme Court of India recognizing privacy as a fundamental right under Article 21 of the Constitution,[10] the DPDPA establishes comprehensive obligations for data fiduciaries regarding consent, data minimization, and security. However, the Act does not specifically address AI-related challenges such as algorithmic profiling and automated decision-making, leaving regulatory gaps in AI governance.

Equality and Non-Discrimination

Human rights law mandates equal treatment and prohibits discrimination based on protected characteristics including race, sex, religion, national origin, and other status. AI systems, however, can perpetuate and amplify discriminatory patterns through biased training data, flawed algorithms, or discriminatory applications. Research has documented AI systems exhibiting racial bias in facial recognition, gender bias in recruitment tools, and socioeconomic discrimination in credit scoring. The EU AI Act directly addresses discrimination risks, defining as prohibited AI practices those that deploy "subliminal techniques beyond a person's consciousness" to manipulate behavior in ways causing harm, exploit vulnerabilities of specific groups, or enable social scoring by governments.[11] The Act's risk-based framework classifies AI systems with significant discrimination potential as "high-risk," subjecting them to strict requirements including data governance, transparency, human oversight, and accuracy standards.

Freedom of Expression and Information

---

[6] Article 12, Universal Declaration of Human Rights, 1948.

[7] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation), Official Journal L 119, 4 May 2016.

[8] Article 22(1), GDPR.

[9] Digital Personal Data Protection Act, 2023, No. 22 of 2023, Act of Parliament, India (11 August 2023).

[10] Justice K.S. Puttaswamy (Retd.) v. Union of India, (2017) 10 SCC 1, Writ Petition (Civil) No. 494 of 2012, Supreme Court of India.

[11] Article 5, Regulation (EU) 2024/1689 (EU AI Act).

Article 19 of the UDHR establishes that "everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas through any media and regardless of frontiers."[12] AI technologies present complex implications for this right. While AI can enhance information access and facilitate communication, it simultaneously enables unprecedented content moderation, censorship, and manipulation. AI-powered content recommendation systems shape information environments, potentially creating filter bubbles and echo chambers that limit exposure to diverse perspectives. Generative AI systems raise concerns about misinformation, deepfakes, and synthetic media that can undermine public discourse. Conversely, AI-enabled surveillance and content monitoring can chill free expression and enable authoritarian control over information flows.

5. Ethical Principles for Artificial Intelligence

Ethics provides normative guidance for AI development and deployment, complementing legal frameworks with principles that reflect societal values and aspirations. Multiple international initiatives have articulated ethical principles for AI, demonstrating substantial convergence around core values.

UNESCO's Ethical Framework

UNESCO's Recommendation on the Ethics of Artificial Intelligence, adopted by 193 member states, establishes four foundational values: respect for human rights and human dignity; living in peaceful, just and interconnected societies; ensuring diversity and inclusiveness; and environment and ecosystem flourishing.[13] These values translate into ten core principles including proportionality, safety and security, privacy and data protection, multi-stakeholder governance, accountability, and transparency. The UNESCO framework emphasizes a human rights-centered approach, explicitly stating that "respect, protection and promotion of human rights and fundamental freedoms and human dignity" constitutes the cornerstone of ethical AI. The Recommendation mandates that member states ensure AI systems do not displace ultimate human responsibility and accountability, and requires ethical impact assessments throughout AI system lifecycles.

Transparency and Explainability

Transparency in AI systems encompasses multiple dimensions: transparency about the existence and use of AI systems; transparency regarding data collection and processing; and explainability of AI decision-making processes. The "black box" problem in complex machine learning systems creates fundamental challenges for accountability and rights protection. When individuals cannot understand how AI systems reach decisions affecting them, meaningful contestation and remedy become difficult or impossible. The EU AI Act addresses transparency through multiple provisions. Article 13 requires that high-risk AI systems be designed to enable deployers to understand outputs and use systems appropriately. Article 50 establishes transparency obligations for certain AI systems, including requirements that individuals be informed when interacting with AI systems such as chatbots, and that AI-generated content be clearly labeled.[14]

Human Oversight and Accountability

Ethical AI governance requires meaningful human oversight of automated systems, particularly those making decisions with significant impacts on individuals. The principle of "human-in-the-loop" reflects the understanding that critical decisions affecting human rights should not be delegated entirely to machines. The EU AI Act mandates human oversight for high-risk AI systems, requiring that natural persons assigned

---

[12] Article 19, Universal Declaration of Human Rights, 1948.
[13] UNESCO Recommendation on the Ethics of Artificial Intelligence, Section III (Values and Principles), 2021.
[14] Article 50, Regulation (EU) 2024/1689 (EU AI Act).C

oversight roles have necessary competence, training, and authority.[15] Accountability mechanisms must ensure that individuals harmed by AI systems have effective remedies. This requires clear allocation of responsibility among AI system developers, deployers, and other actors in the AI value chain. The Council of Europe Framework Convention emphasizes remedies, requiring parties to ensure individuals have effective remedies for violations of their rights, including the right to an explanation of decisions, the right to challenge decisions, and the right to seek correction of personal data.[16]

6. Emerging Legal Frameworks for Ai Governance

The past several years have witnessed rapid development of AI-specific legal frameworks at international, regional, and national levels. These frameworks reflect diverse regulatory philosophies while exhibiting substantial convergence around core principles.

European Union AI Act

The EU AI Act, which entered into force on 1 August 2024, represents the world's first comprehensive, binding legal framework for AI regulation. The Act adopts a risk-based approach, categorizing AI systems according to the risks they pose to safety and fundamental rights.[17] The Act prohibits certain AI practices deemed to pose unacceptable risks, including AI systems that deploy subliminal manipulation, exploit vulnerabilities, enable social scoring by public authorities, and use real-time remote biometric identification in publicly accessible spaces (with limited exceptions). High-risk AI systems including those used in employment, education, law enforcement, migration management, and administration of justice must comply with strict requirements before market placement. The Act's purpose, as stated in Article 1, is "to improve the functioning of the internal market and promote the uptake of human-centric and trustworthy artificial intelligence (AI), while ensuring a high level of protection of health, safety, fundamental rights enshrined in the Charter, including democracy, the rule of law and environmental protection."[18] This articulation explicitly links AI regulation to human rights protection, reflecting the understanding that effective AI governance must be rights-respecting. The Act establishes a governance architecture including the European AI Office, national supervisory authorities, an AI Board for coordination, and advisory mechanisms. Enforcement includes substantial penalties for violations, with maximum fines reaching €35 million or 7% of worldwide annual turnover for prohibited AI practices.

Council of Europe Framework Convention

The Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, adopted in May 2024 and opened for signature in September 2024, constitutes the first legally binding international treaty specifically addressing AI and human rights. The Convention is open to Council of Europe member states and non-member states, with initial signatories including the European Union, United States, United Kingdom, and Japan. The Convention's purpose is ensuring that activities within AI system lifecycles are "fully consistent with human rights, democracy and the rule of law."[19] It establishes principles including respect for human dignity, individual autonomy, equality and non-discrimination, transparency, accountability, and reliability. Parties must conduct risk and impact assessments regarding AI systems' actual and potential impacts on human rights, democracy, and the rule of law. Significantly, the Convention applies to both public authorities and private actors, though it offers parties flexibility in regulating the private sector. Parties may opt to apply Convention obligations directly to private actors or implement alternative measures achieving equivalent rights protection. The Convention excludes activities related to national security interests and national defense matters,

---

[15] Article 14, Regulation (EU) 2024/1689 (EU AI Act).
[16] Chapter IV, Council of Europe Framework Convention on AI and Human Rights, Articles 14-15.
[17] Article 6, Regulation (EU) 2024/1689 establishes classification rules for high-risk AI systems.
[18] Article 1(1), Regulation (EU) 2024/1689 (EU AI Act).
[19] Article 1, Council of Europe Framework Convention on AI and Human Rights.

though these exclusions have attracted criticism from human rights organizations concerned about creating accountability gaps.

National Regulatory Approaches

Beyond regional frameworks, individual nations are developing diverse approaches to AI governance. India, while lacking comprehensive AI-specific legislation, has taken initial steps through data protection law and sectoral guidelines. The Digital Personal Data Protection Act, 2023, establishes consent-based data processing requirements applicable to AI systems that process personal data. However, the Act does not address critical AI-specific issues such as algorithmic bias, automated decision-making transparency, or AI system accountability. The United States has pursued a more fragmented approach, relying on sector-specific regulation, agency guidance, and executive action rather than comprehensive federal legislation. The White House issued an Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence in October 2023, establishing principles and directives for federal agencies. However, the absence of comprehensive federal AI legislation creates regulatory uncertainty and potential gaps in rights protection. China has implemented multiple AI-specific regulations including rules on algorithmic recommendations (2022) and generative AI services (2023), emphasizing content control, data security, and social stability alongside some consumer protection measures. The Chinese approach reflects distinct governance priorities emphasizing state control and social management alongside innovation promotion.

8. Critical Challenges in AI and Human Rights

Despite emerging regulatory frameworks, significant challenges persist in ensuring AI technologies respect and promote human rights.
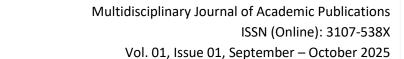
The Problem of Bias and Discrimination

AI systems can perpetuate and amplify existing societal biases through multiple pathways. Training data reflecting historical discrimination can lead AI systems to learn and reproduce discriminatory patterns. Algorithmic design choices, feature selection, and optimization objectives can encode bias. Discriminatory deployment contexts can result in disparate impacts even from neutral algorithms. Addressing AI bias requires technical, legal, and organizational interventions. Technical approaches include diverse training data, fairness-aware algorithms, and bias testing. Legal frameworks increasingly mandate bias mitigation, with the EU AI Act requiring high-risk AI systems to use representative, error-free training data and undergo conformity assessments. However, technical solutions face inherent limitations: different fairness definitions can be mathematically incompatible, and eliminating bias entirely may be impossible in systems learning from biased social realities.

Opacity and Explainability

The complexity of modern AI systems, particularly deep learning models, creates fundamental challenges for transparency and explainability. When AI systems make decisions through processes opaque even to their developers, effective accountability becomes difficult. The "right to explanation" implicit in human rights frameworks and explicit in some regulations faces practical limitations when dealing with complex AI systems. This opacity problem implicates multiple human rights. Fair trial rights require that individuals understand the basis of decisions affecting them. Equality rights require ability to identify and challenge discriminatory treatment. Privacy rights include understanding how personal data is processed. Yet current technical capabilities for explaining complex AI decisions remain limited, creating tension between what legal frameworks demand and what technology can deliver.

Power Concentration and Digital Divide

AI development and deployment capabilities are concentrated among a limited number of corporations and countries, raising concerns about power imbalances and equitable access to AI benefits. This concentration affects rights protection in multiple ways. Dominant AI providers can effectively set standards through market power. Regulatory capture risks emerge when regulators lack technical expertise or resources to effectively oversee sophisticated AI systems. The digital divide means AI benefits accrue primarily to wealthy nations and privileged populations, while risks disproportionately affect vulnerable groups. The UNESCO Recommendation explicitly addresses these concerns, emphasizing that AI development should be inclusive, non-discriminatory, and promote cultural diversity. However, translating these principles into effective policies that redistribute AI capabilities and benefits remains an ongoing challenge.
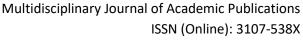
8. Balancing Innovation and Rights Protection

A central tension in AI governance involves balancing innovation promotion with rights protection. Overly restrictive regulation might stifle beneficial innovation, while insufficient regulation might enable rights violations. Different jurisdictions strike this balance differently, reflecting distinct regulatory philosophies and priorities. The EU approach emphasizes precaution and rights protection, establishing strict requirements before AI systems are deployed, particularly in high-risk contexts. Critics argue this approach might disadvantage European AI developers relative to competitors in more permissive regulatory environments. Proponents counter that rights-respecting AI builds trust and sustainable markets. Alternative approaches emphasize innovation and market self-regulation, intervening primarily when harm occurs. This approach may enable rapid innovation but risks allowing rights violations before regulatory response. The optimal balance likely lies between these extremes, with regulation calibrated to risk levels and evolving as AI capabilities advance.

9. Conclusion

The intersection of artificial intelligence, ethics, and human rights represents one of the defining challenges of contemporary law and governance. As AI systems increasingly mediate critical aspects of human life, ensuring these technologies respect fundamental rights and ethical values is imperative. The emergence of comprehensive legal frameworks including the EU AI Act, Council of Europe Framework Convention, and UNESCO Recommendation demonstrates growing international consensus that AI governance must be anchored in human rights principles. However, significant challenges persist. Algorithmic bias and discrimination threaten equality rights. AI opacity undermines transparency and accountability. Power concentration raises concerns about equitable access and democratic control. Balancing innovation and rights protection requires careful calibration of regulatory approaches. Effective responses require multiple strategies operating at different levels. International instruments establish shared principles and facilitate cooperation. Regional regulations provide enforceable requirements tailored to specific contexts. National laws address local priorities and conditions. Ethical frameworks complement legal regulation by articulating aspirational values and guiding voluntary practices. The foundational insight guiding AI governance must be that technology serves humanity, not the reverse. Human dignity, autonomy, equality, and freedom constitute non-negotiable values that AI systems must respect. Legal frameworks must evolve to address AI-specific challenges while remaining grounded in timeless human rights principles established in instruments like the Universal Declaration of Human Rights. As AI capabilities continue advancing, governance frameworks must remain adaptive, ensuring emerging technologies enhance rather than undermine human rights and democratic values. The path forward requires sustained commitment from governments, international organizations, civil society, academia, and industry to construct AI governance architecture that is effective, rights-respecting, and conducive to innovation that benefits all humanity. Only through such comprehensive, principled, and collaborative approaches can societies harness AI's transformative potential while safeguarding fundamental human rights for present and future generations.

10. Bibliography

1. Universal Declaration of Human Rights, G.A. Res. 217A (III), U.N. Doc A/810 at 71 (1948)

2. International Covenant on Civil and Political Rights, 999 U.N.T.S. 171 (1966)
3. Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, Council of Europe Treaty Series No. 225 (2024)
4. Regulation (EU) 2024/1689 (Artificial Intelligence Act), Official Journal of the European Union (2024)
5. Regulation (EU) 2016/679 (General Data Protection Regulation), Official Journal of the European Union (2016)
6. Digital Personal Data Protection Act, 2023, No. 22 of 2023, India
7. Constitution of India, Article 21
8. UNESCO, Recommendation on the Ethics of Artificial Intelligence (2021)
9. Justice K.S. Puttaswamy (Retd.) v. Union of India, (2017) 10 SCC 1