Content lists available at SRN Intellectual Resources

International Journal of Advances in Social Sciences and Humanities

Journal homepage: https://journal.srnintellectual.com/index.php/ijassh

Original Article

Artificial Intelligence and Human Rights in the Era of Digital Governance: Exploring Promises and Challenges

Rajagopal Parameswaran Kunjukrishnan [a,*] and Sreekala Krishnakumariamma [b]

[a] NSS Hindu College, Perunna, Changanassery, Kerala 686102, India.
[b] Government College, Attingal, Attingal, Kerala 695101, India; sreekalasreerangom@gmail.com (S.K.)
* Correspondence: pkrgopal@rediffmail.com (R.P.K.)

Citations: Kunjukrishnan, R.P., & Krishnakumariamma, S. (2025). Artificial Intelligence and Human Rights in the Era of Digital Governance: Exploring Promises and Challenges. *International Journal of Advances in Social Sciences and Humanities*, *4*(2), 64-71.

Received: 8 February 2025 Revised: 16 April 2025 Accepted: 7 May 2025 Published: 31 May 2025

Abstract: Artificial intelligence (AI) is rapidly transforming political, economic, and social systems, raising profound implications for human rights. While AI offers opportunities to enhance equality, access to justice, and public service delivery, it simultaneously poses risks, including mass surveillance, algorithmic bias, privacy violations, and gaps in accountability. Despite growing attention to these issues, existing research has largely focused on either the technological capabilities of AI or its ethical challenges, leaving a gap in comprehensive analyses that integrate AI's impact on human rights within governance frameworks. This study aims to examine the dual effects of AI on fundamental rights, exploring both its potential to promote social equity and its capacity to exacerbate vulnerabilities. Using a qualitative methodology, the research draws on case studies, policy analyses, and reviews of ethical and legal frameworks to evaluate AI's influence on healthcare, education, governance, and access to justice. The study also critically assesses regulatory and governance mechanisms designed to mitigate AI-related risks. The findings highlight that while AI can improve service delivery, empower marginalized communities, and foster more inclusive governance, unregulated deployment can deepen inequality, legitimize discrimination, and erode civil liberties. Ethical gaps, transparency deficits, and weak accountability structures continue to be key challenges. The study concludes that realizing AI's potential to advance human rights requires robust governance frameworks, international cooperation, ethical design principles, and ongoing monitoring. Policymakers, engineers, and human rights advocates must collaborate to ensure that AI serves as a tool for social progress rather than a source of harm, striking a balance between innovation and the protection of dignity, equity, and justice.

Keywords: Artificial Intelligence; Human Rights; Algorithmic Bias; Digital Ethics; Data Protection.

1. Introduction

Artificial intelligence (AI) can reshape our way of life utterly, but it is advancing on an almost exponential scale and in the hands of governments and business corporations that cynically disregard the possible impact of AI, oppose efforts at sensible regulation, and have no concern for regulating or directing its impact (Crawford, 2021; Zuboff, 2019; Pasquale, 2015). The research will examine the various ways in which artificial intelligence (AI) affects human rights, both positively and negatively (Floridi et al., 2018; Jobin, Ienca, & Vayena, 2019). It will also highlight how different

e-ISSN: 2948-4723 @ 2025 SRN Intellectual Resources
DOI: 10.56225/ijassh.v4i2.401

regulatory bodies are responsible for the misuse of AI and how they need to strive to implement safeguards or caps on AI development and usage (UNESCO, 2021; Bryson, 2022). Privacy and bias concerns have become essential, instinctive considerations (Buolamwini & Gebru, 2018; Noble, 2018). But solving privacy and discrimination problems is not the end goal, even though these have been achieved. They are, instead, milestones along the way to the ultimate goal, which is a comprehensive analysis of the effects of AI systems on human rights in general (Russell, 2019; Eubanks, 2018). Artificial intelligence (AI) is revolutionizing the world at an unprecedented rate, affecting almost every aspect of human life (Topol, 2019; Crawford, 2021).

AI has the potential to revolutionize healthcare, education, governance, and security (Topol, 2019; Holmes, Bialik, & Fadel, 2022; Ashley, 2017). While AI has the potential to enhance human welfare, it also presents real challenges that threaten basic human rights. Algorithmic bias, mass surveillance, destruction of privacy, and lack of responsibility all raise serious ethical as well as legal issues (Pasquale, 2015; Buolamwini & Gebru, 2018; Zuboff, 2019). At the interface of AI and human rights is a challenging terrain where innovation and ethics converge. On one hand, AI solutions have the capacity to enhance access to justice, empower marginalized groups, and foster equitable development (Floridi et al., 2018; Eubanks, 2018). On the other hand, AI solutions can improve access to justice, empower underprivileged groups, and foster equitable development. Meanwhile, unregulated deployment of AI technology could exacerbate inequality, legitimize institutional discrimination, and compromise freedoms (Noble, 2018; Crawford, 2021).

The rapid development of artificial intelligence (AI) has transformed practically every aspect of civilization, from public administration and criminal justice to medicine and education (Meijer, Wessels, & van der Voort, 2021; Russell, 2019). Digital technologies are being increasingly used in public decision-making as their accessibility, equity, and inclusiveness continue to expand. Although not all the potential advantages of progress have been equally distributed, AI significantly enhances the digital governance system, and the advanced AI-driven digital governance system offers most of the opportunities. Technology advancements present both new risks and opportunities to guarantee the advancement and defense of universal human rights in an era where algorithms may impact the exercise of personal freedom, privacy, equal protection, and due process rights (Jobin, Ienca, & Vayena, 2019; UNESCO, 2021). We should approach the integration of AI into our government in a thoughtful and responsible manner. We must ensure that ethical issues and human rights are at the forefront of the development and application of these technologies, rather than something that is reluctantly accepted. Together, legislators and technologists must create frameworks that safeguard against prejudice and discrimination, promoting accountability and transparency in the decision-making process (Bryson, 2022; United Nations, 2023).

The current study focuses on the intersection of digital governance (DG), human rights (HR), and artificial intelligence (AI), as well as the implications and risks of AI for both development and democracy. The study examines the emergence of literature on the subject, evaluates the extent to which legal and ethical frameworks provide adequate safeguards, and suggests potential paths toward a more inclusive, rights-based approach to governing AI technology. The current study focuses on the intersection of digital governance (DG), human rights (HR), and artificial intelligence (AI), as well as the implications and risks of AI for both development and democracy. The study examines the legal and ethical frameworks that provide adequate safeguards and suggests potential paths toward a more inclusive, rights-based approach to governing AI technology. Through an analysis of case studies and regulatory models, the argument emphasizes the importance of ethical governance and human-centric design to ensure that AI development aligns with universal human rights principles. In an era of rapid technological advancement, understanding the implications of AI for human rights is more than an intellectual exercise; it is a global imperative.

Artificial intelligence (AI) refers to the simulation of human intellectual processes by computers, mainly computer systems. These processes include learning (gathering data and inferring patterns), reasoning (applying rules to reach approximate or certain conclusions), and self-correction (improving over time based on feedback). AI encompasses various subdomains, including machine learning (ML), natural language processing, computer vision, and robotics, each of which imparts unique abilities for solving complex problems. Functions at the intersection of technological advancement and social change; thus, conceptual knowledge is central while considering its implications for human rights. In this article, AI is analyzed within a multidimensional framework that encompasses technological, ethical, sociological, and legal aspects. By situating AI within this conceptual framework, the article enables a critical evaluation of its contribution to human rights. This is not only useful for understanding what AI is but also provides the foundation for analyzing its advantages and limitations in advancing or diminishing rights in a rapidly digitizing world.

The swift integration of artificial intelligence (AI) into social arrangements has had significant implications for human rights (Crawford, 2021; Zuboff, 2019). Although AI technologies are hailed for their potential to speed up procedures and enhance decision-making, they also present significant issues such as privacy intrusions, biased decision-making, and the deterioration of individual rights (Pasquale, 2015; Buolamwini & Gebru, 2018; Noble, 2018). The lack of robust legal structures and moral principles exacerbates these problems, resulting in a disconnection between technological advances and governance (Jobin, Ienca, & Vayena, 2019; UNESCO, 2021). In addition, marginalized communities often suffer disproportionately from the negative effects of AI, worsening the disparities

(Eubanks, 2018; Crawford, 2021). The dual nature of AI's power to do good and evil demands intense research into how it shapes the destiny of human rights. Solving these issues is crucial to ensure that AI development adheres to the principles of dignity, equity, and justice for all (Floridi et al., 2018; Russell, 2019).

This paper examines the interplay between artificial intelligence (AI) and human rights, specifically the advantages and disadvantages AI poses in promoting and safeguarding basic rights. The effects of AI on privacy, equality, freedom of expression, and access to justice are all significant issues that warrant exploration (Ashley, 2017; Topol, 2019). The research also examines the ethical and legal implications of AI devices, particularly for marginalized and vulnerable groups (Bryson, 2022; Rajagopal et al., 2024). This study also evaluates existing regulatory regimes and governance frameworks to ascertain their efficacy in addressing AI-related human rights issues (UNESCO, 2021; United Nations, 2023). Through the examination of these factors, this essay aims to contribute to the broader conversation on creating a well-balanced approach to AI innovation, one that respects human dignity and ensures accountability (Rajagopal et al., 2024).

2. Materials and Methods

In this multidisciplinary study, the connections between digital governance, human rights, and artificial intelligence (AI) are explored qualitatively at the global level through the research process. To investigate the issues of AI and human rights, this study conducted an extensive literature review from 2010 to 2024, drawing on academic sources, including books, policy papers, UN documents, peer-reviewed journal publications, and legal texts. Reputable sources, including the UN, UNESCO, the European Union, and several human rights organizations, provided the data under investigation. It should be highlighted that because this study is theoretically based (and uses secondary data), there is no empirical fieldwork or technical evaluation of the AI systems found in the data. Given conceivable and/or actual limitations in language or access to specific policy paper material, the literature reviewed seeks to incorporate as many global perspectives as possible. However, this study employs a qualitative research approach that integrates document analysis and thematic inquiry to enhance understanding of the intersection between artificial intelligence (AI) and human rights. An initial framework was established through an examination of primary sources, including international human rights treaties, AI ethical codes, and legislative reports. Secondary sources, including peer-reviewed journals, policy briefs, and case studies, offer insight into the practical applications of artificial intelligence technologies in the context of human rights. A comparative analysis was conducted to assess the effectiveness of existing regulatory systems in various nations. This approach enables a complete understanding of the advantages and dangers of AI.

3. Results and Discussion

3.1. Artificial Intelligence Origin and Development

Human beings are intelligent entities capable of doing complex and innovative jobs. They have always sought faster, easier, more efficient, and more convenient ways to do their work. As a result, they search for novel and innovative methods and technologies. John McCarthy, one of the pioneers of AI, was the first to define the term in 1955. The aim of AI is to build machines that behave intelligently (Bakiner, 2023). The Encyclopedia Britannica defines AI as the ability of digital computers to solve issues that are typically related to the more advanced intellectual processing abilities of humans (Ertel, 2017). AI is primarily concerned with automating intelligent behavior. An example of a technical discipline that is transforming human civilization into a society of robots and machines is AI. AI encompasses a variety of technologies, such as algorithms, big data analytics, natural language processing, and ML. AI, by encouraging discrimination, has the potential to significantly impact human rights, both positively and negatively. To avoid potential harms and optimize the benefits of AI, human rights impact assessments built specifically for the technology are required.

Artificial Intelligence (AI) opened a new vista for human progress despite concerns and prospects. AI assists people in various ways. AI has the potential to be applied in various ways, such as assisting senior individuals in leading more fulfilling and enjoyable lives, as well as aiding in the diagnosis and treatment of patients. They were particularly motivated by AI's participation in comprehensive public health initiatives that rely on vast amounts of data that can be collected in the future, including personal genomes and dietary information. Several experts argued that AI would encourage long-awaited enhancements in education systems (Anderson & Rainie, 2018). AI has also made substantial contributions to the field of medicine, with applications that encompass drug discovery, clinical trials, diagnosis, and therapy. Doctors and researchers can benefit from the utilization of AI-powered systems to analyze patients' data, determine potential health hazards, and develop personalized treatment regimens. This can enhance patient health outcomes and accelerate the development of novel medical technologies and therapies (Duggal, 2023). Human intelligence and AI complement each other. AI can manage repetitive tasks and produce data-driven results. The basic purpose of AI is to support human beings in accomplishing tasks that need intelligence. Most individuals are expected to benefit from the expansion of AI in the coming decade, according to experts. However, there is considerable

apprehension regarding the potential implications of AI advancements on fundamental concepts of human identity, productivity, and autonomy.

Despite the numerous benefits of AI, it also has its drawbacks. Some of the benefits of AI include the development of autonomous vehicles, assistance with medical diagnostics, data analysis to facilitate informed decisions, and increased efficiency through task automation. Displacement of jobs, ethical concerns regarding privacy and bias, security issues resulting from cyber-attacks, and a dearth of human-like creativity and empathy are some of the drawbacks of AI (Duggal, 2023). Currently, AI is contributing to and exacerbating severe breaches of privacy and data protection, as the vast collection of personal information increases the likelihood of exploitation. AI has already been implemented to facilitate mass surveillance and to perpetuate bias in the criminal justice system, healthcare, education, employment market, access to housing, and banking. This has led to an exacerbation of discrimination against already marginalized groups in ways that are nearly imperceptible to the average individual.

3.2. AI and Human Rights: Opportunities and Challenges

AI is revolutionizing contemporary society with its transformational capabilities in areas such as healthcare, education, and economic growth. Although AI has been shown to enhance efficiency, accessibility, and innovation, it also raises significant concerns regarding its compatibility with fundamental human rights. Human rights are a helpless sidekick to the expansion of wealth inequality, says Samuel Moyn (Moyn, 2018). The connection between human rights and AI is a two-edged sword. On the one hand, AI offers considerable advantages in many fields, enhancing society's progress, welfare, and competitiveness by automating human work. On the other hand, it poses a severe threat to many human rights and basic freedoms when misapplied (Quintavalla & Temperman, 2023). Human rights, inherent in all individuals irrespective of nationality, gender, or other status, struggle to keep pace with the rapid changes in fast-developing technologies. As AI changes rapidly, the development of human rights law and jurisprudence can sometimes be slower, making the human rights framework inadequate at times as an ethical regulatory mechanism for AI. However, the current human rights framework can serve as a basis for establishing proper limits of AI activities.

Human rights of marginalized individuals and groups are disproportionately impacted by AI, producing a new type of technological dictatorship. As artificial intelligence becomes increasingly integrated into society, issues like racism and prejudice have surfaced as important concerns in political debates around technological growth. This concern is consistent with the principles outlined in Article 2 of the Universal Declaration of Human Rights and Article 2 of the International Covenant on Civil and Political Rights, which affirm everyone’s right to rights and freedoms without discrimination. Addressing these gaps is critical for ensuring that AI development and deployment promote fairness and inclusion rather than reinforcing systemic inequities. The right to privacy and freedom from discrimination is not the only fundamental right at risk from algorithmic systems. AI technologies also raise significant concerns about free expression and association. However, the impact of AI on these rights is often less immediately apparent compared to privacy and anti-discrimination issues, which are easier to conceptualize. This lack of visibility can obscure the broader implications of AI on human rights. Both the Universal Declaration of Human Rights and the EU Charter of Fundamental Rights emphasise the importance of freedom of opinion and expression. Similarly, in the US, the First Amendment to the Constitution protects these freedoms. These rights safeguard individuals' ability to think and express themselves freely, as well as their right to access and engage with diverse opinions and information. Ensuring these freedoms in the context of AI deployment and development is crucial to preserving democratic values and human dignity.

Human rights are universal and enforceable, and they are enshrined in an international body of law. Although governments are obligated to safeguard and enforce human rights, both corporations and government agencies share a responsibility to uphold these rights. National, international, and domestic agencies and organisations provide established systems for redress and explain the application of human rights law to changing conditions, especially developments in technology. Artificial intelligence technologies are transforming the way businesses and governments operate globally, accompanied by the risk of significant human rights violations. Data protection regulations, transparency, and accountability safeguards can help limit some of the worst-known uses, but more effort is needed to protect human rights as AI technology advances and expands into other sectors. The recent surge in AI applications has raised concerns about potential and actual threats, including discrimination, bias amplification, privacy violations, end-user manipulation, disinformation, and unaccountable decision-making. Since the mid-2010s, managing AI to limit or mitigate its harmful effects has been recognized as one of the most pressing concerns of our time, as evidenced by the numerous AI ethics guidelines. Furthermore, the data gathering and processing techniques encompassed by the term AI threaten to exacerbate other significant issues of social, political, economic, and legal nature, such as environmental degradation, structural racism, and social media content moderation, as well as sexism, among others (Bakıner, 2023).

AI can enhance human rights in various ways, from monitoring supply chain compliance to supporting medical advancements. However, because human rights are indivisible and interdependent, AI has potential to affect nearly every recognised human right, including right to freedom of expression, thought, assembly, association, and movement; right to privacy and data protection; rights to health, education, work, and an adequate standard of living; and non-

discrimination and equality (Human Rights & Artificial Intelligence, 2023). Under international human rights law, nations are required to protect the human rights of all people within their territory and jurisdiction from abuses caused by their own policies, practices, or conduct, as well as that of third parties, such as enterprises or individuals. Human rights considerations are relevant to all aspects of government practice and policy, including the policies and applications of AI employed by governments and other entities in society. New ethical frameworks for AI governance have been proposed by certain individuals.

In certain cases, there are attempts to bypass human rights frameworks altogether, particularly those aspects that governments find inconvenient. In other cases, efforts are made to exceed human rights norms by offering even greater protections. Importantly, nothing prevents governments from going beyond the existing human rights frameworks in their approach to AI governance. However, there are compelling reasons to utilize an established human rights framework as a foundation for AI law and administration. Human rights should serve as the cornerstone for all subsequent governance models for several reasons. First, the international human rights framework places the individual at the centre of governance, ensuring a people-focused approach. Second, it addresses the most critical societal challenges through established mechanisms. Third, it clearly defines the roles and responsibilities of both governments and the private sector, an essential feature for managing a technology like AI, which demands collaborative oversight. Finally, despite variations in its interpretation and implementation worldwide, no newly emerging ethical framework is likely to attain the level of geopolitical recognition and legal authority that the international human rights framework enjoys under international law.

A National Artificial Intelligence Strategy (NAS) is a document developed by the government that outlines a comprehensive, strategically oriented approach to AI, including key fields of concentration and specific actions. The primary goal of a NAS is to coordinate government initiatives to maximise AI's potential benefits for the economy and society while minimising associated risks. Since 2017, over twenty-five states and regional intergovernmental organisations have introduced such strategies, with several others in development. Globally, governments can take several measures to provide a framework for a human rights-centered approach to AI policy. The European Commission's Guidelines for Trustworthy AI address not only the technical soundness and safety of AI but also its sociopolitical impact on human rights. Salil Shetty, former Secretary-General of Amnesty International, emphasised this perspective in a 2017 address, stating, "There are huge possibilities and benefits to be gained from artificial intelligence if human rights are part of the core design and use of this technology" (Su, 2022). The government should guarantee that the NAS-creation process includes a critical component of experts in human rights and the impact of AI on human rights. In addition to generic human rights organizations, these should encompass a diverse array of civil society organizations and other stakeholders that represent communities that may be disproportionately benefited from specific uses of AI or may be negatively impacted by AI.

3.3. Key Human Rights Concerns with AI

Artificial intelligence (AI) has raised numerous human rights concerns as its applications become increasingly integrated into social, political, and economic systems. One major issue is the invasion of privacy, particularly using AI in data analytics, facial recognition, and mass surveillance. Governments and corporations can collect, analyze, and store vast amounts of personal data without sufficient consent or oversight, undermining individual autonomy and the right to privacy (Crawford, 2021; Zuboff, 2019; Eubanks, 2018). Such practices enable intrusive monitoring that threatens civil liberties and human freedom. Another critical challenge is the presence of discrimination and bias in AI systems. Algorithms trained on biased or incomplete datasets tend to reproduce and even amplify existing social inequalities, resulting in discriminatory outcomes in hiring, credit scoring, criminal justice, and healthcare (Buolamwini & Gebru, 2018; Noble, 2018; O'Neil, 2016). These biases can marginalize already vulnerable groups, reinforcing structural injustices. In addition, the use of AI in content moderation poses risks to freedom of expression. Automated moderation systems deployed by social media platforms often misclassify or remove legitimate content, inadvertently suppressing public discourse and cultural diversity (Gillespie, 2018; Balkin, 2020). This automated censorship threatens open communication, a cornerstone of democratic societies. The lack of accountability in AI decision-making processes further exacerbates these concerns. Due to the "black box" nature of many AI systems, it is often difficult to trace how and why specific decisions are made, making it nearly impossible to assign responsibility in cases of harm or human rights violations (Pasquale, 2015; Selbst & Barocas, 2018). Lastly, job displacement and economic inequality represent another dimension of human rights risk. The automation of labor-intensive and routine tasks through AI technologies can disproportionately affect low-income and less-skilled workers, leading to unemployment, wage suppression, and growing socioeconomic inequality (Frey & Osborne, 2017; Susskind, 2020). These shifts threaten the right to work and economic security, deepening existing disparities across societies.

3.4. Opportunities to Promote Human Rights with AI

Artificial intelligence (AI) presents substantial opportunities to advance and protect human rights when developed and applied responsibly. One of the most promising areas is enhancing access to justice. AI-driven legal tools, such as automated case management systems, online dispute resolution platforms, and virtual legal assistants, can provide affordable and accessible legal support to marginalized and vulnerable populations (Ashley, 2017; Remus & Levy, 2016). These innovations help reduce barriers to justice, especially in regions with limited legal infrastructure. In the field of healthcare, AI technologies contribute significantly to the realization of the right to health. Through advanced diagnostic tools, predictive analytics, and personalized medicine, AI can enhance early disease detection, refine treatment plans, and predict epidemic outbreaks, ultimately improving global health outcomes (Topol, 2019; Esteva et al., 2019). Similarly, in education, AI-powered platforms have the capacity to democratize learning by offering personalized instruction, adaptive learning pathways, and accessible educational resources for underserved communities (Holmes et al., 2022; Luckin et al., 2016). AI also strengthens governance and accountability by improving transparency, detecting corruption, and enhancing efficiency in public administration. For example, AI-driven data analytics can help governments allocate resources more equitably and monitor policy implementation (Bannister & Connolly, 2020; Meijer et al., 2021). Additionally, AI technologies play an essential role in monitoring human rights abuses. Machine learning algorithms can analyze satellite imagery, social media data, and digital evidence to identify patterns of violence, forced displacement, or environmental harm in real time, thereby facilitating timely intervention and advocacy (Altholz & Brown, 2020; Amnesty International, 2021).

To ensure that AI serves as a force for human dignity rather than harm, it is vital to adopt a human rights–centered approach to innovation. This includes developing and enforcing ethical AI guidelines that emphasize fairness, transparency, and accountability (Jobin et al., 2019; Floridi et al., 2018), as well as establishing legal frameworks that enhance data protection, implement anti-discrimination measures, and provide safeguards against misuse (UNESCO, 2021). Moreover, promoting inclusivity is crucial so that marginalized populations can benefit from AI advancements, bridging digital and social divides (Cihon et al., 2021). Ultimately, international cooperation among governments, civil society, and the private sector is crucial for addressing cross-border challenges and ensuring that AI development aligns with universal human rights values (Bryson, 2022; United Nations, 2023). By prioritizing ethical design, robust governance, and social inclusion, AI can be leveraged as a transformative instrument for justice, equality, and human dignity worldwide. As artificial intelligence (AI) becomes more prevalent in everyday life, regulation has become a pressing concern. The regulation of AI presents challenges due to the technology's rapid progress and global reach. Establishing efficient AI legislation and policies is critical to realising AI's benefits while limiting its hazards (RTS Labs, 2024). Establish a well-resourced, independent oversight agency to regulate and supervise the use of AI. Monitoring of open, ex-ante public procurement procedures and the outcomes of AI tool use should be part of the oversight mission. Oversight bodies require clear routes for hearing and resolving public complaints. (Kennedy, 2024).

4. Conclusions

Artificial intelligence has the transformative power to promote human rights by providing tools to advance justice, equality, and access to vital services. Nevertheless, this potential is overshadowed by serious dangers, systemic prejudices, privacy violations, and the stripping of people's agency, which undermine the rights it aims to safeguard. Navigating this complex landscape requires a proactive and multifaceted approach, one that prioritizes transparency, accountability, and inclusivity at every stage of AI development and deployment. By centring human rights within the frameworks guiding AI, stakeholders, governments, technologists, and civil society can harness its promise while mitigating its perils. Ultimately, the role of AI in shaping human rights is not predetermined; it is a choice made through collective action and ethical resolve. The future of human rights in the era of AI hinges on our capacity to balance innovation with responsibility; to ensure technology serves humanity and not erodes it.

Artificial intelligence has significant potential to enhance human rights by facilitating access to justice, reducing unfairness, and ensuring that everyone has equal access to essential services. Artificial intelligence has significant potential to enhance human rights by facilitating access to justice, reducing unfairness, and ensuring that everyone has equal access to essential services. These problems are not only technical failures; they are also failings of ethics and governance that need quick and concerted action. These problems need to be resolved with an emphasis on human rights. Policymakers must pay close attention to robust regulatory mechanisms that ensure accountability, transparency, and data protection. This is especially true for apps used in the public sphere. Developers and technologists must adhere to ethical design principles and conduct detailed impact analyses to avoid discriminatory outcomes. Increased public awareness, especially among marginalized communities, is crucial for tracking AI deployment and digital governance; human rights defenders and civil society organizations play a central role in this endeavor.

AI poses significant challenges to human rights, particularly for at-risk, vulnerable, and marginalised communities. Concrete efforts must be made to safeguard human rights in the age of AI. AI governance plans must incorporate clear goals, commitments, and actionable steps that aim to protect human rights. This includes integrating human rights

impact assessments into the development and deployment of AI systems. Governments should ensure these objectives are specific, measurable, achievable, relevant, and time-bound (SMART) to promote accountability and effectiveness. Governments should establish incentives or specific criteria to encourage the adoption of human rights-respecting practices across all sectors. These incentives could include tax benefits, grants, or public recognition for companies and organisations demonstrating a strong commitment to ethical AI practices. Establish accessible grievance and redressal mechanisms to address human rights violations arising from the use of AI. These mechanisms should include clear processes for reporting, investigating, and resolving issues, with remedies such as compensation, policy changes, or system improvements. Establish mechanisms for the continuous monitoring and adaptation of AI policies to address emerging risks and technological advancements. This ensures that protections remain effective in the face of rapid innovation. By integrating these elements into AI governance, governments can adopt a proactive and holistic approach to protecting human rights while promoting trust and accountability in AI systems.

Future work must prioritize building stronger international cooperation, consolidating AI policy with existing human rights standards, and creating inclusive forums where all stakeholders, particularly Global South voices, can contribute to the AI governance of the future. Finally, AI's influence on human rights will be a function of the decisions we make today. The way forward calls for a shared commitment to making technological innovation responsive to ethical foresight and basing it on the principles of dignity, equality, and justice. Provided those values hold true, AI can become a force to be reckoned with in furthering and eradicating our common human rights.

Author Contributions: Conceptualization, R.P.K. and S.K.; methodology, R.P.K. and S.K.; validation, R.P.K.; formal analysis, R.P.K. and S.K.; investigation, R.P.K. and S.K.; resources, R.P.K.; data curation, S.K.; writing—original draft preparation, R.P.K. and S.K.; writing—review and editing, R.P.K. and S.K.; project administration, R.P.K.; funding acquisition, R.P.K. Author has read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank NSS Hindu College and Government College, Attingal, India, for supporting this research and publication. The authors would also like to thank the reviewers for their constructive comments and suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

References

Anderson, J., Rainie, L., & Luchsinger, A. (2018). Artificial intelligence and the future of humans. Pew Research Center: Internet, Science & Tech. *Retrieved from https://www.pewresearch.org/internet/2018/12/10/artificial-intelligence-and-the-future-of-humans/*

Altholz, R., & Brown, D. (2020). Artificial intelligence and human rights fact-finding. *Human Rights Quarterly*, 42(3), 526–554.

Amnesty International. (2021). *Algorithmic accountability: Applying human rights standards to AI governance*.

Ashley, K. D. (2017). *Artificial intelligence and legal analytics: New tools for law practice in the digital age*. Cambridge University Press.

Balkin, J. M. (2020). *The free speech century: Technology, law, and the future of expression*. Harvard University Press.

Bakiner, O. (2023). The promises and challenges of addressing artificial intelligence with human rights. *Big Data & Society*, *10*(2), 20539517231205476.

Bannister, F., & Connolly, R. (2020). The future ain't what it used to be: Artificial intelligence and public sector ethics. *Government Information Quarterly*, 37(4), 101-410.

Bryson, J. (2022). The standard model of governance for AI. *Science and Engineering Ethics, 28*(2), 17. https://doi.org/10.1007/s11948-022-00381-2

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, 81, 1–15.

Cihon, P., Maas, M. M., & Kemp, L. (2021). Should artificial intelligence governance be centralised? *Science and Public Policy*, 48(5), 635–648.

Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.

Duggal, N. (2023). Advantages and Disadvantages of Artificial Intelligence [AI]. simplilearn. *August*, *24*, 2023.

Ertel, W. (2024). *Introduction to artificial intelligence*. Springer Nature.

Esteva, A., et al. (2019). A guide to deep learning in healthcare. *Nature Medicine*, 25(1), 24–29.

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines, 28*(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Floridi, L., et al. (2018). AI4People—An ethical framework for a good AI society. *Minds and Machines*, 28(4), 689–707.

Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, 114, 254–280.

Gillespie, T. (2018). *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press.

Holmes, W., Bialik, M., & Fadel, C. (2022). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson Education.

Meijer, A., Wessels, M., & van der Voort, H. (2021). Algorithmic governance and public values: Balancing transparency, accountability, and efficiency. *Public Administration Review, 81*(5), 941–951.

Moyn, S. (2018). *Not enough: Human rights in an unequal world*. Harvard University Press.

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.

Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Harvard University Press.

Quintavalla, A., & Temperman, J. (2023). *Artificial intelligence and human rights*. Oxford University Press.

Rajagopal, S., Popat, K., Meva, D., & Bajeja, S. (Eds.). (2024). *Advancements in Smart Computing and Information Security: Second International Conference, ASCIS 2023, Rajkot, India, December 7–9, 2023, Revised Selected Papers, Part III* (Vol. 2039). Springer Nature.

Remus, D., & Levy, F. (2016). Can robots be lawyers? Computers, lawyers, and the practice of law. *Georgetown Journal of Legal Ethics*, 30(3), 501–558.

Russell, S. (2019). *Human compatible: Artificial intelligence and the problem of control*. Viking.

Selbst, A. D., & Barocas, S. (2018). The intuitive appeal of explainable machines. *Fordham Law Review*, 87(3), 1085–1139.

Su, Anna. (2022). The Promise and Perils of International Human Rights Law for AI Governance. *Law, Technology and Humans, 4*(2), 166-182. https://doi.org/10.5204/lthj.2332

Susskind, D. (2020). *A world without work: Technology, automation, and how we should respond*. Metropolitan Books.

Topol, E. (2019). *Deep medicine: How artificial intelligence can make healthcare human*

UNESCO (2021). *Recommendation on the ethics of artificial intelligence*. United Nations Educational, Scientific and Cultural Organization.

United Nations. (2023). *Human rights, AI and digital technology: A global framework for cooperation*.

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.