

Human Rights and Artificial Intelligence: Addressing Emerging Threats to Humanity in Indian Context

Varsha Gaur¹, Prof. Naved Jamal²

¹Research Scholar, Department of Political Science, Jamia Millia Islamia

Email id: v.gaur1350@gmail.com

²Professor, Department of Political Science Jamia Millia Islamia

Email id: navedjamal@gmail.com

Abstract

The rapid convergence of artificial intelligence (AI), social media, and human rights has fundamentally transformed the global landscape, offered unprecedented opportunities while posed significant challenges. Human rights, inherent to all individuals regardless of nationality, gender, or socio-economic background, serve as universal safeguards against infringements on dignity and freedom. However, as technology reshapes how we interact, work, and communicate, ethical dilemmas surrounding privacy, equality, and freedom of expression have emerged. This article will underline the challenges and threats in respect to Human rights integration with Artificial intelligence. It will also underline the role of global institutions in digital age. Moving towards the conclusion this paper with highlight the importance of ethical codes as well as the way forward for India in addressing the challenges and threats in such emerging scenarios.

Keywords: Human Rights, Artificial Intelligence (AI), cyber security, Privacy, social media, Digital Age Technology.

INTRODUCTION

The fundamental tenet of human rights approaches to development is that states and the international community have a duty to guarantee that the economic, social, and cultural rights of developing nations' citizens, as well as their civil and political rights, are all upheld. Human rights are regarded as inherent to all people, regardless of our nationality, place of residence, sex, national or ethnic origin, colour, religion, language, or any other status. They are meant to be universal legal guarantees that shield individuals and groups from acts that infringe upon fundamental freedoms and human dignity. The intersection human relations and digital technology has grown in importance in today's quickly changing digital environment. This future that is both exciting and dangerous as technology continues to transform and model how we live and work. Our exploration of the worlds of artificial. Open-source innovation and artificial intelligence (AI) in cybersecurity have led us to a pivotal point: the nexus of AI and human rights. This new frontier forces us to think about not only what AI is capable of, but also what it ought to be able to achieve, and how its potential might be used to uphold rather than undermine our core liberties. Imagine a society in which artificial intelligence (AI) both protects and threatens to usurp our liberties. But as technology develops further, it also presents several moral and human rights issues that must be resolved. Technology and human rights

cross in a dynamic and complicated environment that needs to be carefully navigated and considered. (Muoma, 2024)

One important document in the history of human rights is the Universal Declaration of Human Rights (UDHR). The Declaration, which was drafted by representatives from all over the world with varying legal and cultural backgrounds, was adopted as a universal benchmark for success by all peoples and all countries on December 10, 1948, when the UN General Assembly met in Paris and adopted General Assembly resolution 217 A (III). For the first time, it calls for the universal protection of fundamental human rights. The UDHR has been translated into more than 500 languages since it was adopted in 1948, making it the most translated document in the world. It has also served as an inspiration for the constitutions of numerous newly independent states and emerging democracies. The UDHR, the International Covenant on Civil and Political Rights, the International Covenant on Economic, Social, and Cultural Rights, and its two Optional Protocols (on the death penalty and the complaints process). International Bill of Human Rights is made up of these rights and their optional protocol.¹ A key document in the international promotion and defence of human rights is the 1948 United Nations General Assembly ratification of the Universal Declaration of Human Rights (UDHR). Universal Declaration of Human Rights (UDHR) has had a major impact on how social, legal, and political institutions have developed around the world, including how social media platforms are regulated. It is based on the ideas of equality, freedom, and dignity. By providing a fundamental framework to safeguard individual liberties and advance democratic values, the UDHR has influenced attempts to resolve the difficulties and disputes related to social media regulation in the digital era. (Raj, 2024)

Recognizing each person's inherent rights and dignity is the main goal of the UDHR concepts over any communication channel, without regard to geographical limitations. The Universal Declaration of Human Rights (UDHR) expressly guarantees the right to freedom of expression in Article 19, which encompasses the freedom to actively seek, receive, and distribute information and concepts over any communication channel, without regard to geographical limitations. Because it emphasizes how crucial it is to preserve the openness, diversity, and accessibility of digital communication platforms, this principle has significant implications for social media management. With the ability to significantly affect every sphere of society, artificial intelligence is set to emerge as one of the most revolutionary forces in human history.

Digital Age and Human Rights interventions with Artificial Intelligence

The impact of technology on human rights is unprecedented in our modern culture, which is fuelled by technical innovation and digital literacy. The introduction of new technology has changed how we live, work, and communicate with one another, posing both opportunities and challenges. On the one hand, technology developments have improved communication, education, and information access, empowering people and communities globally. People can now use the internet as a potent tool for freedom of expression to hold governments responsible, campaign for social change, and share their voices. Beyond AI, our world is changing quickly due to technology breakthroughs. Automation and robotics, where robots do

¹ Universal Declaration of Human Rights | United Nations

progressively more difficult tasks with accuracy and independence, are influencing sectors like logistics, healthcare, and manufacturing.

Social media has profoundly changed how individuals connect, interact, and express themselves. For millions of individuals, Facebook, Instagram, YouTube, and Twitter are becoming essential components of their daily life. But social media's quick development has brought about significant changes to the field of free speech, with both advantages and disadvantages. The way that knowledge is shared and consumed has changed significantly because of social media's rise. Social media platforms offer almost unfettered access to everyone with an internet connection, in contrast to traditional media channels that are subject to editorial scrutiny and control systems.

Social media has also made it easier for online groups centred on common identities, interests, or viewpoints to emerge. These online spaces facilitate idea exchange, conversation, and social concern advocacy, giving voice to underrepresented groups and supporting neighbourhood organizations in organizing and enacting change. Divergent viewpoints have been amplified and worldwide social and political movements like the Arab Spring, Black Lives Matter, #Metoo, and others have been sparked in large part by social media. The development of social media has ultimately had a significant impact on the ability to express oneself, drastically changing how we interact with information, communicate, and make contact. Social media has empowered marginalized people and provided fair access to information, but it has also created difficulties in controlling online debate and maintaining uncensored discussions maintain the ideals of unfettered communication in the digital age while simultaneously preventing damaging speech.

The AI might monitor the creation and dissemination of narratives to sway public opinion in the political sphere. It could spot cases where political campaigns employ specific propaganda, disinformation, or deliberate disclosures to influence voters or undermine rivals. The AI may also be able to identify ways that foreign organizations try to sway elections or policy choices through clandestine means like disinformation campaigns, cyberattacks, or covert political movement sponsorship. A degree of transparency that is frequently absent in human society would be provided by the AI's capacity to identify these deceptive tactics. By shedding light on the unseen factors influencing political and economic outcomes, the AI would help make the public more aware and watchful, which could lessen the impact of similar manipulations in the future. (Ergest Alite)

Human Rights and Artificial Intelligence – Advantages

As demonstrated by the United Nations' use of AI in times of crisis, AI improves disaster response and resource optimization in humanitarian circumstances. Through technologies like blockchain, AI also improves accountability and transparency in supply chains and governance, thwarting corruption and guaranteeing adherence to moral labor standards. Even though AI has several advantages for promoting human rights, ethical supervision and responsible application are essential to guaranteeing that technology continues to be a positive force that empowers people and addresses structural injustices worldwide.

Through the analysis of enormous volumes of data, including social media activity and satellite imagery, AI also plays a crucial role in monitoring and stopping violations of human rights. AI is used by groups like Amnesty International to monitor atrocities and spur international action. Additionally, by spotting biased trends in systems, AI promotes equality by assisting legislators in addressing prejudice and guaranteeing equity in domains such as employment procedures. AI-powered educational platforms offer customized learning opportunities, closing gaps and facilitating underprivileged groups access to high-quality education.

One of the most revolutionary effects of artificial intelligence (AI) on the advancement of human rights is the promotion of freedom of expression. AI-powered platforms have transformed communication by giving a range of voices especially those from underrepresented or marginalized communities -a forum to express their opinions and have meaningful conversations. By allowing people to get over more conventional obstacles like geographic limitations, restricted media access, or sociopolitical silence, these platforms promote diversity in public discourse.

AI-powered social media algorithms are crucial in raising awareness of social inequalities and human rights abuses. AI can recognize and rank important issues by examining trends and patterns, guaranteeing that they are seen by people all around the world. For instance, the employment of AI tools to monitor and draw attention to instances of discriminatory behaviour, forced relocation, or police brutality has raised public awareness and sparked international campaigning. AI algorithms are used by social media sites like Facebook and Twitter to suggest posts, hashtags, or content that aligns with current movements in order to spur group activity.

AI also makes real-time translation and transcription possible, removing linguistic barriers and facilitating international discussions. In areas where local languages are infrequently included in the media, this has a particularly significant effect. Additionally, AI tools help journalists and human rights advocates by giving them the tools they need to more effectively record and report violations, even in circumstances that are restrictive.

One crucial area where artificial intelligence (AI) shows great promise to promote equality and fairness is in the fight against prejudice. Large datasets can be analysed by AI techniques to find latent bias and discriminating patterns that might otherwise go undetected. These prejudices are frequently present in fields including law enforcement, housing distribution, employment practices, and educational access. By seeing these trends, AI offers useful information that businesses and legislators may utilize to enact more equal systems and guarantee adherence to anti-discrimination regulations.

Emerging threats and Challenges of Artificial Intelligence and social media affecting Human Rights

Social media and artificial intelligence (AI) have helped to advance human rights and empower people, but they also bring new risks and difficulties that go against these very principles. Serious questions of equality, freedom of speech, privacy, and human welfare are brought up by the quick development of technology without sufficient regulation. (Channi Sachdeva)

Violations of Privacy

One of the major issues with artificial intelligence is data privacy because it involves sensitive information or histories. Social media sites and AI-powered algorithms gather and examine enormous volumes of user data, frequently without the users' express agreement. Because sensitive personal information can be used for targeted advertising, political manipulation, or spying by authoritarian regimes, this widespread data surveillance poses a challenge to the right to privacy. The risk of intrusive monitoring and control is increased when biometric technology, such as facial recognition, are misused, especially in repressive countries.

Algorithmic Bias

AI systems have the potential to reinforce and even worsen discrimination if they are trained on biased datasets. Vulnerable groups are disproportionately affected by algorithmic bias in fields such as hiring, credit rating, and law enforcement, which violates their right to equality. Like this, biased content may be amplified by social media algorithms, normalizing discrimination and hate speech. Algorithmic bias based on race, gender, or socioeconomic status can perpetuate societal inequalities. Data privacy violations raise concerns about surveillance and individual autonomy. Responsible development frameworks and robust regulations are crucial to address these ethical challenges.

Misinformation and Disinformation

Content that encourages interaction is given priority by AI algorithms, which frequently magnify propaganda or misleading information. Misinformation, which can instigate violence, sway public opinion, and undermine democratic processes, has flourished on social media platforms. Examples include disinformation operations that target underprivileged areas and cause societal divisions, or the dissemination of fake news before elections or exist polls. Even WhatsApp forwards these days leads to severe cybercrime and misleading offences.

Cyberbullying/Cyber crime

Rapid advancements in digital technology and the internet have revolutionized commerce, education, and communication, but they have also led to the increase of cybercrimes and cyberbullying, which present serious problems for people, society, and judicial systems around the globe. Cyberbullying is the practice of harassing, intimidating, or degrading someone via digital platforms like social media, messaging applications, online forums, and gaming platforms. Cyberbullying has a greater emotional and psychological impact than traditional bullying since it may occur anywhere, at any time, and frequently anonymously.

Cyberbullying Types-

Harassment: Consistently sending offensive remarks or messages.

Doxxing: is the public disclosure of a person's personal information without that person's consent. Making false profiles in order to damage someone's reputation is known as impersonation.

Exclusion: Wilfully keeping someone out of online communities or events.

Spreading rumours: Disseminating inaccurate or damaging information about another individual.

Cybercrimes are unlawful actions taken against people, businesses, or governments using digital technology. These offenses have serious repercussions and span from financial fraud to terrorism.

Cybercrime Types-

Identity theft is the theft of personal data for fraudulent or financial purposes.

Hacking: Unauthorized entry into networks or computer systems. Phishing is the fraudulent practice of using phony communications to try and steal private information.

Ransomware: Using malware to infect computers and demand money to unlock data. Cyberterrorism is the disruption of national security through digital means.

In the digital age, human trust is seriously threatened by the development of deepfake² technology and the growing use of malware to mislead people. While malware targets financial systems and personal data, resulting in fraud and identity theft, deepfakes combine audio and video footage to generate convincing fictional narratives that are frequently used for extortion or disinformation. These technologies undermine trust in the veracity of information and online transactions by taking advantage of our increasing reliance on digital platforms.

Social well-being and mental health

Social media platforms addictive qualities, which are fueled by AI algorithms intended to increase user interaction, have serious repercussions for mental health. Among users, particularly teenagers, problems including anxiety, depression, and body image issues are common. Additionally, cyberbullying and online harassment foster a hostile environment that negatively impacts people's mental and emotional health. In today's digital era, young influencers who post or upload content on social media platforms face increasing exposure to hatred, trolling, and abuse. This negative engagement often leads to severe mental health challenges, including stress, anxiety, and even depression. The competitive nature of social media has transformed the platform into a numbers-driven environment, where the game of "follow-unfollow" has become a critical factor for success.

This digital competitiveness fosters a culture where false narratives, misbehaviour, and toxic interactions overshadow the creative and innovative potential of these platforms. Instead of focusing on meaningful content creation or personal growth, many young influencers become fixated on achieving higher follower counts, likes, and engagement metrics. This undue pressure not only hampers their creativity but also distorts their perception of self-worth and success.

The obsession with numbers and validation often pushes influencers to adopt unhealthy behaviours, such as producing content solely for virality or engaging in trends that contradict their personal values. Moreover, the fear of losing followers or facing backlash discourages authenticity, further contributing to a toxic digital culture.

² Deepfake - Wikipedia

Ethical code of conduct in India

Protecting cyberspace from risks like cybercrimes, data breaches, and digital fraud is a major concern for India, one of the fastest-growing digital economies. The government has created a thorough legal framework to handle these issues, with the goals of guaranteeing cybersecurity, safeguarding privacy, and building confidence in digital ecosystems. The Information Technology Act, 2000 (IT Act)³, which establishes the legal basis for preventing cybercrimes and controlling online activity, is a pillar of this system. One important piece of law that regulates social media platform operations in India is the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021⁴. These rules have been implemented by the Ministry of Electronics and Information Technology to regulate digital media platforms and social media intermediaries, such as messaging apps, social networking sites, and streaming services to regularly release compliance reports outlining their attempts to uphold the rules and the established guidelines. 175 social media companies who do business in India must follow extra relevant laws and rules, such as the Indian Penal Code and the Information Technology Act of 2000, in addition to the intermediary standards. These rules create a legal framework to deal with a range of illegal online acts, including hate speech, cyberbullying, harassment, and defamation. It is essential that social media companies abide by these rules and work with law enforcement to identify and bring criminals to justice.

Provisions to address identity theft, phishing, hacking, illegal access, and cyberstalking are included in the IT Act of 2000, which was modified in 2008. While protecting sensitive data, it gives law enforcement the ability to investigate cybercrimes and bring perpetrators to justice. Cyber offenses are explicitly targeted by important parts including Section 66 (penalty for hacking) and Section 43 (unauthorized access and damage to systems). Penalties for more recent dangers, such as identity theft (Section 66C) and cyberterrorism (Section 66F), have been added by amendments. India created the Indian Computer Emergency Response Team (CERT-In), a national nodal agency for tracking, identifying, and handling cybersecurity issues, to improve cybersecurity even more. To coordinate reactions to threats, issue advisories, and raise awareness of safe online practices, CERT-In is essential⁵.

The Ministry of Electronics and Information Technology had also launched the "Stay Safe Online"⁶ campaign during India's G20 presidency to educate the public, including people with disabilities, about the dangers of using social media and the quick uptake of digital payments. The program focuses on educating users of all ages about online risk and safety measures and promoting cyber hygiene, which will strengthen citizens' cyber safety as India works to become a trillion-dollar digital economy. The Stay Safe Online program is open to people of all ages, living in both rural and urban regions, with a focus on women, children, students, teachers,

³ itbill2000.pdf

⁴ Revised-IT-Rules-2021-proposed-amended.pdf

⁵ it_amendment_act2008.pdf

⁶ Stay Safe Online | MyGov.in

faculty, older citizens, people with disabilities, and representatives of the federal and state governments.

India has taken initiative by creating new laws in addition to the IT Act. Protecting personal information, maintaining openness, and controlling how businesses process data are the main goals of the planned Digital Personal Data Protection Act, 2023⁷. The goal of this legislation is to strike a balance between user privacy and the expanding demand for data use in business and governance. To increase officials and enterprises understanding of cybersecurity, the government has also started programs like Cyber Surakshit Bharat⁸. Additionally, to enhance investigative capacities, specialist units such as the Cyber Crime Coordination Centre (I4C)⁹ and cyber forensic labs have been developed.

A Way Forward

Despite these advancements, challenges remain in enforcing laws uniformly, addressing jurisdictional issues, and managing the rapid evolution of Artificial intelligence and its threats. India's push toward digitization has increased the adoption of Artificial Intelligence (AI) in sectors like healthcare, education, governance, and commerce. However, ensuring that AI development remains ethical is crucial to prevent unintended biases and discriminatory practices. AI algorithms, often trained on large datasets, risk reflecting societal prejudices if not monitored carefully. For instance, biases in hiring algorithms or AI-enabled credit scoring systems can reinforce existing inequalities. To address this, India must establish transparent, accountable, and fair guidelines for AI systems. The NITI Aayog's National Strategy on AI, which emphasizes "AI for All," underscores the need for inclusive and unbiased AI development. This involves auditing AI systems for fairness, using representative datasets, and embedding ethical principles in their design. Additionally, frameworks like the proposed Data Protection Act should mandate accountability for AI-driven decisions to ensure that they do not infringe on individual rights or perpetuate discrimination.

Widespread digital literacy is necessary in India's quickly expanding digital economy to guarantee that individuals can utilize the internet sensibly. Many people are still ignorant of the dangers posed by AI, social media, and false information, particularly in rural and semi-urban areas. Internet accessibility has improved because to initiatives like the Digital India Campaign, but more work has to be done to inform users about ethical behavior and online safety. Government and commercial sector-backed awareness initiatives can instruct users on how to spot false information, safeguard their personal information, and utilize social media in an appropriate manner.

India is a major player in promoting worldwide cooperation on moral AI and social media usage since it is a global centre for technology. Cross-border problems including misinformation, cybercrimes, and AI abuse call for concerted action. India is already involved in international projects that support the development and application of ethical AI, such as the

⁷ Digital Personal Data Protection Act 2023.pdf

⁸ CISO Training Programme Brochure (1).pdf

⁹ Cyber Crime Portal

Global Partnership on Artificial Intelligence (GPAI)¹⁰. India can handle cross-border issues like data breaches and disinformation campaigns with the support of bilateral agreements with nations like the US and the EU as well as cooperative efforts with international organizations like the UN. A strong ecosystem for the development of ethical AI can be established domestically by encouraging collaborations between tech firms, government agencies, and civil society. In order to use social media and artificial intelligence (AI) in a way that respects human rights, India must take a holistic approach that strikes a balance between ethical responsibility and technological innovation.

With strong measures like the Digital Personal Data Protection Bill, 2023, and amendments to the Information Technology Act, 2000, aimed at addressing issues like AI misuse, cybercrimes, and disinformation, it is essential to strengthen legal and regulatory frameworks. In order to prevent technology from undermining fundamental rights, these frameworks must place a high priority on protecting equity, freedom of expression, and privacy. To protect human rights in the digital age, public awareness and digital literacy are essential. Citizens, particularly in rural and semi-urban regions, can be empowered to use digital platforms ethically and identify hazards like disinformation and data exploitation by extending projects like Digital India to incorporate digital ethics into educational curriculum and community outreach activities. In India, digital arrest technologies like facial recognition software and surveillance systems driven by artificial intelligence are also being utilized more frequently to find and capture cybercriminals. Although these advances improve public safety and law enforcement's capacity to combat cybercrimes, their implementation requires strict control to guard against abuse and guarantee adherence to constitutional protections. Protecting individual rights requires striking a balance between their usefulness and moral considerations.

India can use AI and social media to promote human rights, promote social and economic inclusion, and establish itself as a global leader in responsible technology governance by adopting an all-encompassing and moral approach. Strengthening international cooperation, updating existing laws, and enhancing the digital literacy of citizens are essential steps to ensure that India's legal framework remains robust and adaptive to emerging challenges in cyberspace.

Conclusion

In Global world, digital future is rapidly being shaped by the convergence of social media, artificial intelligence (AI), and human rights. Technology presents both enormous opportunities and formidable problems as it continues to disrupt how people interact with society, communicate, and work. In this quickly changing digital world, protecting human rights necessitates a well-rounded strategy that encourages creativity while guaranteeing responsibility and moral technology use. Through programs like Digital India, revisions to the Information Technology Act of 2000, and initiatives like the Digital Personal Data Protection Bill of 2023, India has achieved significant strides. However, the execution of these safeguards

¹⁰ Global Partnership on Artificial Intelligence - GPAI

and their flexibility in responding to new threats such as misinformation, cybercrimes, and misuse of AI will determine how effective they are.

To make sure AI systems and digital platforms don't reinforce prejudice, discrimination, or privacy abuses, transparent and inclusive frameworks are crucial. The increasing use of AI-powered digital arrest tools, such facial recognition and predictive policing, emphasizes the necessity of strict regulation to guard against abuse and preserve individual liberties. Maintaining the equilibrium between ethical considerations and technological utility requires that these instruments adhere to constitutional protections. Additionally, promoting public awareness and digital literacy, particularly in rural and semi-urban regions, can enable residents to use digital platforms responsibly, identify hazards, and defend their rights.

India's dedication to bringing technology breakthroughs into line with global ethical norms is demonstrated by its active engagement in international projects such as the Global Partnership on Artificial Intelligence (GPAI). Addressing cross-border issues like data breaches and disinformation campaigns would require cooperation between governments, businesses, and civil society organizations. India can use social media and AI to further human rights, encourage social inclusion, and boost economic growth by taking a holistic and human-centered approach. In line with the universal values of equality, freedom, and human dignity, this approach not only establishes India as a pioneer in responsible technology governance but also guarantees a safe and just online environment for all its residents.

References

1. Channi Sachdeva, V. G. (n.d.). Balancing Technology and Humanity: Perspectives on AI in Healthcare. <https://doi.org/10.4018/979-8-3693-5468-1.ch001>
2. Ergest Alite, A. T. (n.d.). AI – The Future Of Humanity And The Biggest Dilemma . *interdisciplinary journal of architecture and built environment*. <https://doi.org/10.37199/F40002706>
3. Han, S. (n.d.). Bridging Today And The Future Of Humanity: Ai Safety In 2024 And Beyond.
4. Muoma, V. C. (2024, December). Technology and Human rights: navigating the digital frontier. *African Law Students' Association – Nigerian Chapter*(October – November – December Edition), 16-25.
5. Otair, M. (2024). Bridging the Gap: How AI & Technology are Reshaping the Landscape of Humanity and Social Sciences. *Al-Noor Journal for Humanities, Vol. 2*. [https://doi.org/ https://doi.org/10.69513/jnfh.v2.n4.en18](https://doi.org/https://doi.org/10.69513/jnfh.v2.n4.en18)
6. Raj, P. (2024). Upholding Human rights in the Digital age: Exploring the interface between the UDHR and social media regulation in India. In *UNIVERSAL VALUES & NATIONAL REALITIES: EXPLORING UDHR IN THE INDIAN CONTEXT* (Pp. 187-200). ATOMIC Owl DIGITAL MEDIA PRIVATE LIMITED (OPC).
7. Youvan, D. C. (2024). Illuminating Intelligence: Bridging Humanity and Artificial Consciousness. [https://doi.org/ 10.13140/RG.2.2.23383.89767](https://doi.org/10.13140/RG.2.2.23383.89767)
8. Youvan, D. C. (2024). The Post-Human Paradigm: Exploring the Implications of AI-Augmented Humanity.

9. Youvan, D. C. (September 1, 2024). The Perception of Humanity through the Eyes of an Advanced AI: A Speculative Exploration.