The Indian Journal for Research in Law and Management

Open Access Law Journal – Copyright © 2026
Editor-in-Chief – Dr. Muktai Deb Chavan; Publisher – Alden Vas; ISSN: 2583-9896

LEGAL ASPECTS OF DEEPFAKES AND AI-GENERATED CONTENT IN INDIA: CHALLENGES, LIABILITY AND THE NEED FOR REGULATORY REFORM

Samikhya Patro

ABSTRACT

Artificial intelligence has significantly transformed digital content creation, giving rise to deepfakes and other forms of AI-generated content capable of replicating human identity with remarkable accuracy. While these technologies offer legitimate benefits in sectors such as entertainment, education, and communication, they also present serious legal challenges relating to privacy, reputation, electoral integrity, and platform accountability. India currently lacks a dedicated legal framework regulating deepfakes and instead relies on a combination of cyber, criminal, data protection, and constitutional laws. This article examines the legal implications of deepfakes in India, evaluates the adequacy of existing regulatory mechanisms, analyzes the complex issue of liability among creators, platforms, and AI developers, and draws insights from international approaches adopted by the European Union, China, and the United States. The article argues that the present legal framework remains fragmented and reactive, necessitating a comprehensive and rights-based regulatory response. It concludes by proposing reforms that balance technological innovation with the protection of privacy, dignity, democratic values, and digital trust.

KEYWORDS

Deepfakes; Artificial Intelligence; AI-Generated Content; Privacy; Digital Identity; Liability; Cyber Law; Data Protection; Platform Accountability; Regulatory Reform.

INTRODUCTION

Artificial intelligence has made synthetic media startlingly realistic. Deepfakes and other AI-generated content can now imitate a person's face, voice, writing style, or identity with a degree of precision that blurs the line between truth and fabrication. In India, this creates a serious legal problem: the harm caused by deepfakes is immediate and scalable, but the legal response remains fragmented.

India does not yet have a stand-alone statute dedicated to deepfakes. Instead, the current response relies on a patchwork of the Information Technology Act, 2000, the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 as amended, the Digital Personal Data Protection Act, 2023, the Bharatiya Nyaya Sanhita, 2023, and general civil and criminal principles such as defamation, privacy, and unfair exploitation of personality.[1][2][3][4]

Although existing Indian laws can address some harms caused by deepfakes, they were drafted in a period when such technologies were largely unimaginable. As deepfakes become increasingly realistic and accessible, courts and regulators are often required to stretch traditional legal concepts to address challenges that were never specifically anticipated by lawmakers.

Indian scholarship on artificial intelligence has traditionally focused on privacy, data governance, and intermediary liability. Comparatively less attention has been paid to deepfakes as a distinct legal problem, despite their growing social and political impact. Against this background, the present article examines deepfakes as a distinct category of digital harm and evaluates whether existing Indian laws are capable of responding effectively to the challenges they create.

THE EMERGING THREAT

Deepfakes are not confined to entertainment or harmless parody. They can be used to create non-consensual sexual imagery, impersonate public figures, manipulate elections, commit financial fraud, damage reputations, and distort evidence in legal or investigative settings. The problem extends beyond simple misinformation because deepfakes directly exploit an individual's identity and likeness while often spreading faster than any subsequent correction.

Recent controversies involving manipulated content featuring Indian public figures have shown how easily synthetic media can blur the distinction between reality and fabrication. Even when such content

[1] Information Technology Act, No. 21 of 2000, India Code. https://www.indiacode.nic.in
[2] Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, Ministry of Electronics & Information Technology, https://www.meity.gov.in.
[3] Digital Personal Data Protection Act, No. 22 of 2023, India Code, https://www.indiacode.nic.in/handle/123456789/22037?view_type=browse.
[4] Bharatiya Nyaya Sanhita, No. 45 of 2023, India Code, https://www.indiacode.nic.in.

is later debunked, the initial impact on public perception can be significant. The growing prevalence of these cases highlights the increasing difficulty of distinguishing authentic content from manipulated material, thereby increasing the risk of misinformation, reputational harm, and public confusion.

What makes election-related deepfakes particularly concerning is that they do not operate like ordinary rumours or misleading posts. A fabricated video can appear so convincing that many viewers may accept it as genuine before any fact-checking process has a chance to intervene. Unlike conventional misinformation, synthetic media derives credibility from visual realism, making it significantly harder for voters to distinguish truth from manipulation. In a country such as India, where political information spreads rapidly through digital platforms, the misuse of deepfakes during elections may undermine informed democratic participation and public trust in electoral processes.

The Indian government has expressly acknowledged these risks. MeitY's 2023 advisory[5]to intermediaries stated that the existing IT Rules must be applied to misinformation, false or misleading content, and material that impersonates others, including deepfakes. Parliamentary responses in 2025 further confirmed[6] that the government views deepfakes as threats to dignity, reputation, privacy, and platform accountability, while maintaining that current laws are technology-neutral and therefore applicable to AI-generated harms.

EXISTING LEGAL FRAMEWORK IN INDIA

The first layer of regulation is the IT Act, 2000. Where a deepfake involves identity theft, impersonation, violation of privacy, or the publication of obscene or sexually explicit material, Sections 66C, 66D, 66E, 67, and 67A may become relevant. Section 69A also provides a mechanism for blocking access to unlawful content, and Section 79 becomes important in determining when intermediaries may seek safe-harbour protection.

The second layer is the IT Rules, 2021. Rule 3(1)(b), as described in the MeitY advisory, requires intermediaries to clearly prohibit users from uploading or sharing content that is false, misleading, invasive of privacy, obscene, or impersonatory. The advisory further emphasized that platforms must communicate these prohibitions clearly to users, issue reminders, and act with due diligence in relation to unlawful content.

[5] Ministry of Electronics & Information Technology, Advisory on Deepfakes and Compliance with IT Rules (2023), https://www.meity.gov.in.
[6] Parliamentary Responses on Deepfakes and Artificial Intelligence Regulation, Lok Sabha Secretariat (2025).

The third layer is the DPDP Act, 2023. A deepfake often depends upon the use of personal data such as facial images, voice samples, or biometric-like identifiers. Government statements in Parliament have indicated that where personal data is used without consent, the DPDP framework may be attracted, especially in relation to lawful processing and reasonable safeguards.

The fourth layer is the Bharatiya Nyaya Sanhita, 2023. Parliamentary material indicates that false or misleading reports capable of causing public mischief may trigger penal consequences, and organised cybercrime involving deepfake content may also be prosecuted under the BNS. In addition, conventional criminal and civil actions such as defamation, cheating, intimidation, and reputational harm may operate depending on the facts of the case.

CONSTITUTIONAL DIMENSIONS OF DEEPFAKES IN INDIA

The constitutional implications of deepfakes extend beyond ordinary cyber harms. In *Justice K.S. Puttaswamy v. Union of India*[7], the Supreme Court recognized privacy as a fundamental right under Article 21 and emphasized the importance of informational self-determination. Deepfakes frequently involve the unauthorized use of an individual's facial image, voice, and personal data, thereby threatening privacy, dignity, and autonomy.

Similarly, in *R. Rajagopal v. State of Tamil Nadu*[8], the Supreme Court acknowledged an individual's right to control the publication of personal information. The unauthorized creation of AI-generated content that appropriates a person's identity is inconsistent with the principles articulated in Rajagopal.

Reputational harm presents an additional concern. In *Subramanian Swamy v. Union of India*[9], the Court upheld criminal defamation and recognized reputation as an integral aspect of dignity protected under Article 21. Deepfakes capable of depicting individuals in false and damaging situations can cause serious reputational injury.

However, regulatory responses must also respect freedom of expression. In *Shreya Singhal v. Union of India*[10], the Supreme Court emphasized the importance of protecting online speech and struck down Section 66A of the Information Technology Act. Any future regulation of deepfakes must therefore carefully balance the need to prevent harm with constitutional guarantees of free expression.

[7] Justice K.S. Puttaswamy (Retd.) v. Union of India, (2017) 10 SCC 1.
[8] R. Rajagopal v. State of Tamil Nadu, (1994) 6 SCC 632.
[9] Subramanian Swamy v. Union of India, (2016) 7 SCC 221.
[10] Shreya Singhal v. Union of India, (2015) 5 SCC 1

THE LIABILITY PROBLEM

The original creator may generate the synthetic content; a second actor may edit or package it deceptively; a third may upload it; and a platform may amplify it through recommendation systems. In many cases, the greatest harm is caused not at the moment of creation but at the point of virality.

Perhaps the most difficult issue is determining who should ultimately bear responsibility when a harmful deepfake causes damage. In practice, a harmful deepfake may pass through several hands before reaching the public, making it difficult to identify where legal accountability should begin and end. While the creator of a deepfake may be the primary wrongdoer, the eventual harm is often amplified by distributors, platform algorithms, and in some cases the design choices of AI developers themselves. Yet Indian law still addresses these wrongs indirectly, by fitting them into older categories such as impersonation, obscenity, or misinformation, rather than recognising synthetic identity manipulation as a distinct legal harm.

For intermediaries, the debate turns on due diligence and safe harbour. India's current approach expects platforms not to host or facilitate unlawful content once it falls within prohibited categories, and the 2023 advisory signalled a stricter compliance stance in relation to deepfakes. However, the law still does not provide a detailed statutory protocol for high-risk AI content, mandatory provenance disclosure, or time-bound takedown duties tailored specifically to deepfake harm.

For AI developers and model deployers, the law is even less clear. If a tool is designed or released without adequate guardrails, watermarking, abuse prevention, or complaint channels, the resulting injury may be socially foreseeable, but legal attribution remains uncertain. This is one of the biggest weaknesses in the present framework: it focuses on unlawful content after publication more than systemic responsibility before deployment.

WHY THE PRESENT FRAMEWORK IS INADEQUATE

A closer examination of the existing framework reveals a number of practical weaknesses that may limit its effectiveness in addressing deepfake-related harms. One major concern is the absence of a statutory definition of deepfakes or AI-generated impersonation. This creates uncertainty for both enforcement agencies and courts.

Another concern relates to the speed of available remedies. In practical terms, the effectiveness of a legal remedy often depends on timing. A manipulated video that remains online for even a short period may be copied, downloaded, and redistributed across multiple platforms. By the time legal action is

initiated, the reputational or emotional harm suffered by the victim may already be difficult to reverse. Deepfake-related harm often spreads within hours, whereas legal processes typically operate at a much slower pace. The current legal structure offers blocking, complaint, and criminal processes, but it does not yet provide a specialised rapid-response system for synthetic sexual abuse, political manipulation, financial fraud, or viral reputational attacks.

Responsibility for deepfake-related harm remains difficult to allocate. The law does not clearly distinguish between malicious creators, negligent deployers, compliant intermediaries, and platforms that algorithmically boost harmful content after notice. A modern regulatory model must distribute responsibility according to role, knowledge, capacity to prevent harm, and speed of response.

Practical evidentiary difficulties further complicate enforcement. Deepfake cases raise practical questions about authenticity, forensic detection, chain of custody, burden of proof, and preservation of metadata. A legal system that treats synthetic media as ordinary digital content may struggle when the central issue is not merely publication, but the manipulation of identity and authenticity at scale.

These shortcomings indicate that existing legal provisions alone may be insufficient to address the unique harms associated with deepfakes. Instead, they require a coherent legal framework specifically designed to regulate synthetic media and AI-enabled identity manipulation.

COMPARATIVE REGULATORY APPROACHES

Similar concerns have emerged in several other jurisdictions, leading governments to experiment with different regulatory approaches.

Among recent international developments, the European Union's AI Act represents one of the most significant attempts to regulate artificial intelligence through a comprehensive legal framework.[11] The legislation adopts a risk-based approach, imposing stricter obligations on AI systems that present significant risks to individuals and society. In relation to deepfakes, the Act emphasizes transparency by requiring AI-generated or manipulated content to be clearly disclosed to users. This transparency obligation reflects the principle that individuals have a right to know when they are interacting with synthetic rather than authentic content. The objective is not to prohibit synthetic media altogether but to ensure that individuals are informed when they are interacting with AI-generated content. This

[11] 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).

approach seeks to balance technological innovation with the protection of fundamental rights such as privacy, dignity, and democratic participation.

China has adopted a more interventionist regulatory model through its Deep Synthesis Provisions.[12] These regulations require providers of deep synthesis services to clearly label AI-generated content and implement measures to prevent misuse. Service providers are obligated to verify user identities, maintain records, and establish mechanisms for detecting and addressing unlawful content. This approach places significant responsibility on technology providers to prevent harm before it occurs.

In the United States, there is currently no comprehensive federal legislation specifically regulating deepfakes. However, several states have enacted targeted laws addressing specific harms. States such as California[13] and Texas[14] have introduced measures prohibiting the dissemination of deceptive deepfake content during election periods, while others have criminalized the creation and distribution of non-consensual deepfake pornography. The American approach demonstrates a sector-specific strategy focused on protecting electoral integrity and individual privacy without imposing broad restrictions on AI innovation.

Although these jurisdictions have adopted different strategies, each has recognised that traditional legal tools are often inadequate when dealing with synthetic media. Their experiences suggest that transparency obligations and clearly allocated responsibilities are becoming central features of modern AI regulation. India can draw valuable lessons from these approaches while developing a framework tailored to its constitutional values, democratic institutions, and rapidly expanding digital ecosystem.

While India has attempted to address deepfake harms through existing cyber, criminal, and data protection laws, a purely technology-neutral approach may prove inadequate in the long term. Deepfakes differ fundamentally from traditional forms of misinformation because they weaponize authenticity itself. The harm arises not merely from the dissemination of false information but from the deliberate manipulation of identity, consent, and trust. Consequently, future regulation should move beyond content moderation and focus on safeguarding digital identity as an independent legal interest deserving of statutory protection.

THE CASE FOR REGULATORY REFORM

[12] Provisions on the Administration of Deep Synthesis Internet Information Services (China) (2023).
[13] CAL. ELEC. CODE § 20010 (West).
[14] TEX. ELEC. CODE ANN. § 255.004.

India should move toward a dedicated statutory or quasi-statutory framework for deepfakes and high-risk AI-generated content. That reform need not prohibit innovation; rather, it should separate legitimate synthetic creativity from deceptive, exploitative, and manipulative uses.

A strong reform package should include five core elements. First, a statutory definition of deepfakes and AI-generated impersonation, with special categories for sexual content, election-related deception, fraud, and child safety harms. Second, a consent-and-disclosure rule requiring clear labelling or provenance markers for synthetic content in specified contexts. Third, graded liability for creators, deployers, and intermediaries based on intent, negligence, and non-compliance after notice. Fourth, a rapid takedown and emergency relief mechanism for victims. Fifth, preservation, audit, and traceability obligations for high-risk generative AI systems.

India should also consider a personality-rights based remedy structure. Deepfakes frequently misuse a person's face, voice, and identity capital, even where traditional copyright claims are unavailable. A clearer statutory recognition of image, voice, and likeness misuse would help bridge the current gap between privacy, defamation, and data protection.

At the institutional level, reform should encourage coordination among MeitY, data protection authorities, law-enforcement agencies, cyber forensics units, and courts. A rights-based and technologically informed framework would be more effective than relying on scattered post-hoc enforcement across unrelated statutes.

CONCLUSION

India is not operating in a legal vacuum with respect to deepfakes. Existing cyber, criminal, and privacy laws provide certain remedies. The difficulty, however, is that these remedies were developed for earlier forms of digital harm and therefore do not always fit the unique challenges posed by synthetic media.

While existing statutes provide partial remedies, they remain reactive rather than preventive. Deepfakes challenge not only privacy and reputation but also the very foundations of trust upon which democratic societies depend. Deepfakes raise concerns that extend beyond technology and increasingly engage constitutional values such as privacy, dignity, and freedom of expression. When artificial intelligence can convincingly manufacture speech, identity, and evidence, the law must evolve to preserve autonomy, dignity, and democratic accountability. Deepfakes have transformed what was once a technological concern into a legal and constitutional challenge. As synthetic media

becomes increasingly accessible, the ability of legal institutions to respond quickly and effectively will play a crucial role in maintaining public trust in digital information. The real challenge is no longer recognising the risks posed by deepfakes, but ensuring that the law evolves quickly enough to address them. In the age of synthetic media, preserving trust may become one of the most important objectives of modern law.