DEEPFAKE TECHNOLOGY AND THE NEED FOR A COMPREHENSIVE LEGAL FRAMEWORK IN INDIA

Vaishnavi Agrahari, B.A.LL.B., United University Prayagraj

ABSTRACT

The rapid development of artificial intelligence has transformed various aspects of human life, offering unprecedented opportunities for innovation and communication. However, the emergence of deepfake technology has simultaneously generated serious legal, ethical, and social concerns. Deepfakes are synthetic media created through sophisticated machine learning techniques capable of manipulating audio, images, and videos to produce highly realistic yet deceptive content. While this technology has legitimate applications in education, entertainment, and accessibility services, its misuse has become increasingly prevalent. Deepfakes have been employed to disseminate misinformation, facilitate financial fraud, undermine electoral integrity, violate privacy, and perpetrate gender-based cyber exploitation through the creation of non-consensual intimate imagery. In India, the increasing circulation of manipulated digital content has exposed significant inadequacies within the existing legal framework. Although provisions under the Information Technology Act, 2000 , the Bharatiya Nyaya Sanhita, 2023 , the Digital Personal Data Protection Act, 2023 , and the Copyright Act, 1957 may provide partial remedies, there remains an absence of legislation specifically addressing the unique challenges posed by deepfake technology. This regulatory gap raises important constitutional questions concerning freedom of speech and expression, the right to privacy, dignity, and the State's obligation to protect individuals from technologically facilitated harm.

Using a doctrinal research methodology, this paper examines the evolution of deepfake technology, analyses the existing legal mechanisms governing its misuse in India, and evaluates the constitutional and ethical concerns arising from its widespread proliferation. The article argues that India requires a comprehensive and technologically responsive legal framework that balances innovation with accountability while safeguarding fundamental rights and preserving public trust in digital information systems.

Keywords: Deepfakes, Artificial Intelligence, Cyber Law, Misinformation, Constitutional Rights, Digital Governance.

I. INTRODUCTION

Artificial intelligence has transformed the way information is created, shared, and consumed in modern society. While advancements in machine learning have contributed positively to sectors such as healthcare, education, and entertainment, they have also enabled the development of technologies capable of manipulating digital content with striking accuracy. Among these, deepfake technology has emerged as a significant concern due to its ability to generate highly realistic yet fabricated audio, video, and image content.

The term "deepfake" combines "deep learning" and "fake" and refers to synthetic media created using artificial intelligence techniques, particularly Generative Adversarial Networks (GANs).[1] By analysing vast datasets, these systems can imitate facial expressions, voices, and mannerisms, making manipulated content appear authentic. Although initially associated with research and entertainment, deepfake tools have become increasingly accessible through open-source software and online platforms. The misuse of deepfake poses serious legal and social challenges. They have been used to spread misinformation, damage reputations, commit fraud, and interfere with democratic processes. Particularly alarming is the rise of non-consensual sexually explicit deepfake targeting women, which intensifies existing forms of gender-based cyber violence. Additionally, the growing presence of synthetic media risks eroding public trust in genuine evidence, creating what scholars describe as the "liar's dividend."[2] In India, incidents involving deepfake have exposed the limitations of existing legal frameworks. Current laws addressing defamation, identity theft, obscenity, and cyber offences were enacted before the emergence of sophisticated AI-generated content and therefore provide only fragmented remedies. This raises important questions regarding accountability, victim protection, and effective regulation.

The regulation of deepfake also involves balancing competing constitutional values. While Article 21 of the Constitution protects privacy and informational autonomy,[3] Article 19(1)(a) guarantees freedom of speech and expression, subject to reasonable restrictions under Article 19(2). Any regulatory framework must therefore safeguard individual dignity and democratic interests without unnecessarily restricting innovation or legitimate expression. Against this

[1] Ian J. Goodfellow et al., Generative Adversarial Nets, in Advances in Neural Information Processing Systems 2672 (2014).
[2] Robert Chesney & Danielle Citron, Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security, 107 Calif. L. Rev. 1753, 1768–69 (2019).
[3] Justice K.S. Puttaswamy (Retd.) v. Union of India, (2017) 10 S.C.C. 1.

backdrop, this article examines the legal and constitutional implications of deepfake technology in India. It evaluates the adequacy of existing laws, explores the ethical concerns associated with synthetic media, and argues for a comprehensive legislative framework capable of addressing the unique harms posed by deepfake while remaining consistent with constitutional principles and democratic values.

II. HISTORICAL EVOLUTION OF DEEPFAKE TECHNOLOGY AND ITS EMERGENCE IN INDIA

2.1 Historical Evolution

The alteration of images and videos has existed for decades through conventional editing techniques used for artistic, commercial, and deceptive purposes. However, advances in artificial intelligence have significantly enhanced the realism and accessibility of such manipulation. Deepfake technology enables the creation of synthetic media that convincingly replicates human appearance, speech, and behaviour. The development of deepfake is closely linked to progress in machine learning, particularly the introduction of Generative Adversarial Networks (GANs) by Ian Goodfellow and his team in 2014[4]. These systems generate increasingly realistic content by using one neural network to create synthetic material and another to assess its authenticity. Initially, deepfake technology had several beneficial applications, including medical imaging, film production, language translation, and accessibility services. However, concerns intensified in 2017 when non-consensual manipulated videos involving celebrities appeared online. The subsequent availability of user-friendly software made the creation of convincing deepfake accessible even to individuals with limited technical expertise.

Today, deepfake encompass voice cloning, synthetic identities, speech modification, and real-time body manipulation. While these innovations offer legitimate opportunities, they have also facilitated political misinformation, financial fraud, identity theft, reputational harm, and the creation of non-consensual intimate content. In India, the widespread use of digital platforms has amplified these risks, exposing gaps in existing legal protections. The emergence of deepfake reflects the broader challenge of regulating technologies that have both beneficial and harmful potential. The Indian experience demonstrates that deepfake pose serious threats to

[4] Ian J. Goodfellow et al., Generative Adversarial Nets, in *Advances in Neural Information Processing Systems 2672 (2014).

privacy, democratic integrity, gender justice, and public trust. Addressing these concerns requires a comprehensive legal framework that balances innovation with accountability and the protection of fundamental rights.

III. EXISTING LEGAL FRAMEWORK GOVERNING DEEPFAKES IN INDIA

3.1 Introduction

India presently lacks a dedicated statute specifically regulating deepfake technology. Instead, harms arising from synthetic media are addressed through a combination of provisions contained in cyber laws, criminal legislation, data protection laws, intellectual property frameworks, and constitutional principles. Although these mechanisms may provide remedies in particular circumstances, their fragmented nature frequently results in uncertainty regarding applicability, enforcement, and accountability.

The absence of a specialised legislative framework has become increasingly problematic as artificial intelligence-generated content assumes greater sophistication and prevalence. Existing provisions were generally enacted to address traditional forms of cybercrime and therefore may not adequately capture the unique characteristics of deepfake-related harms. Nevertheless, an examination of the present legal framework remains necessary to understand both its strengths and limitations.

3.2 Information Technology Act, 2000

The Information Technology Act, 2000 constitutes the principal legislation governing cyber activities in India. Several provisions of the Act may be invoked in cases involving harmful deepfake. Section 66C criminalizes identity theft involving fraudulent use of electronic signatures, passwords, or unique identification features.[5]Deepfake creators impersonating individuals through synthetic media may potentially attract liability under this provision. Similarly, Section 66D penalises cheating by personation through computer resources or communication devices.[6]Where deepfake are employed to deceive individuals for financial gain or other fraudulent purposes, this provision may provide an avenue for prosecution. Furthermore, Sections 67 and 67A prohibit the publication or transmission of obscene and

[5] Information Technology Act, 2000, s 66C, No. 21, Acts of Parliament, 2000 (India).
[6] Information Technology Act, 2000, s 66D, No. 21, Acts of Parliament, 2000 (India).

sexually explicit material in electronic form.[7] These provisions assume particular relevance in addressing non-consensual intimate deepfake targeting women. Despite their utility, these provisions do not explicitly recognises synthetic media as a distinct category of harmful content. Consequently, enforcement authorities may encounter challenges in establishing appropriate evidentiary standards and determining the scope of liability.

3.3 Bharatiya Nyaya Sanhita, 2023[8]

Although the BNS does not specifically regulate deepfake, its provisions relating to defamation, cheating, fraud, public mischief, obscenity, and offences against women may apply depending on the nature of the harm caused. However, the absence of explicit recognition of deepfake, particularly AI-generated sexual content, creates enforcement challenges.

3.4 Digital Personal Data Protection Act, 2023[9]

The DPDP Act strengthens privacy protection through consent-based processing of personal data. Since deepfake often involve the unauthorised use of images, videos, and voice recordings, the Act offers some safeguards. However, its focus on data processing and the difficulty of regulating anonymous creators limit its effectiveness in addressing deepfake-related harms.

3.5 Copyright Act, 1957[10]

The Copyright Act may provide remedies where deep fakes involve the unauthorised use of copyrighted material. Nevertheless, AI-generated content raises unresolved questions regarding authorship and ownership, areas where Indian copyright law lacks clarity.

3.6 IT Rules, 2021[11]

The IT Rules impose due diligence obligations on intermediaries, including grievance redressal and content removal mechanisms. While these measures enhance platform accountability, the

[7] Information Technology Act, 2000, §§ 67, 67A, No. 21, Acts of Parliament, 2000 (India).
[8] The Bharatiya Nyaya Sanhita, No. 45 of 2023, India Code (2023).
[9] Digital Personal Data Protection Act, No. 22 of 2023, INDIA CODE (2023).
[10] Copyright Act, No. 14 of 1957, INDIA CODE (1957).
[11] Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, Gazette of India, Extraordinary, Part II, sec. 3(i) (Feb. 25, 2021).

rapid spread of deepfake often outpaces notice-and-takedown procedures, limiting their practical effectiveness.

3.7 Constitutional Dimensions

Deepfake regulation must balance freedom of expression under Article 19(1)(a) with the need to protect privacy, dignity, and reputation under Article 21. Any restrictions should be reasonable and proportionate. The recognition of privacy as a fundamental right in Justice K.S. Puttaswamy v. Union of India[12] provides a constitutional basis for addressing non-consensual deepfake.

3.8 Critical Evaluation

India's existing laws provide partial remedies against deepfake-related harms through criminal, privacy, copyright, and intermediary regulations. However, the framework remains fragmented and reactive. The absence of a specific legal definition, preventive safeguards, platform standards, and victim-centred remedies highlights the need for a comprehensive and future-oriented regulatory framework.

IV. Judicial Developments and Emerging Challenges Relating to Deepfake Technology in India

4.1 Introduction

In the absence of dedicated deepfake legislation, Indian courts have relied on existing constitutional principles relating to privacy, dignity, free speech, and intermediary liability to address challenges arising from digital technologies. Judicial decisions concerning online harms and fundamental rights provide important guidance for regulating deepfake while balancing innovation and constitutional values.

4.2 Justice K.S. Puttaswamy (Retd.) v. Union of India(2017)[13]

In this landmark judgment, the Supreme Court recognised privacy as a fundamental right under Article 21. The Court held that privacy includes informational self-determination, autonomy,

[12] Justice K.S. Puttaswamy (Retd.) v. Union of India, (2017) 10 S.C.C. 1.
[13] Justice K.S. Puttaswamy (Retd.) v. Union of India, (2017) 10 S.C.C. 1.

and control over personal identity. These principles are highly relevant to deepfake, as the unauthorised use of a person's image or voice undermines dignity and privacy. The decision therefore provides a strong constitutional basis for regulating harmful synthetic media.

4.3 Shreya Singhal v. Union of India(2015)[14]

The Supreme Court struck down Section 66A of the Information Technology Act for violating freedom of speech under Article 19(1)(a). The judgment emphasised that restrictions on expression must be reasonable and narrowly tailored under Article 19(2).[15] In the context of deepfake, this highlights the need for legislation that targets harmful content without restricting legitimate uses such as satire, artistic expression, or political commentary. The Court also cautioned against imposing disproportionate monitoring obligations on intermediaries, underscoring the importance of procedural safeguards and balanced regulation.

4.4 Intermediary Responsibility and Emerging Judicial Trends

Indian courts have increasingly acknowledged the influential role played by intermediaries in facilitating digital communication. Judicial observations in cases involving harmful online content have highlighted the necessity of establishing effective grievance redressal mechanisms capable of responding to evolving technological threats. Deepfake-related harms often require urgent intervention due to the speed with which manipulated content may spread across digital platforms. Delayed removal may substantially aggravate reputational damage, psychological trauma, and financial losses suffered by victims. Consequently, courts may increasingly confront questions concerning the extent to which intermediaries should proactively detect and remove synthetic media that violates legal standards.

However, imposing stringent monitoring obligations presents practical and constitutional challenges. Automated detection systems may produce inaccurate results, potentially resulting in the removal of legitimate content. Balancing victim protection with freedom of expression therefore remains a central judicial concern.

[14] Shreya Singhal v. Union of India, (2015) 5 S.C.C. 1.
[15] INDIA CONST. art. 19, cl. 2.

V. ETHICAL ISSUES AND CONSTITUTIONAL PERSPECTIVE

5.1 Ethical Dimensions of Deepfake Technology

The emergence of deepfake technology presents complex ethical challenges extending beyond conventional legal considerations. Although artificial intelligence has the potential to enhance creativity, accessibility, and innovation, its misuse raises significant concerns regarding consent, autonomy, trust, and human dignity. The ethical debate surrounding deep fakes centres upon whether technological advancement can be permitted to operate without adequate safeguards against foreseeable harm. One of the most pressing ethical concerns relates to the absence of meaningful consent. Deepfakes frequently utilise an individual's facial features, voice recordings, or mannerisms without authorisation. The creation and dissemination of manipulated content involving a person's identity directly undermines individual autonomy and the ability to exercise control over one's own representation in society. The unauthorised appropriation of personal attributes therefore raises serious questions concerning respect for human dignity and personal agency.Another ethical issue concerns the erosion of public trust. Democratic societies depend upon the reliability of information to facilitate informed decision-making and meaningful public participation. Deepfakes possess the capacity to distort reality by presenting fabricated events as authentic occurrences. As synthetic media becomes increasingly sophisticated, individuals may struggle to distinguish truth from deception. This phenomenon threatens confidence in journalism, governmental institutions, and digital communication systems.

Scholars have additionally identified the emergence of the "liar's dividend,"[16] whereby the existence of deepfake technology enables individuals to dismiss genuine evidence as fabricated. This development may significantly impede accountability mechanisms by allowing perpetrators of misconduct to challenge the authenticity of legitimate recordings. Consequently, deepfake not only generate false information but also undermine the evidentiary value of authentic material. The disproportionate impact of deepfake upon women represents another critical ethical concern. Non-consensual sexually explicit deepfake have become increasingly prevalent across digital platforms, frequently targeting women without their knowledge or consent. Such conduct reinforces harmful gender stereotypes, facilitates cyber

[16] Robert Chesney & Danielle Citron, Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security, 107 Calif. L. Rev. 1753, 1768–69 (2019).

exploitation, and contributes to psychological distress. The gendered nature of these harms necessitates responses grounded in principles of equality, dignity, and social justice.

Furthermore, ethical questions arise concerning the responsibilities of technology developers and digital intermediaries. Entities involved in designing artificial intelligence systems may possess obligations to incorporate safeguards capable of preventing foreseeable misuse. Similarly, online platforms facilitating content dissemination may bear ethical responsibilities to implement effective reporting mechanisms and cooperate with enforcement authorities. The principle of corporate accountability therefore assumes growing importance within discussions surrounding deepfake governance.

5.2 Constitutional Perspective

Article 14: Equality Before Law and Equal Protection of Laws[17]

Deepfakes raise concerns under Article 14 due to their disproportionate impact on women and vulnerable groups, particularly through non-consensual intimate content. The State has a constitutional obligation to provide equal protection by addressing emerging technological harms through effective legal measures.

Article 19: Freedom of Speech and Expression[18]

Article 19(1)(a) protects freedom of speech and expression, while Article 19(2) permits reasonable restrictions on grounds such as defamation, public order, and morality. Deepfake regulation must balance legitimate uses, such as satire and artistic expression, with the need to prevent harmful content involving fraud, misinformation, or reputational harm. Any restrictions must satisfy the test of proportionality.

Article 21: Right to Life and Personal Liberty[19]

Deepfakes implicate the rights to privacy, dignity, autonomy, and reputation protected under Article 21. Non-consensual manipulation of an individual's identity can cause significant emotional and reputational harm. The recognition of informational privacy in Justice K.S.

[17] INDIA CONST. art. 14.
[18] INDIA CONST. art. 19, cl. 1(a).
[19] INDIA CONST. art. 21

Puttaswamy (Retd.) v. Union of India[20] underscores the State's responsibility to establish safeguards against such technological threats.

5.3 Human Rights Perspective

International human rights instruments recognise privacy, dignity, and freedom of expression as essential components of human flourishing. Article 12 of the Universal Declaration of Human Rights protects individuals against arbitrary interference with privacy, family, and reputation.[21] Similarly, Article 19 recognises the right to freedom of opinion and expression.[22]The challenge presented by deepfake lies in ensuring that these rights remain mutually reinforcing rather than mutually destructive. Protecting individuals from technological abuse should not necessitate the suppression of legitimate expression. Conversely, the protection of free speech cannot justify violations of privacy, dignity, and equality.

Human rights principles therefore support the development of balanced regulatory frameworks grounded in transparency, accountability, proportionality, and access to effective remedies. Victims of harmful deepfake should possess opportunities to seek redress through accessible legal processes, while creators engaging in legitimate expressive activities should remain protected against arbitrary restrictions.

VI. CRITICAL ANALYSIS

Deepfake technology has exposed significant shortcomings in India's existing legal framework. While laws such as the Information Technology Act, 2000, the Bharatiya Nyaya Sanhita, 2023, the Digital Personal Data Protection Act, 2023, and the Copyright Act, 1957 offer limited remedies, they do not specifically address the unique challenges posed by AI-generated synthetic media. A major concern is the absence of a statutory definition of deepfake, which creates uncertainty in enforcement, classification, and the determination of liability. Existing laws are largely reactive, providing remedies only after harm has occurred, often after manipulated content has already caused irreversible damage. Regulating deepfake also requires balancing freedom of expression under Article 19(1)(a) [23]with the rights to privacy, dignity,

[20] Justice K.S. Puttaswamy (Retd.) v. Union of India, (2017) 10 S.C.C. 1.
[21] G.A. Res. 217 (III) A, Universal Declaration of Human Rights art. 12 (Dec. 10, 1948).
[22] G.A. Res. 217 (III) A, Universal Declaration of Human Rights art. 19 (Dec. 10, 1948).
[23] INDIA CONST. art. 19, cl. 1(a).

and equality under Articles 14 and 21.[24] Any legal framework must therefore be proportionate and narrowly tailored. Additionally, intermediary accountability remains a challenge, requiring effective grievance mechanisms and timely content moderation without imposing excessive monitoring obligations.

The misuse of deepfake in electoral contexts further threatens democratic integrity by enabling the spread of misinformation. These concerns highlight the need for specialised legislation incorporating clear definitions, victim-centric remedies, intermediary responsibilities, and procedural safeguards to effectively address the evolving risks posed by deepfake technology.

VII. CONCLUSION

Deepfake technology has emerged as a significant legal and societal challenge in the digital era. Although artificial intelligence has driven innovation and economic growth, the misuse of synthetic media threatens privacy, reputation, dignity, democratic processes, and national security. The ability of deepfake to manipulate reality highlights the urgent need for effective regulation.India's current legal framework addresses certain aspects of deepfake-related harms through laws on cybercrime, defamation, obscenity, data protection, and intellectual property. However, these provisions are fragmented and insufficient to tackle the unique challenges posed by deepfake. The absence of dedicated legislation has created regulatory gaps and limited remedies for victims.

Any legal response must balance the constitutional guarantee of freedom of speech under Article 19(1)(a)[25] with the rights to privacy, dignity, and equality under Articles 14 and 21.[26] Comparative regulatory approaches adopted by other jurisdictions, including disclosure obligations and intermediary accountability, offer valuable guidance for India.A comprehensive framework should provide clear statutory definitions, effective grievance redressal mechanisms, stronger privacy safeguards, intermediary responsibilities, and specialised enforcement measures. Alongside legislative reforms, public awareness and technological detection tools are essential to counter misinformation and digital manipulation.

Ultimately, regulating deepfake is not merely a technological concern but a constitutional

[24] INDIA CONST. arts. 14, 21.
[25] INDIA CONST. art. 19, cl. 1(a).
[26] INDIA CONST. arts. 14, 21.

imperative aimed at preserving trust, authenticity, and human dignity in an increasingly digital society.

Declaration: The author has used an AI-assisted writing tool solely for language editing and grammar refinement purposes. All legal research, analysis, and conclusions presented in this manuscript are the author's own work.

BIBLIOGRAPHY

Cases:

i. Justice K.S. Puttaswamy (Retd.) v. Union of India.

ii. Shreya Singhal v. Union of India.

Statutes and Legislative Materials:

i. Information Technology Act, 2000.

ii. Bharatiya Nyaya Sanhita, 2023.

iii. Digital Personal Data Protection Act, 2023.

iv. Copyright Act, 1957.

v. Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021.

vi. Constitution of India.

International Instruments

i. Universal Declaration of Human Rights.

Journal Articles and Academic Sources

i. Ian J. Goodfellow et al., Generative Adversarial Nets, in Advances in Neural Information Processing Systems 2672 (2014).

ii. Robert Chesney & Danielle Keats Citron, Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security, 107 California Law Review 1753 (2019).

Books

Law, Policy and Artificial Intelligence (Edward Elgar Publishing, 2020).

i. The Age of Surveillance Capitalism (Profile Books, 2019).

ii. The Future of Reputation (Yale University Press, 2007).