International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 15, Issue 02, February 2026)

Artificial Intelligence, Human Rights, and the Right to Privacy: The Deepfake Challenges

Dr. Ikhlaq Ahmed[1], Dr. Rizvana Choudhary[2]

[1]*Assistant Professor, Dept of Political Science BGSB University, Rajouri, J&K, India.*
[2]*Assistant Professor, Dept of Sociology BGSB University, Rajouri, J&K, India.*

Abstract **-- Artificial Intelligence (AI) has resulted in a revolutionary change in the field of governance, business, healthcare, and communication. Meanwhile, generative AI, such as deepfakes (AI-generated fake images, audio, and video) has heightened threats to human rights, particularly to the right to privacy. This paper examines how the AI system data practices, inferential and generative outputs challenge crucial human rights norms. The analysis identifies weaknesses of the existing legal and ethical systems, which struggle to cope with the size of AI, its lack of transparency and dual-use. It applies different areas of research and the current shifts in policies to use deepfakes as a prime illustration of the privacy threats and the analysis of its effects on identity, autonomy, dignity, and democratic participation. This paper concludes that the existing regulatory practices are not effective unless the proactive governance approach is taken based on rights, effective enforcement and ethical design.**

Keywords**--Artificial Intelligence, Right to Privacy, Deepfakes, Human Rights, Synthetic Media, Digital Identity.**

I. INTRODUCTION

Nowadays AI technologies are crucial in all digital activities, including decision-support systems in government and personalized content in social media. Machine learning systems based on generative AI, particularly those that apply sophisticated deep learning algorithms such as Generative Adversarial Networks (GANs) and diffusion models, have the capability to generate hyper-realistic computer-generated content, and it is very similar to actual individuals. These artificial innovations, also generally referred to as deepfakes, comprise AI-generated photographs, voices, and videos that mimic people without their consent. Although AI has significant advantages, including enhanced efficiency, automated and routine non-creative work, and new creative opportunities, it also brings significant threats to human rights, particularly, the right to privacy. The right to privacy, including the authority to regulate personal data and the non-violation of invasions is an international human rights law and constitutional law that is recognized across the globe in all member countries of the United Nations.

Nevertheless, with the generative abilities of AI anybody can reproduce faces and voices without their permission, which brings up complex issues of consent, identity, trust, and autonomy. This paper presents the issue of AI and deepfake technology that endangers privacy rights and other commercial human rights. It discusses the technology of deep fakes, defines what damage they have already brought, discusses how the law and ethics are insufficient, and proposes ways of governance that address human dignity and responsibility.

II. ARTIFICIAL INTELLIGENCE AND THE RIGHT TO PRIVACY

Privacy is the right of human beings which is expressed in the documents such as the Universal Declaration of Human Rights and the International Covenant on Civil and Political Rights. According to the Constitution of India right to privacy is a fundamental right. Traditional definition of privacy in the digital age entails protection of personal communications and giving people the opportunity to regulate the collection, processing, and sharing of their personal data. Artificial intelligence systems, particularly the ones that are based on large data volumes, do not subscribe to the old concept of privacy. These systems often collect and process vast quantities of personal data, such as social media usage, tracking location, biometric and behavioral data. The personal inferences produced by them, such as predictions of their personal characteristics, preferences, or their future behavior, can take place unnoticed or uncontrollably by the users. According to scholars, AI has an inferential power that threatens the privacy of various individuals by disclosing concealed parts of their life (Levitt, 2024).

The ability of AI to produce new information on the basis of the existing information poses a threat to the informational privacy. Huge language models, facial recognition, and behavioral analytics are able to recognize trends that the user never explicitly provided, and make assumptions regarding sensitive issues such as health status, sexual orientation, or political views.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 15, Issue 02, February 2026)

Such inference processes pose privacy threats that are not merely limited to the data collection process, but may result in personal insights, as a culmination of aggregated behavior and digi-footprint. The traditional laws are hard to conform to as they tend to concentrate on particular type of data and fail to respond to these new inferences. Also, AI systems are not always developed and utilized with transparency and accountability. Proprietary algorithms are black-boxed, which means that it is not clear how data is used, the decisions made, and the inferences on people. Such non-transparency does not give people an opportunity to question the use of data or manage their online identities.

III. Deepfakes: Tech And Privacy Ills

Deepfakes are a specific branch of AI-created media that is created by using machine learning algorithms such as GANs, in which two neural networks the generator and the discriminator operate together. The generator creates artificial content and the discriminator determines its realism. The generator is trained on large collections of real-world images, audio or video and learns to generate very realistic fake content that is difficult to distinguish between synthetic media and real media. Deepfakes can imitate facial features, movements, as well as voice of a person. A deep learning model can also generate doctored videos of individuals uttering words that they never uttered or existing in circumstances they have never been in once it has sufficient training data, and this is a substantial advancement over the old-fashioned photo editing techniques.

With the development of AI models and their increased power and accessibility, deepfakes have become easier to produce. Easy to use software is used to create synthetic media with simple instructions as well as by non experts which has now contributed to the volume of deepfake content being produced online. Although not all applications of generative AI are malicious, as with artistic or educative projects, it is the same technologies that are applied in malicious ways to produce an unauthorized representation of a real individual, who may be unaware of it or without his or her consent. Deepfakes have a direct impact on the privacy of individuals directly because of the application of personal identifiers, facial features, voiceprints, and mannerisms to produce credible yet fake images. Illegal synthetic media violates informational privacy because it reveals individual characteristics before people without permission. When their digital identity can be replicated easily and shared among many people, individuals lose the control of their likenesses.

In more severe cases, deepfakes are now utilized to create explicit sexual imagery and content involving people against their will, including women and minors. UNICEF has been reporting the increased number of AI-created sexual images among children, which has triggered the demand to ban such content worldwide as it has a serious negative impact on privacy, dignity, and even the mental health of people. This exploitation demonstrates how deepfakes can destroy the autonomy and body integrity of individuals. In addition to loss of control of one image, deepfakes may be psychologically and socially harmful. An analysis of the deepfake impacts revealed that there are indications of distress, anxiety, and distrust towards people who are infected with the influence of manipulated media, especially when the information is intimate or harmful. Deepfakes may pervert the minds of the population, misinformation, and reduce confidence in digital media, in which individuals doubt the validity of credible visual information.

Women and marginalized groups frequently suffer the consequences of deepfakes being used improperly. The study of AI and women rights has revealed that deepfakes have been extensively utilized to harass, sexually exploit, and damage the reputation of women, which demonstrates the weaknesses of the digital environment. More extensive research suggests that deepfakes as gendered violence instruments are a type of online abuse that is related to the already existing tendencies of mistreatment in the online sphere. Collective privacy and democratic governance are also at risk because of deepfakes. False videos of popular individuals can change elections, mislead the mass discourse, and undermine confidence in the institutions. A UN report pointed to the risks of deepfakes to the trust of global media, and there is a need to enforce stricter criteria on the detection and verification of information in order to maintain informational integrity. Democracy is undermined when human beings cannot be certain about visual evidence.

IV. Commercial And Legal Issues

Existing privacy and data protection legislation finds it difficult to cope with the harm of deepfakes. Such laws as the European Union General Data Protection Regulation (GDPR) are primarily aimed at the ways identifiable controllers and processors treat personal data. They are not very specific in the production of synthetic media. As a result, the creation of deepfakes tends to enter the gray area of the law as the conventional privacy norms cannot suffice.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 15, Issue 02, February 2026)

In India, the legal system is challenged by the misuse of deepfakes in the protection of privacy and personal autonomy using constitutional law. Although privacy is considered to be one of the fundamental rights, loopholes in the legal regulations against AI-generated synthetic content hinder proper protection particularly when criminals cross the border of various jurisdictions.

It is legally difficult to establish liability of deepface harms. The lifecycle of synthetic media involves creators, platforms, developers of AI tools and distributors. Current liability systems are ill-equipped to identify responsibility where content is created automatically and expeditiously disseminated on decentralized systems. The protections of the intermediary liability complicate the implementation of the enforcement of laws as platforms can also be excused of liability due to the storage and transmission of content, rather than active regulation. Enforcement is also restricted by issues in cross-border jurisdiction. The creators of deepfakes might be based in a different country than their victims, and it is difficult to use just one legal framework. These problems demonstrate that there is a necessity to coordinate international measures to control synthetic media.

V. Ethical Dilemmas And Human Rights Systems

Ethically, producing deepfakes compromises autonomy by skipping the informed consent, which is one of the most important human rights principles. People being the subject of deepfakes cannot control how their image is being utilized, which is harmful to personal sovereignty and dignity. Information technology ethics emphasize the need to respect other people and seek their permission before acting on their personal data, but generative AI fails to do so in many cases. Deepfakes are more likely to widen the social inequalities that can exist by unfairly affecting vulnerable individuals and groups. Women, ethnic minorities and public figures risk being exploited and defamed more. This tendency indicates more general criticalities of digital harms that tend to support structural injustices, and a response to human rights demands that take into account the intersectional vulnerabilities.

Artificial intelligence is not always transparent, and individuals can hardly comprehend or ask questions about the way their data or pictures are utilized. Ethical AI models acknowledge that transparency, explainability, and accountability must be placed in AI system design. Devoid of such safeguards, deepfakes may harm the trust and lead to the scenarios in which the infringement of rights can increase undetected.

VI. Reactions And Style Of Governance

Technological solutions are AI detection systems that are designed to identify synthetic media. The study of deepfake detection, including zero-shot detection, and real-time detection can provide a chance to detect deepfakes early and minimize their damages (Sar et al., 2025). Nonetheless, it is not sufficient to be detected but continuous adaptations and the serious competition between the creation and detection tools are issues. Besides, innovative ways such as digital watermarking, blockchain-based provenance, and media authentication tools are being developed to defend authenticity. Such solutions will be able to generate reliable logs of original material and it will be simpler to track the manipulated media.

The regulatory tools should be altered to address the risks of deepfakes and AI that are unique. A number of states have begun enacting laws that criminalize non-consensual deepfakes and make platforms eliminate harmful content. As an example, recently the United Kingdom enacted a law that makes it a criminal offense to create non-consensual explicit deep fake images and imposes more responsibility on platforms (Reuters, 2026). Such legislative modifications are significant measures in the direction of acknowledging the evils of deepfakes as specific offenses. Nevertheless, isolated reforms cannot solve the problem unless privacy-oriented systems can be unified to insert deepfake concerns into broader digital rights safeguards. It has been analyzed that existing regulations on data protection such as GDPR and new privacy regulations should be reformed to clearly cover synthetic media, identity abuse, and unauthorized reproduction of personal characteristics (Levitt, 2024). Since the digital media is a global entity, foreign collaboration is needed. Intelligence work by governments, technologies, civil society, and international agencies can come up with a standard of content authenticity, ethical design of AI, and enforcement practices. Combined efforts to curb AI-related evils and enhance human rights safeguards have already been demanded by the United Nations and other international bodies.

VII. Conclusion

The issue of Artificial Intelligence is complex in respect of privacy rights and other associated human rights standards. This challenge is noted through deepfakes and AI-generated images, voice, and video since it enables unauthorized copying of personal identities, compromising autonomy, and causing more psychological, social, and democratic damages.

The interdisciplinary studies indicate that the existing legal and ethical standards can no longer address these emerging risks, and gaps that endanger the personal dignity and confidence in the online environment exist. Addressing these issues will require a rights-based approach to governance integrating technological solutions, laws, and moral standards. Deepfake harms should be identified by regulatory systems, accountability systems must be established, and people need to have some control over their online identity. Legal protections must not be substituted with technological solutions such as detection tools and authentication systems. In conclusion, it is important to keep privacy and human dignity safe in the face of AI advancements, and this effort is unattainable without the involvement of all stakeholders, rather than only focusing on the technical and scientific aspects of the matter.

REFERENCES

[1] Bhatia, A. (2025). *Deepfakes as a weapon of gendered terror: Non-consensual synthetic intimacy and systematic harassment of women and marginalised communities*. IJLSSS, 3(6), 148–156.

[2] Diel, A., Lalgi, T., Mellis, F. S., & Teufel, A. (2025). The harm of deepfakes: A scoping review of deepfakes' negative effects on human mind and behavior. *AI & Society*.

[3] Gagliardi, M. P. (2023). Artificial intelligence and women's rights: Deepfake technology. In *Artificial Intelligence and Human Rights* (pp. 235–247). Oxford University Press.

[4] Levitt, A. (2024). Navigating privacy in the age of AI: Evaluating the adequacy of the GDPR and CCPA to combat data exploitation and deepfake technology. NHSJS.

[5] Nowak, A., Tóth, B., & Ionescu, A. (2025). Legal challenges of deepfakes: Liability, harm, and regulatory responses. *Legal Studies in Digital Age*.

[6] Reuters. (2025, July 11). *UN report urges stronger measures to detect AI-driven deepfakes*.

[7] Reuters. (2026, February 5). *UNICEF calls for criminalization of AI content depicting child sex abuse*.

[8] Sar, A., Roy, S., Choudhury, T., & Abraham, A. (2025). *Zero-shot visual deepfake detection: Can AI predict and prevent fake content before it's created?* arXiv.

[9] Sukhram, A. (2025). Legal challenges of deepfake and synthetic media in India. *Indian Journal for Research in Law and Management*.

[10] UN General Assembly. (1948). *Universal Declaration of Human Rights*.