TRUTH IN THE AGE OF DEEPFAKES: EVALUATING THE ADEQUACY OF INDIA'S CYBER LAW FRAMEWORK

Ishita Rai, LL.B., Department of Law,
Bharati Vidyapeeth (Deemed to be University), New Delhi

ABSTRACT

The emergence of deepfake technology has transformed the digital landscape by enabling the creation of highly realistic yet fabricated audio, video, and image content. While deepfakes have legitimate applications in entertainment, education, and accessibility, their misuse has intensified concerns relating to identity theft, misinformation, cyber fraud, non-consensual intimate imagery, and electoral manipulation. These developments have exposed significant challenges for legal systems that were not designed to regulate AI-generated synthetic media. This paper examines whether India's existing cyber law framework adequately addresses the creation and dissemination of deepfakes. Adopting a doctrinal research methodology, it analyses the Information Technology Act, 2000, the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, the Digital Personal Data Protection Act, 2023, and relevant provisions of the Bharatiya Nyaya Sanhita, 2023. The paper argues that although these laws offer fragmented remedies against specific harms, they fail to provide a coherent regulatory framework for deepfakes. It concludes by proposing targeted legal reforms that strengthen accountability, improve victim protection, and preserve innovation and freedom of expression in India's evolving digital ecosystem.

Keywords: Deepfakes, Artificial Intelligence, Cyber Law, Information Technology Act, Digital Personal Data Protection Act, Intermediary Liability, India.

1. Introduction

The rapid advancement of artificial intelligence has transformed the creation and dissemination of digital content, making it increasingly difficult to distinguish authentic material from manipulated media. Among the most significant developments is the emergence of *deepfakes* AI-generated audio, video, and images that convincingly imitate real individuals. Once confined to research laboratories and niche online communities, deepfake technology is now widely accessible through commercially available applications and open-source tools. While the technology offers legitimate benefits in fields such as filmmaking, education, accessibility, and digital communication, its misuse has raised serious legal and ethical concerns across the world.

In India, the growing circulation of deepfakes has exposed individuals, public institutions, and digital platforms to new forms of cyber harm. Fabricated videos of public figures, non-consensual intimate imagery, financial scams using cloned voices, and manipulated political content demonstrate how synthetic media can undermine personal privacy, reputation, public trust, and democratic discourse. Unlike conventional forms of digital manipulation, modern deepfakes are often capable of deceiving even careful viewers, making detection increasingly difficult and amplifying the scale of potential harm. As the technology becomes more sophisticated, existing legal mechanisms are being tested in ways that were not anticipated when India's principal cyber legislation was enacted.

The Information Technology Act, 2000 remains the cornerstone of India's cyber law framework. Together with the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, the Digital Personal Data Protection Act, 2023, and relevant provisions of the Bharatiya Nyaya Sanhita, 2023, it offers remedies for offences such as identity theft, cheating by personation, privacy violations, and the circulation of unlawful digital content. However, none of these laws specifically regulate AI-generated deepfakes or comprehensively address issues such as algorithmic content generation, attribution of liability, platform accountability, or the rapid removal of harmful synthetic media. The result is a fragmented legal framework that often requires courts and enforcement agencies to adapt provisions designed for earlier forms of cyber misconduct.

This paper examines whether India's existing cyber law framework is capable of effectively regulating the creation and dissemination of AI-generated deepfakes. It adopts a doctrinal

research methodology based on the analysis of statutes, judicial precedents, government policy documents, and comparative legal developments in other jurisdictions. The paper argues that although existing laws provide partial protection against certain consequences of deepfakes, they do not adequately address the distinctive legal challenges posed by AI-generated synthetic media. It therefore contends that India should strengthen its existing cyber law framework through targeted legislative amendments, clearer intermediary obligations, and regulatory safeguards that promote both accountability and technological innovation.

2. Understanding Deepfakes and Their Legal Implications

Deepfakes are a form of synthetic media created using artificial intelligence techniques, particularly deep learning models such as Generative Adversarial Networks (GANs), to generate or manipulate images, audio, or video with a high degree of realism. Unlike conventional digital editing, deepfakes are capable of replicating a person's facial expressions, voice, and mannerisms so convincingly that fabricated content may be perceived as authentic. Robert Chesney and Danielle Keats Citron describe deepfakes as a technological development that significantly lowers the cost of deception by enabling the rapid creation and dissemination of realistic yet false digital content.[1] Their observation highlights that the legal concern surrounding deepfakes lies not merely in the technology itself but in its capacity to undermine trust in digital evidence and online communication.

Deepfake technology is not inherently unlawful. It has been employed for legitimate purposes, including film production, language translation, digital restoration of historical archives, accessibility tools for persons with disabilities, and educational content. However, the widespread availability of AI-powered content generation tools has substantially increased the risk of misuse. Henry Ajder notes that the democratisation of deepfake technology has reduced technical barriers, allowing individuals with minimal expertise to create convincing synthetic media capable of causing significant personal and societal harm.[2]

The legal implications of deepfakes extend beyond traditional cyber offences because the harm frequently affects multiple legally protected interests simultaneously. A manipulated video may infringe an individual's privacy, damage reputation, facilitate financial fraud, violate data

[1] Robert Chesney & Danielle Keats Citron, Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security, 107 *California Law Review* 1753 (2019)
[2] Henry Ajder, The State of Deepfakes: Landscape, Threats, and Impact (Deeptrace Labs, 2019).

protection principles, and interfere with electoral integrity. One of the most concerning forms of misuse involves non-consensual intimate deepfakes, where a person's likeness is digitally superimposed onto explicit content without consent. Such material can cause severe psychological, social, and reputational harm despite being entirely fabricated. Similarly, AI-generated voice cloning has enabled sophisticated impersonation scams in which fraudsters imitate family members, company executives, or public officials to deceive victims into transferring money or disclosing confidential information.

These developments present significant legal challenges because deepfakes blur the distinction between authentic and fabricated digital evidence. Danielle Keats Citron argues that synthetic media increasingly threatens individual autonomy, privacy, and democratic discourse by eroding public confidence in the authenticity of online information.[3] This "liar's dividend" where genuine evidence may be dismissed as fake simply because convincing fabrications exist further complicates judicial proceedings, criminal investigations, and public accountability. Existing legal frameworks, including India's Information Technology Act, 2000, primarily address the consequences of harmful digital conduct rather than the unique technological characteristics of AI-generated synthetic media. Consequently, questions relating to attribution of liability, intermediary responsibility, evidentiary reliability, and timely removal of harmful content remain only partially addressed under the current legal regime.

3. Existing Indian Legal Framework

India does not presently have legislation specifically regulating deepfakes. Instead, legal action against AI-generated synthetic media depends upon a combination of cyber law, criminal law, data protection law, and intermediary regulation. Although these statutes provide remedies against particular forms of harm caused by deepfakes, they were enacted to address broader categories of digital misconduct rather than AI-generated content. Consequently, the current legal framework is largely reactive, requiring existing provisions to be adapted to technological developments that were not contemplated when these laws were enacted.

A. Information Technology Act, 2000

The Information Technology Act, 2000 ("IT Act") remains the principal legislation governing

[3] Danielle Keats Citron, The Fight for Privacy: Protecting Dignity, Identity, and Love in the Digital Age (W.W. Norton & Co. 2022).

cyber offences in India. While the Act does not refer to artificial intelligence or deepfakes, several of its provisions may be invoked depending on the nature of the harm caused. Section 66C criminalises identity theft through the fraudulent use of another person's electronic signature, password, or unique identification feature, whereas Section 66D penalises cheating by personation using computer resources.[4] These provisions may apply where deepfakes are used to impersonate individuals for financial fraud or deception. Likewise, Sections 67, 67A, and 67B prohibit the publication or transmission of obscene and sexually explicit electronic material, making them particularly relevant in cases involving non-consensual intimate deepfakes.[5]

Despite their utility, these provisions regulate the consequences of deepfake misuse rather than the technology itself. They neither define synthetic media nor establish standards for liability where AI-generated content causes harm without a clearly identifiable perpetrator. As Citron observes, laws drafted to address traditional cyber offences often struggle to capture the distinctive features of AI-generated deception, particularly where the technology enables large-scale dissemination with minimal human intervention.[6]

B. Information Technology Rules, 2021 and Intermediary Liability

The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 supplement the IT Act by prescribing due diligence obligations for intermediaries. Social media platforms are expected to remove unlawful content upon receiving valid legal directions and to exercise greater responsibility in addressing harmful digital material. The significance of intermediary liability was clarified by the Supreme Court in *Shreya Singhal v. Union of India*, which held that intermediaries lose statutory protection only upon receiving actual knowledge through a court order or an authorised governmental notification.[7]

This framework functions reasonably well where unlawful content can be readily identified. Deepfakes, however, introduce additional complexities. AI-generated content may spread across multiple platforms within minutes, while determining whether a video is manipulated often requires technical verification. Delayed detection substantially reduces the effectiveness

[4] Information Technology Act, No. 21 of 2000, §§ 66C, 66D (India).
[5] Information Technology Act, No. 21 of 2000, §§ 67, 67A, 67B (India).
[6] Danielle Keats Citron, *The Fight for Privacy: Protecting Dignity, Identity, and Love in the Digital Age* 181–205 (W.W. Norton & Co. 2022).
[7] *Shreya Singhal v. Union of India*, (2015) 5 SCC 1.

of post-publication remedies, particularly where reputational harm or electoral misinformation has already occurred. Consequently, the existing intermediary framework places greater emphasis on content removal than on proactive risk mitigation or transparency concerning AI-generated media.

C. Digital Personal Data Protection Act, 2023

The Digital Personal Data Protection Act, 2023 represents a significant step towards strengthening informational privacy in India. By regulating the processing of personal data and recognising individuals' rights over their digital information, the Act provides indirect protection against certain forms of deepfake misuse. Where AI systems collect or process facial images, voice samples, or biometric information without lawful consent, affected individuals may invoke protections available under the Act.[8]

Nevertheless, the legislation primarily governs the processing of personal data rather than the creation or circulation of synthetic media. A deepfake may be generated using publicly available photographs or lawfully obtained recordings without necessarily violating every requirement relating to data processing. Consequently, important questions concerning consent, attribution, and remedies for AI-generated impersonation remain outside the Act's direct scope.

D. Bharatiya Nyaya Sanhita, 2023

Certain provisions of the Bharatiya Nyaya Sanhita, 2023 may also become relevant where deepfakes facilitate offences such as cheating, forgery, defamation, identity-based deception, or the dissemination of false information causing public harm. However, similar to the IT Act, these offences continue to operate on the assumption that criminal conduct originates from identifiable human actors. They do not specifically address situations where sophisticated AI systems autonomously generate or modify deceptive content that is subsequently disseminated through digital platforms.

The cumulative effect of these legislative instruments is a fragmented regulatory framework. Existing laws undoubtedly provide avenues for prosecuting particular harms resulting from deepfakes, yet they fail to establish a coherent legal regime governing the creation, distribution,

[8] Digital Personal Data Protection Act, No. 22 of 2023 (India).

detection, and accountability of AI-generated synthetic media. As Chesney and Citron caution, the unprecedented realism and accessibility of deepfake technology challenge legal systems that rely upon conventional assumptions regarding authenticity, attribution, and evidentiary reliability.[9] India's current cyber law framework therefore addresses the symptoms of deepfake misuse more effectively than the technological phenomenon itself, underscoring the need for a more comprehensive and future-oriented regulatory approach.

4. Judicial Responses and Comparative Perspectives

Although Indian courts have not yet delivered a comprehensive judgment specifically addressing AI-generated deepfakes, several landmark decisions have established legal principles that are directly relevant to issues of privacy, intermediary liability, and digital evidence. When examined alongside comparative international approaches, these decisions demonstrate that existing legal doctrines can address certain consequences of deepfake misuse but remain inadequate for regulating the technology itself.

One of the most influential judgments in India's cyber law jurisprudence is *Shreya Singhal v. Union of India*, in which the Supreme Court interpreted Section 79 of the Information Technology Act, 2000 and clarified the scope of intermediary liability.[10] The Court held that intermediaries are required to remove unlawful content only upon receiving actual knowledge through a court order or a notification issued by the appropriate government. This interpretation sought to balance freedom of speech with intermediary accountability by preventing arbitrary takedowns of online content. However, the decision predates the emergence of generative AI and does not address situations where platforms host or distribute highly realistic AI-generated content capable of causing immediate and irreversible harm. In cases involving viral deepfakes, reliance on post-publication notice-and-takedown mechanisms may prove inadequate because harmful content often spreads across multiple platforms before legal intervention becomes possible.

The constitutional importance of informational privacy was firmly recognised in *Justice K.S. Puttaswamy (Retd.) v. Union of India*, where the Supreme Court held that privacy forms an integral part of the right to life and personal liberty under Article 21 of the Constitution.[11] The

[9] Robert Chesney & Danielle Keats Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 *Calif. L. Rev.* 1753 (2019).
[10] *Shreya Singhal v. Union of India*, (2015) 5 SCC 1.
[11] *Justice K.S. Puttaswamy (Retd.) v. Union of India*, (2017) 10 SCC 1.

judgment acknowledged that technological developments significantly increase the capacity to collect, process, and exploit personal information. Deepfakes intensify these concerns by enabling the unauthorised use of an individual's facial features, voice, and other identifiable characteristics to create fabricated content. While *Puttaswamy* provides an important constitutional foundation for protecting personal autonomy and dignity, it does not prescribe specific remedies for AI-generated impersonation or synthetic media. Consequently, victims must continue to rely upon a combination of statutory provisions that were not designed with deepfake technology in mind.

Questions relating to evidentiary reliability also become increasingly significant in the context of synthetic media. In *Anvar P.V. v. P.K. Basheer*, the Supreme Court emphasised the importance of maintaining the authenticity and integrity of electronic records before they are admitted as evidence.[12] Although the judgment focused on compliance with evidentiary requirements under the law relating to electronic records, its reasoning assumes renewed significance in an era where AI-generated videos and audio recordings may be virtually indistinguishable from genuine material. As deepfake technology continues to improve, courts may face increasing difficulty in assessing the authenticity of digital evidence, thereby placing greater reliance on forensic verification and expert testimony.

Legal scholars have similarly recognised that deepfakes pose challenges extending beyond conventional cybercrime. Robert Chesney and Danielle Keats Citron argue that deepfakes threaten democratic institutions, individual privacy, and national security because they erode public confidence in the authenticity of digital information.[13] They further observe that the existence of convincing fabricated media creates what has been described as the "liar's dividend," whereby genuine evidence may be dismissed as fake, allowing wrongdoers to evade accountability. Danielle Keats Citron likewise emphasises that existing legal remedies often respond only after substantial harm has already occurred, highlighting the need for preventive regulatory measures capable of addressing AI-generated deception before it causes widespread damage.[14]

[12] *Anvar P.V. v. P.K. Basheer*, (2014) 10 SCC 473.
[13] Robert Chesney & Danielle Keats Citron, Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security, 107 *Calif. L. Rev.* 1753 (2019).
[14] Danielle Keats Citron, *The Fight for Privacy: Protecting Dignity, Identity, and Love in the Digital Age* (W.W. Norton & Co. 2022).

Comparative legal developments provide valuable guidance for addressing these concerns. The European Union has adopted a comprehensive risk-based approach through the Artificial Intelligence Act, which imposes transparency obligations on providers of AI systems that generate or manipulate synthetic content.[15] Rather than prohibiting deepfakes outright, the legislation generally requires appropriate disclosure so that users are informed when content has been artificially generated or altered, while imposing stricter compliance obligations on high-risk AI systems. This approach recognises that regulation should focus on transparency and accountability rather than restricting technological innovation itself.

The United Kingdom has adopted a different regulatory strategy by favouring a principles-based framework that encourages existing sectoral regulators to oversee AI within their respective fields while emphasising safety, accountability, fairness, transparency, and contestability.[16] Singapore's Model AI Governance Framework similarly promotes responsible AI development through organisational accountability, explainability, human oversight, and effective risk management without creating an entirely separate legal regime for artificial intelligence.[17]

These comparative approaches demonstrate that effective regulation of deepfakes requires more than traditional criminal sanctions or intermediary obligations. They prioritise preventive governance through transparency requirements, institutional oversight, technical standards, and clearly defined responsibilities for AI developers and digital platforms. By contrast, India's existing cyber law framework remains largely reactive, addressing deepfake-related harms only after they materialise. While judicial decisions such as *Shreya Singhal*, *Puttaswamy*, and *Anvar P.V.* provide important constitutional and legal principles, they cannot by themselves resolve the complex issues of attribution, platform responsibility, and evidentiary reliability created by rapidly evolving deepfake technology. The comparative experience therefore suggests that India should supplement its existing legal framework with AI-specific safeguards capable of addressing the unique risks posed by synthetic media while preserving innovation and freedom of expression.

[15] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act), 2024 O.J. (L).
[16] Department for Science, Innovation and Technology, *A Pro-Innovation Approach to AI Regulation* (U.K. Gov't 2023).
[17] Infocomm Media Development Authority & Personal Data Protection Commission Singapore, *Model AI Governance Framework* (2d ed. 2020).

5. Recommendations

The existing legal framework provides remedies for many of the harms caused by deepfakes but does not adequately regulate the technology itself. Rather than replacing the current framework, India should adopt targeted reforms that strengthen accountability while preserving innovation and freedom of expression.

To begin with, the Information Technology Act, 2000 should be amended to expressly recognise AI-generated synthetic media and define the legal responsibilities of creators, developers, and digital platforms involved in its dissemination. A statutory definition of "deepfake" would reduce ambiguity and ensure consistent enforcement across different forms of AI-generated content.

The due diligence obligations of intermediaries should also be strengthened by requiring large digital platforms to implement reasonable mechanisms for detecting, labelling, and expeditiously removing malicious deepfake content upon verification. Such measures should be accompanied by transparent grievance redressal procedures to ensure timely relief for affected individuals.

Given the increasing misuse of deepfakes for identity theft, financial fraud, and non-consensual intimate imagery, specialised forensic capabilities should be developed within cybercrime investigation agencies to improve the detection and authentication of AI-generated content. Regular collaboration between government agencies, technology companies, and academic institutions would further support the development of reliable detection tools and technical standards.

Finally, India should formulate a comprehensive regulatory framework for artificial intelligence that complements existing cyber laws. Such a framework should emphasise transparency, accountability, human oversight, and risk-based regulation while safeguarding constitutional values, including privacy and freedom of expression.

6. Conclusion

Deepfake technology presents one of the most significant legal challenges emerging from recent advances in artificial intelligence. While the Information Technology Act, 2000, the Digital Personal Data Protection Act, 2023, the Information Technology Rules, 2021, and the

Bharatiya Nyaya Sanhita, 2023 collectively provide remedies for certain forms of digital harm, they do not comprehensively regulate AI-generated synthetic media. Existing legal provisions primarily respond to the consequences of deepfakes rather than addressing the unique issues of attribution, platform accountability, and evidentiary reliability created by the technology. As deepfakes become increasingly sophisticated and accessible, India's cyber law framework must evolve beyond reactive enforcement towards preventive regulation. Carefully designed legislative reforms, supported by institutional oversight and technological safeguards, can strengthen public trust while ensuring that innovation in artificial intelligence develops within a framework of legal accountability and respect for fundamental rights.

Bibliography

A. Statutes and Legislation

Bharatiya Nyaya Sanhita, No. 45 of 2023 (India).

Digital Personal Data Protection Act, No. 22 of 2023 (India).

Information Technology Act, No. 21 of 2000 (India).

Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, G.S.R. 139(E), Gazette of India, Extraordinary, pt. II, sec. 3(i), Feb. 25, 2021 (India).

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act), 2024 O.J. (L).

B. Cases

Anvar P.V. v. P.K. Basheer, (2014) 10 SCC 473 (India).

Justice K.S. Puttaswamy (Retd.) v. Union of India, (2017) 10 SCC 1 (India).

Shreya Singhal v. Union of India, (2015) 5 SCC 1 (India).

C. Books

Citron, Danielle Keats. *The Fight for Privacy: Protecting Dignity, Identity, and Love in the Digital Age*. W.W. Norton & Company, 2022.

D. Journal Articles

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

E. Reports and Policy Documents

Department for Science, Innovation and Technology (United Kingdom). *A Pro-Innovation*

Approach to AI Regulation. 2023.

Henry Ajder. *The State of Deepfakes: Landscape, Threats, and Impact*. Deeptrace Labs, 2019.

Infocomm Media Development Authority & Personal Data Protection Commission Singapore. *Model AI Governance Framework*. 2d ed. 2020.

NITI Aayog. *National Strategy for Artificial Intelligence: #AIForAll*. Government of India, 2018.

Organisation for Economic Co-operation and Development (OECD). *OECD Principles on Artificial Intelligence*. 2019.

UNESCO. *Recommendation on the Ethics of Artificial Intelligence*. 2021.