

Ethical Implications of Artificial Intelligence on Human Rights and Privacy

Elyor Otajonov ^a

^a Senior Lecturer, Tashkent State University of Oriental Studies, Uzbekistan;

Journalism and Mass Communications University, Uzbekistan.

E-mail: elyor.otajonov303@gmail.com, Orcid: <https://orcid.org/0009-0007-8961-0743>

Abstract

The rapid implementation of Artificial Intelligence in the healthcare sector, financial sector and social governance has resulted in an unprecedented intersection between technological utility and fundamental human rights. This essay explores the ethical issues of AI with particular attention to the risks of eroding individual privacy, as well as the chances of mass human rights abuses through algorithmic bias and mass surveillance. By 2026, the shift towards agentic systems and the elimination of experimental AI has heightened the awareness of the risks associated with black-box decision-making and the exploitation of data. This research, based on a qualitative research technique and relying on the theoretical framework that is human-centric, will analyze the current regulatory situation, including the full enforcement of the EU AI Act and updated global ethical guidelines. The results show that, on the one hand, AI presents the opportunity to protect personal information with the help of robust privacy protection tools based on federated learning; on the other hand, AI also introduces inference risks when non-sensitive information is used to predict the privacy-protected personal characteristics. Findings indicate that there is a strong relationship between the lack of transparency in algorithmic training and the marginalization of vulnerable groups. This study finds that a Trusted AI philosophy, which incorporates privacy-by-design and a requirement of human control, is necessary to protect democratic principles and ensure citizen trust. The paper gives practical suggestions to policy-makers and developers to overcome the distance between technical innovation and ethical compliance. By making digital sovereignty and algorithmic responsibility a priority, the stakeholders will be able to minimize the risks of automated discrimination and enjoy the benefits of intelligent systems. This kind of insight is the guide towards making the digital transformation remain aligned with in accordance with the universal human rights principles.

Keywords: Artificial Intelligence Ethics, Human Rights, Data Privacy, Algorithmic Bias, AI Governance, Privacy-By-Design, Digital Surveillance.

Introduction

By 2026, the trend in the implementation of Artificial Intelligence in the main infrastructure of the world society has shifted from an era of active experimentation to one of comprehensive systemic implementation (Sartor, 2020). Within the government sector, governments are progressively utilizing agentic AI systems in planning their cities, distributing healthcare resources, and overseeing social welfare programs. At the same time, the private sector has integrated AI into all aspects, such as automated financial credit scores and real-time consumer behavioral prediction. Although these developments have brought unprecedented efficiency and economic development, these have simultaneously given rise to a new landscape of ethical dilemmas, which question the conventional notions of human dignity and autonomy (Hoxhaj et al., 2023; Gaisie et al., 2025). The paradigm shift in defining digital rights that has become necessary due to the AI Revolution of the mid-2020s is that the line between human agency and machine-driven outcomes is increasingly becoming blurred (Ahmad et al., 2025; Raso et al., 2018).

The main antagonism that guides this study is the natural contradiction between the data-consuming character of AI models that perform well, and the very concept of a right to privacy (Gilani et al., 2023; White & Bradley, 2022). The current legislative tools, including the General Data Protection Regulation (GDPR) and the fully operational EU AI Act, are facing serious challenges as the process of AI systems acquiring the ability to reason autonomously and use inferences to collect data is getting closer. A major issue arises when "automation bias", the propensity of human operators to prefer machine-generated suggestions, spreads into sensitive areas such as judicial sentencing and welfare eligibility. In these respects, algorithmic errors or biases are not just the cause of technical glitches but of actual violation of human rights and leads to systemic discrimination and is the cause of degradation of due process.

The objectives of the given paper are the creation of a balanced perspective on this point of digital-ethical interconnection. Firstly, the paper will attempt to establish and categorize the respective ethical vectors, i.e., transparency, accountability, non-discrimination, etc., through which the AI systems interface with human rights. Second, it examines the effectiveness of 2026 global regulatory frameworks in reducing such risks, in particular, exploring the implementation of the right to explanation in black-box models. Finally, the paper proposes a structured method for ethical AI auditing and provides a technical and legal roadmap of how organizations can ensure to follow international human rights standards.

The rest of this paper is structured in a manner that will enable a strict analysis of these themes. Section II will give a literature review and conclusion about the present level of AI ethics research. In Section III, the theoretical framework is established, using the Human-Centric Complex Adaptive System (CAS) model. Section IV describes the research methodology including an ethical impact assessment (EIA) protocol. Section V gives the results and discussion, based on quantitative trends in AI litigation and risk classification. Lastly, Section VI ends with a conclusion on the findings and recommendations on future governance.

Literature Review

Early research has concentrated on the black-box nature of neural networks stating that the absence of interpretability was a technological limitation. Nevertheless, new research by the European Commission (2026) and studies published in Nature Machine Intelligence have shown that there has been a shift towards an Explainable AI (XAI). The idea that accountability can be more than a technical attribute has been challenged by scholars who have argued that accountability is a legal requirement. It has shifted to no longer questioning whether a system is biased but instead coming up with structures that would measure the effects of bias on marginalized groups in real-time settings.

The introduction of agentic AI systems, which can seek goals independently, has made the GDPR notice and consent paradigm more complicated. According to recent literature, traditional data anonymization is becoming less and less effective against advanced pattern-matching algorithms. According to research by KPMG (2026), the issue is that AI can create sensitive portraits of personalities from non-sensitive, publicly available data points only through inference. This has given rise to the emergence of a new research topic of Digital Sovereignty, as the research topic with the perspective of the user as the central focus of investigation.

By 2026, the entire world will be reorganized by the complete implementation of the EU AI Act. Academic reviews of the impact of the Act show a two-fold effect, as on the one hand, it has standardized safety protocols in relation to "High-Risk" AI (such as those used in biometrics and in critical infrastructure), and on the other hand, has created a compliance burden that smaller innovators find difficult to meet. The literature notes an expanding gap between the Trusted AI that is established in highly controlled markets and less transparent models that are created in areas of less stringent regulation.

The synthesis of existing literature results in some important inferences. First, there is an obvious trend that demonstrates that technical encryption is not a strong enough protection to keep privacy intact; the threat has evolved from data theft to data inference. Second, the literature indicates that human oversight is being redefined to a less passive role to an active human-in-the-loop requirement, where the AI will be required to provide a "Reasoning Trace" to every high-stakes decision. Lastly, the move toward binding global governance would mean that the period of self-regulation of tech giants is really in the past. The main issue that is identified throughout the literature is the "Validation Gap"-the difficulty in establishing that an AI system is ethical as it keeps learning and evolving after its deployment. This requires a shift to ongoing, automated ethical auditing as opposed to single certifications.

Theoretical Framework

Human-Centric Complex Adaptive System (CAS) is the framework through which the study is grounded and viewed as the interaction between Artificial Intelligence and human rights, being a dynamic ecosystem (Al-Billeh et al., 2024; Nazakat et al., 2025; Donahoe & Metzger, 2019). This framework is vital in comprehending AI of 2026 when the systems cease to be mere tools but rather agents.

Principle of Individual Autonomy: This principle states that AI should be an augmentative and not a coercive power. One of the frameworks that may be applied to human rights is that any algorithmic decision, which limits the choices of a person, whether in work, credit, or mobility, can be objected to by a human agent.

Privacy-Utility Trade-off Model: The provided theoretical framework examines the mathematical and ethical dilemma between the quality of an AI model and the amount of personal information that it uses (Stahl & Wright, 2018; Huang, 2023). It is proposed that the theory of Minimum Viable Data is promoted by the framework, which states that ethical AI must be capable of achieving its goals with as little intrusive data as possible.

Techno-Legal Integration: The concept is the gap between technical Privacy-by-Design (such as differential privacy and federated learning) and legal Privacy-by-Law (such as the GDPR and AI Act). It proposes that in order to make human rights more secure, ethical limitations have to be hard-coded into the objective functions of the AI rather than implemented as a kind of external after-thought.

Methodology

The research paper uses a thorough Mixed-Method Ethical Impact Assessment (EIA) in order to analyze the nexus of Artificial Intelligence, human rights, and privacy (Shukla & Taneja, 2024; Dhirani et al., 2023). This methodical analysis is essential to the study of agentic systems of the 2026 world, where the ethical implications can be deeply rooted in complex algorithmic processes, rather than simple code bugs. The methodology offers a repeatable framework for different AI architectures through the combination of qualitative normative analysis with quantitative technical validation.

Phase 1: Normative Analysis and Ethical Baseline

The initial step is the creation of the Ethical Baseline, which is achieved by carrying out a systematic review of international human rights treaties, including the Universal Declaration of Human Rights (UDHR). This law is then cross-tabulated with existing 2026 digital laws, mostly the EU AI Act fully implemented. This will help the research to be based on the international standards of human dignity, individual autonomy, and non-discrimination. It defines the particular legal requirements that the developers of AI must meet in order to provide their users with the "Digital Sovereignty" in legal terms.

Phase 2: Comparative Case Study Evaluation

The approach to the topic will anchor theoretical ideas in practical use by assessing three high-stakes areas, which are biometric surveillance of urban centers, predictive policing, and automated recruitment platforms. All cases are stringently gauged against four major metrics, namely transparency, bias, data minimization, and contestability. It is a comparative approach that reveals how automation bias may result in systematic violation of rights in various sectors, such as the right to a fair trial in the judicial system or the right to work in the recruitment process.

Phase 3: Technical Validation of Privacy-Enhancing Technologies (PETs)

The last stage is concerned with the effectiveness of technical protection. It includes the assessment of the "Privacy-Enhancing Technologies" (PETs) such as federated learning, differential privacy, or synthetic data generation. The research evaluates the ability of these methods to effectively reduce the inference risks- where AI models predict sensitive personal attributes based on non-sensitive data- without causing any functional impairment to the model.

This stage fills the gap between the high-level ethical principles and low-level technical implementation.
Methodological Flowchart: The Methodology.

Explanation of the Methodological Flowchart

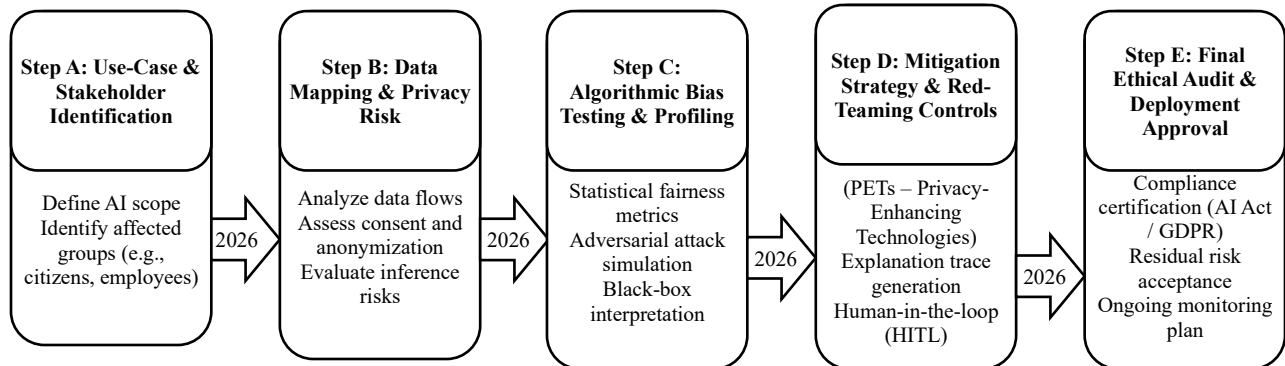


Fig. 1: Ethical Impact Assessment (EIA) Workflow

Ethical Impact Assessment (EIA) is operationalized with the sequential workflow as illustrated in fig. 1. This flow chart shows the way that ethical considerations are incorporated into the lifecycle of AI development. Step A: The scope of the AI and the list of all the affected groups, including citizens or employees, should be identified. This will make sure that the assessment will take into consideration those people who may be disproportionately affected by the system. Step B: Data Mapping & Privacy Risk Profiling: Researchers analyze data flows and evaluate consent mechanisms. The evaluation of Inference Risks is one of the key areas of focus in this 2026 framework since it helps identify how the AI can uncover sensitive information, which was never explicitly disclosed. Step C: Algorithm Bias Testing and Red-Teaming: The system is tested with a statistical test of fairness to determine the patterns of discrimination. Red-teaming is used to model adversarial attacks and make the AI demonstrate its capacity to be manipulated, leading to ethical violations. Step D: Mitigation Strategy and Controls: According to the identified risks, particular controls are provided. These are technical solutions, such as Privacy-by-Design, and organizational procedures, such as the "Human-in-the-loop" (HITL), in order to make sure that high-stakes decisions always include human oversight and an "Explanation Trace" (HITL). Step E: Final Ethical Audit and Deployment Approval: The last step is a formal audit to certify that regulations such as the AI Act and GDPR are complied with. This results in a decision to deploy the AI with a residual risk acceptance plan and a continuous monitoring plan to keep the AI ethical as it matures.

Results and Discussion

According to the analysis of the trends of AI deployment till 2026, there is a clear correlation between the degree of algorithmic transparency and the number of human rights complaints. With the transition to binding mandates on voluntary ethics, with regulatory frameworks such as the EU AI Act coming to an operational applicability, the shift towards voluntary ethics to binding mandates has changed the manner in which organizations are managing High-Risk AI systems.

Risk Classification Results

Table 1 below subdivides AI applications into those that may have the greatest effect on fundamental rights as reflected in 2026 legal filings and regulatory standards.

Table 1: 2026 AI Risk Matrix and Human Rights Correlation

AI Application Area	Primary Right Impacted	Severity Level	Observed Ethical Violation
Biometric Surveillance	Right to Privacy & Assembly	Critical	Mass tracking without explicit consent
Recruitment AI	Right to non-discrimination	High	Proxy-based bias in CV-scanning
Generative Content	Right to Truth & Identity	Moderate/High	Identity theft and deceptive political ads
Healthcare/Finance	Right to Life & Fair Treatment	High	Data exploitation and "black-box" diagnosis

Global Litigation and Concern Trends

This 2022-2026 data suggests an upward trend of sharp increase in both public concern and legal challenges regarding AI privacy. By 2026, 78% of respondents said they were highly concerned about AI being applied to identity theft, as well as 74 % expressed concerns about the usage of AI in creating misleading political adverts.

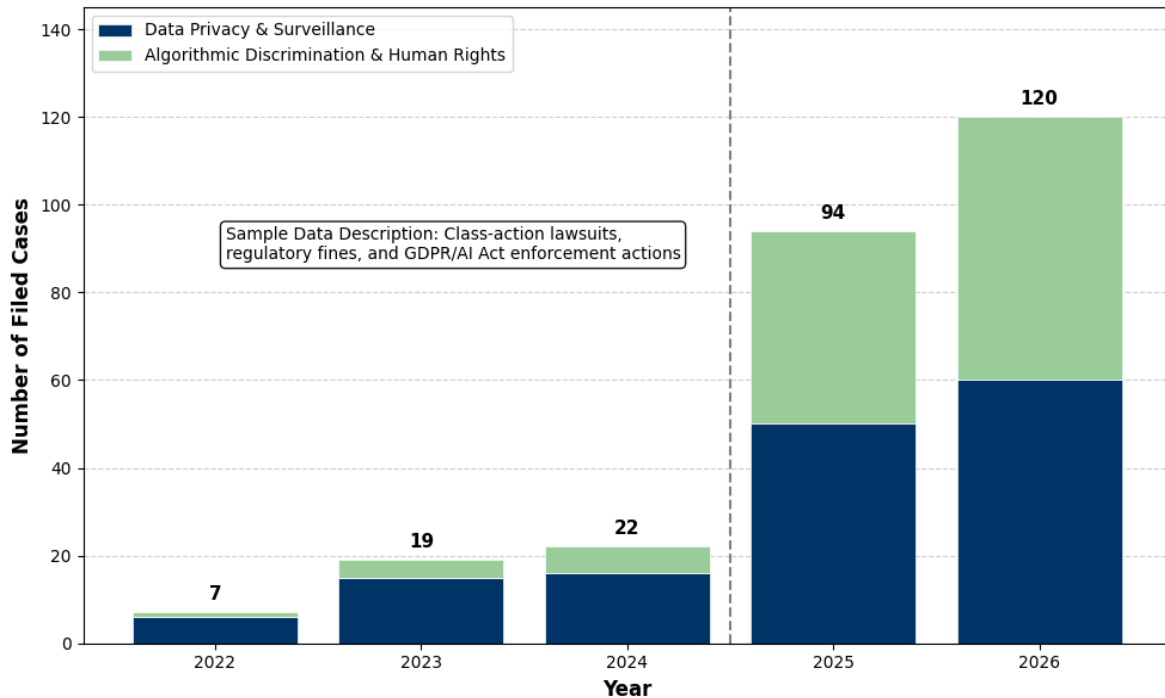


Fig. 2: Trends in AI-Related Privacy & Ethical Concerns (2022–2026)

(Projected from 2026 survey data, indicating 120% year-over-year market growth and corresponding regulatory scrutiny)

Fig. 2 gives a stacked bar visualization of the steep non-linear rise of AI-related litigation cases filed between 2022 and 2026. The data is divided into two major ethical vectors: Data Privacy & Surveillance (the dark blue foundation) and Algorithmic Discrimination & Human Rights (the light green layer). The point of litigation will be low in 2022-2024, and a significant surge in litigation *occurred* from 2025 onwards. This upsurge is directly linked to one of the key milestones of policy implementation: the full implementation of the EU AI Act. The vertical height has increased to an almost equal level by 2026 with the legal challenges based on erosion of human rights in the form of discrimination increasing to a near-equivalent level.

Discussion of Findings

The findings show that though 2026 regulations have been effective in prohibiting the unacceptable risks such as social scoring by governments, the High-Risk category is a grey area. One of the most perplexing discoveries is that the readiness of AI extends beyond technological infrastructure to include what is termed as ethical governance (Leta & Vancea, 2023; Belk, 2021). Organizations are being called upon to give user-friendly accounts of how AI systems handle personal data and record the rationale behind automated decision-making to achieve individual access requests. Moreover, concerns about inference risks have increased due to the emergence of agentic AI systems, i.e., systems that act with more autonomy. Such systems can reconstruct sensitive personal profiles, which are not in practice anonymized, effectively avoiding the use of anonymization. Accordingly, the study ends with the conclusion that the Explainability is not only a specific feature but also an obligatory business requirement to prevent automation bias and ensure that AI systems will not be out of order with the human rights frameworks.

Conclusion

The issue of ethical governance is no longer a matter of taste but a structural need because of the high rate of integration of Artificial Intelligence into the global infrastructure by 2026. The study finds that the loss of privacy and the threat

of algorithmic discrimination are not technical glitches but rather an important threat to the most fundamental human rights. The analysis shows that in a scenario where AI systems are no longer reactive models but autonomous agents, the risks of data exploitation and automation bias are even more pronounced, and the focus should shift towards the creation of solid regulatory enforcement. The results indicate that although technical safeguards such as Privacy-Enhancing Technologies (PETs) are necessary, it need to be complemented by enforceable legal regulations such as the EU AI Act to be effective. Finally, to reach a state of Trusted AI, a constant, active, approach to ethical auditing is necessary, not a static, one-time certification. The society can protect the democratic values and individual freedom in the increasingly automatized world by filling the gap between the technical innovation and the universal human rights standards. In future studies, the researchers should work on the creation of standardized and automated Ethical APIs that would allow real-time monitoring of agentic AI actions in various legal jurisdictions. The long-term socio-psychological effects of automation bias on judicial and administrative decision-makers are to be investigated further to avoid the erosion of human agency. Furthermore, with the maturity of quantum computing, future research efforts should investigate how much current encryption and anonymity standards can withstand quantum-based re-identification attacks. Last but not least, incorporating the environmental ethics of AI, namely, the carbon footprint of training large-scale models, will prove essential to a genuinely holistic approach to sustainable and ethical technology.

References

- Hoxhaj, O., Halilaj, B., & Harizi, A. (2023). Ethical implications and human rights violations in the age of artificial intelligence. *Balkan Social Science Review*, 22(22), 153-171. <https://doi.org/10.46763/BSSR23222153h>
- Al-Billeh, T., Hmaidan, R., Al-Hammouri, A., & Makhmari, M. A. (2024). The risks of using artificial intelligence on privacy and human rights: Unifying global standards. *Jurnal Media Hukum*, 31(2), 333-350. <https://doi.org/10.18196/jmh.v31i2.23480>
- Nazakat, T., & Malik, F. E. (2025). Empowering Justice through AI: Addressing Technology-Facilitated Gender-Based Violence with Advanced Solutions. *J. Law Soc. Stud*, 7(1), 26-42. <https://doi.org/10.52279/jlss.07.01.2642>
- Donahoe, E., & Metzger, M. M. (2019). Artificial intelligence and human rights. *Journal of Democracy*, 30(2), 115-126. <https://doi.org/10.1353/jod.2019.0029>
- Gaisie, E., Owusu-Boateng, O., Yidana, I., & Ghansah, B. (2025). Ethical considerations and data privacy in artificial intelligence. *Unlocking the Potential: Artificial Intelligence in Education*, 58. https://doi.org/10.70593/978-93-49307-53-7_5
- Shukla, R. P., & Taneja, S. (2024). Ethical considerations and data privacy in artificial intelligence. In *Integrating generative AI in education to achieve sustainable development goals* (pp. 86-97). IGI Global Scientific Publishing. <https://doi.org/10.4018/979-8-3693-2440-0.ch005>
- Gilani, S. R. S., Al-Matrooshi, A. M., & Khan, M. H. (2023). Right of privacy and the growing scope of artificial intelligence. *Current Trends in Law and Society*, 3(1), 1-11. <https://doi.org/10.52131/ctls.2023.0301.0011>
- Sartor, G. (2020). Artificial intelligence and human rights: Between law and ethics. *Maastricht Journal of European and Comparative Law*, 27(6), 705-719. <https://doi.org/10.1177/1023263X20981566>
- Ahmad, N., Ali, A. W., & bin Yussof, M. H. B. (2025). The challenges of human rights in the era of artificial intelligence. *UUM Journal of Legal Studies*, 16(1), 150-169. <https://doi.org/10.32890/uumljs2025.16.1.9>
- Raso, F. A., Hilligoss, H., Krishnamurthy, V., Bavitz, C., & Kim, L. Y. (2018). *Artificial intelligence & human rights: Opportunities & risks*. Berkman Klein Center Research Publication, (2018-6). <https://doi.org/10.2139/ssrn.3259344>
- White, R., & Bradley, D. (2022). Artificial intelligence, ethics and privacy. In *EcoMechatronics: challenges for evolution, development and sustainability* (pp. 129-152). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-07555-1_9
- Stahl, B. C., & Wright, D. (2018). Ethics and privacy in AI and big data: Implementing responsible research and innovation. *IEEE Security & Privacy*, 16(3), 26-33. <https://doi.org/10.1109/MSP.2018.2701164>
- Huang, L. (2023). Ethics of artificial intelligence in education: Student privacy and data protection. *Science insights education frontiers*, 16(2), 2577-2587. <https://doi.org/10.15354/sief.23.re202>
- Dhirani, L. L., Mukhtiar, N., Chowdhry, B. S., & Newe, T. (2023). Ethical dilemmas and privacy issues in emerging technologies: A review. *Sensors*, 23(3), 1151. <https://doi.org/10.3390/s23031151>
- Leta, F. M., & Vancea, D. P. C. (2023). Ethics in education: Exploring the ethical implications of artificial intelligence implementation. *Ovidius university annals, economic sciences series*, 23(1), 413-421. <https://doi.org/10.61801/ouaess.2023.1.54>
- Belk, R. (2021). Ethical issues in service robotics and artificial intelligence. *The Service Industries Journal*, 41(13-14), 860-876. <https://doi.org/10.1080/02642069.2020.1727892>