REGULATING DEEPFAKES IN INDIA: A CRITICAL ANALYSIS OF CRIMINAL LIABILITY, COPYRIGHT PROTECTION, AND DIGITAL RIGHTS IN THE AGE OF ARTIFICIAL INTELLIGENCE

Kunal Aswani, Symbiosis Law School, Noida

ABSTRACT

Deepfakes—synthetic media generated through artificial intelligence that convincingly alter or fabricate a person's appearance, voice or actions—have moved from technical curiosity to a pervasive social and legal problem. In India the response has been incremental rather than comprehensive. Prosecutors currently rely on a patchwork of provisions under the Information Technology Act, 2000 and the Bharatiya Nyaya Sanhita, 2023; copyright and personality rights litigation has filled some of the gaps left by statute; and the Digital Personal Data Protection Act, 2023 together with recent amendments to the IT Rules, 2021 attempt to regulate the platforms that host such content. This paper examines these overlapping regimes, identifies structural weaknesses in the existing framework, and argues that the absence of a coherent digital identity right, rather than the mere lack of a standalone deepfake offence, constitutes the central legislative deficiency. Drawing on recent High Court jurisprudence, comparative developments and enforcement realities, the analysis proposes a calibrated reform agenda that strengthens criminal sanctions where harm is severe while preserving space for legitimate satire, research and artistic expression.

Keywords: Deepfakes, Artificial Intelligence, Bharatiya Nyaya Sanhita, Information Technology Act, Copyright, Personality Rights, Digital Personal Data Protection Act, Intermediary Liability.

I. Introduction

The capacity of generative artificial intelligence to produce photorealistic images, videos and audio recordings of real persons has outpaced the ability of most legal systems to respond with clarity. India is no exception. While deepfake technology can be deployed for creative, educational or satirical purposes, its misuse for non-consensual intimate imagery, financial fraud, electoral manipulation and reputational harm has become sufficiently frequent to demand sustained legal attention.

What distinguishes the Indian regulatory landscape is not the absence of law but its fragmentation. Provisions scattered across the Information Technology Act, 2000,[1] the Bharatiya Nyaya Sanhita, 2023,[2] the Copyright Act, 1957 and the Digital Personal Data Protection Act, 2023,[3] together with evolving intermediary obligations under the IT Rules,[4] create a web of potential liability. Courts, particularly the Delhi and Bombay High Courts, have further expanded the protective reach of personality rights in the absence of a dedicated statute.[5] The result is a regime that is both flexible and unpredictable—capable of addressing many harms yet leaving victims, platforms and creators uncertain about the precise boundaries of liability.

This paper undertakes a critical examination of the three principal axes of regulation: criminal liability, copyright and personality rights protection, and the broader architecture of digital rights and platform accountability. It argues that while the existing toolkit is more robust than is sometimes acknowledged, its effectiveness is undermined by conceptual mismatches, evidentiary difficulties and the absence of a unified right to digital identity. The analysis concludes with a set of reform proposals that seek to close the most serious gaps without over-criminalising speech or innovation.

II. Conceptualising Deepfakes and the Nature of the Harm

Deepfakes are commonly understood as media generated or substantially altered by

[1]Information Technology Act, 2000, ss. 66C, 66D, 66E, 67, 67A.
[2]Bharatiya Nyaya Sanhita, 2023, ss. 318, 319, 336, 340, 353, 356.
[3]Digital Personal Data Protection Act, 2023, ss. 2(x), 11, 12.
[4]Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 (as amended 2025–2026).
[5]Anil Kapoor v. Simply Life India & Ors., Delhi High Court (2023); Arijit Singh v. Codible Ventures LLP, Bombay High Court (2024).

machine-learning techniques so as to depict a person saying or doing something they did not. The technical methods—generative adversarial networks, diffusion models, voice-cloning systems—continue to improve, reducing the skill and cost required to produce convincing fabrications. The legal problem, however, is not primarily technical. It is the combination of three features that generates distinctive harm: the appropriation of a person's unique biometric identifiers (face, voice, mannerisms), the creation of a false representation that is difficult for ordinary viewers to detect, and the potential for rapid, widespread dissemination through digital platforms.

The harms fall into several overlapping categories. Non-consensual intimate deepfakes inflict profound psychological injury and social stigma, disproportionately affecting women. Financial fraud facilitated by voice or video cloning of company executives or family members has already produced substantial losses. Political deepfakes threaten the integrity of electoral discourse by fabricating statements that never occurred. Reputational damage, even when not rising to criminal defamation, can be severe and difficult to reverse once content circulates. Each of these harms engages different legal doctrines, which is both a strength and a source of incoherence.

III. Criminal Liability under the IT Act and the Bharatiya Nyaya Sanhita

A. The IT Act Framework

The Information Technology Act remains the primary statute for computer-related offences. Section 66C criminalises identity theft involving the fraudulent use of another person's unique identification feature, a formulation that can encompass facial or voice data extracted for deepfake generation. Section 66D targets cheating by personation through a computer resource, which covers many of the financial fraud scenarios involving synthetic audio or video. Section 66E addresses the capture or publication of images of a person's private area without consent, and Sections 67 and 67A prohibit the electronic transmission of obscene or sexually explicit material. Together these provisions supply the most frequently invoked charges in deepfake prosecutions involving intimate imagery or impersonation fraud.

Yet the fit is imperfect. Identity theft under Section 66C was drafted with passwords and digital signatures in mind; extending it to biometric likeness requires interpretive effort. Cheating by personation presupposes an intent to induce delivery of property or some other

tangible advantage, which may not capture pure reputational or political deepfakes. Obscenity provisions focus on the nature of the content rather than the violation of consent and identity that characterises non-consensual intimate deepfakes. Prosecutors therefore often charge multiple overlapping sections, producing uncertainty about the precise elements that must be proved.

B. The Bharatiya Nyaya Sanhita

The BNS, which replaced the Indian Penal Code, modernises several relevant offences but does not create a dedicated deepfake crime. Section 318 (cheating) and Section 319 (cheating by personation) can apply where synthetic media is used to induce a victim to part with property. Section 336 and related forgery provisions treat electronic records as documents and can, with some stretching, cover the creation of false synthetic media intended to deceive. Section 353 addresses the circulation of false information likely to cause public mischief, offering a potential route for political or panic-inducing deepfakes. Section 356 continues the offence of defamation.

The difficulty is that each of these provisions was designed for a different factual paradigm. Forgery traditionally involves the making of a false document; a deepfake is better understood as a false representation of a person's actions or words. Defamation focuses on reputation rather than the deeper autonomy and dignity interests implicated by the non-consensual appropriation of identity. Organised-crime provisions under Section 111 may capture large-scale deepfake fraud rings, but ordinary individual perpetrators fall outside their scope. The cumulative effect is that the BNS expands the available charges without resolving the conceptual mismatch between existing offences and the distinctive nature of synthetic-media harm.

C. Enforcement Realities

Even where the statutory language can be stretched to cover a deepfake, successful prosecution remains challenging. Attribution is difficult when content is uploaded anonymously or through encrypted channels. Intent to deceive or harm must usually be inferred from circumstantial evidence. Forensic examination of synthetic media requires specialised expertise that is unevenly distributed across police forces. Victims, particularly of intimate deepfakes, may be reluctant to pursue complaints because of the risk of secondary trauma.

These practical constraints mean that the formal availability of criminal sanctions does not translate automatically into effective deterrence or redress.

IV. Copyright Protection and the Emerging Jurisprudence on Personality Rights

A. Copyright Act Limitations

The Copyright Act, 1957 approaches the problem from a different direction. Where a deepfake incorporates substantial portions of a copyrighted film, performance or sound recording, the rights of the producer or performer are engaged. Performers' rights under Sections 38 and 38A, and moral rights under Section 57, have been invoked to restrain unauthorised voice cloning and synthetic recreations of performances.[6] The definition of author in Section 2(d)(vi) attributes computer-generated works to the person who causes the work to be created, a formulation that becomes ambiguous when the human contribution is limited to a short prompt.[7]

Copyright, however, protects expression, not identity as such. An ordinary individual whose face or voice is used without consent may have no copyright interest to assert. Even celebrities whose performances are protected find that copyright litigation is slow, expensive and focused on commercial exploitation rather than the full range of dignitary harms. Recent litigation concerning the training of generative models on copyrighted news content illustrates the further complication that arises when AI systems ingest protected material at scale.[8]

B. Judicial Construction of Personality Rights

In the absence of a statutory personality or publicity right, Indian courts have developed a body of case law that draws on Articles 19 and 21 of the Constitution, the residual categories of the Copyright Act, passing-off principles under trade-mark law, and the privacy jurisprudence inaugurated by R. Rajagopal[9] and consolidated in Puttaswamy.[10] High Courts have granted injunctions restraining the unauthorised commercial use of a celebrity's name, image, voice, mannerisms and catchphrases, expressly including AI-generated deepfakes and synthetic content. These orders have been framed as John Doe injunctions, enabling rapid

[6]Copyright Act, 1957, s. 2(d)(vi), ss. 38, 38A, 57.
[7]Eastern Book Company v. D.B. Modak, (2008) 1 SCC 1.
[8]ANI Media Pvt. Ltd. v. OpenAI Inc. & Anr., CS(COMM) 1028/2024, Delhi High Court.
[9]R. Rajagopal v. State of Tamil Nadu, (1994) 6 SCC 632.
[10]Justice K.S. Puttaswamy (Retd.) v. Union of India, (2017) 10 SCC 1.

takedown of unidentified uploaders and directing platforms to remove specified content.

The jurisprudence is significant but incomplete. It remains concentrated in a handful of High Courts and is most accessible to public figures with the resources to litigate. Ordinary victims of non-commercial intimate deepfakes or pure reputational fabrications still lack a clear, codified right of action. The absence of a statutory framework also leaves the precise contours of the right duration, exceptions for satire and news reporting, posthumous protection undefined, creating uncertainty for both claimants and content creators.

V. Digital Rights, Privacy and the Digital Personal Data Protection Act

Deepfakes almost invariably involve the processing of personal data, particularly biometric data in the form of facial images or voice recordings. The Digital Personal Data Protection Act, 2023 therefore supplies an independent regulatory layer. Processing requires a lawful basis, typically consent or one of the limited legitimate-use exceptions. Data principals enjoy rights of access, correction and erasure that can, in principle, be invoked to demand the removal of deepfake content that incorporates their personal data.[11]

The DPDP Act's potential remains only partially realised. The detailed rules necessary for effective enforcement are still being notified, and the interaction between data-protection remedies and the more immediate criminal and civil routes is not yet settled. Moreover, the Act is framed around the relationship between data fiduciaries and data principals; it does not directly address the distinctive injury caused by the creation of a false synthetic representation. Nevertheless, the recognition that biometric identifiers constitute personal data, and that their non-consensual use attracts significant penalties, strengthens the overall protective architecture.

VI. Intermediary Liability and the Regulation of Platforms

Because deepfakes achieve scale primarily through digital platforms, intermediary liability rules are central to any effective response. The IT Rules, 2021, as amended in 2025 and 2026, now expressly address "synthetically generated information" content artificially created or altered so as to appear authentic.[12] Significant social media intermediaries are

[11] Digital Personal Data Protection Act, 2023, ss. 2(x), 11, 12.
[12] Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 (as amended 2025–2026).

required to deploy technical measures to detect such content, to label it clearly, and to remove unlawful synthetic material within accelerated timelines, including as little as two hours for certain categories of non-consensual intimate imagery. Failure to comply risks the loss of safe-harbour protection under Section 79 of the IT Act.

These obligations represent a marked tightening of platform responsibility. They also raise familiar concerns about over-removal, the chilling of legitimate speech, and the practical capacity of platforms to distinguish harmful deepfakes from satire, parody or artistic work. The definition of synthetically generated information is deliberately broad, and the pressure to act quickly may incentivise conservative content moderation. Balancing the imperative of rapid removal against the risk of collateral censorship remains an unresolved tension in the current framework.

VII. Judicial Trends and Institutional Responses

Indian courts have not waited for comprehensive legislation. The Delhi High Court, in particular, has issued a series of injunctions protecting the personality rights of public figures against AI-generated misuse, recognising that deepfakes threaten dignity, livelihood and public trust in ways that traditional intellectual-property categories only partially capture. Parallel orders from the Bombay High Court in cases involving voice cloning of performing artists have reinforced the same protective impulse. These decisions demonstrate judicial willingness to adapt existing doctrines to new technological realities.

At the same time, the judiciary has been careful not to invent new criminal offences. The Supreme Court has observed that the creation of novel crimes is a legislative function, underscoring the limits of judicial law-making even while pressing the executive and legislature to address gaps in the statutory framework. The result is a division of labour in which courts provide interim civil protection and interpretive guidance, while the more fundamental questions of criminalisation and institutional design remain with Parliament and the executive.

VIII. Comparative Insights and Structural Gaps

Comparative experience offers both caution and inspiration. The European Union’s AI Act treats certain deepfake applications as high-risk and imposes transparency obligations,

while leaving much of the substantive harm regulation to national criminal and civil law.[13] China's deep-synthesis regulations impose licensing and labelling requirements on providers of synthetic-media services.[14] Several United States jurisdictions have enacted targeted statutes criminalising non-consensual intimate deepfakes or requiring disclosure of synthetic political content. None of these models is directly transplantable, yet each illustrates that effective regulation typically combines platform duties, clear criminal prohibitions for the most serious harms, and civil remedies for identity-based injury.

India's principal structural gaps may be summarised as follows. First, there is no standalone deepfake offence that captures the distinctive combination of identity appropriation and deceptive synthetic representation. Second, ordinary individuals lack a codified personality or digital-identity right that would support civil claims without recourse to expensive constitutional or intellectual-property litigation. Third, forensic and investigative capacity remains uneven, limiting the practical utility of existing criminal provisions. Fourth, the interaction among criminal, data-protection, copyright and intermediary regimes is insufficiently coordinated, producing both overlaps and blind spots.

IX. Towards a Coherent Framework: Recommendations

A coherent response need not begin with an entirely new statute, but it does require deliberate legislative and institutional design. The following measures would address the most pressing deficiencies while preserving necessary space for legitimate expression.

First, the Bharatiya Nyaya Sanhita should be amended to introduce a graduated deepfake offence. A baseline offence could target the creation or dissemination of synthetic media that is intended to deceive and that causes or is likely to cause harm to reputation, privacy or financial interests. Aggravated forms—non-consensual intimate imagery, electoral manipulation, large-scale fraud—should attract higher penalties. Clear mens rea requirements and carefully drafted exceptions for satire, research, journalism and artistic expression would mitigate the risk of overbreadth.

Second, Parliament should consider enacting a statutory digital identity or personality right available to every individual, not only celebrities. Such a right would confer control over

[13]European Union Artificial Intelligence Act, Regulation (EU) 2024/1689.
[14]China's Provisions on the Administration of Deep Synthesis of Internet Information Services (2022).

the commercial and non-commercial use of one's name, image, voice and other distinctive attributes, subject to well-defined exceptions for public-interest speech. Codification would reduce reliance on judicial improvisation and make remedies more accessible.

Third, enforcement capacity must be strengthened. Specialised digital forensic units, standardised protocols for the examination of synthetic media, and training for investigating officers and prosecutors are essential if the formal availability of sanctions is to translate into actual deterrence.

Fourth, platform obligations under the IT Rules should be refined to require not only rapid removal of clearly unlawful content but also transparent reporting on detection methods, error rates and the volume of legitimate content that is mistakenly restricted. Independent oversight of these processes would help maintain public confidence.

Finally, coordination among the ministries responsible for information technology, home affairs, law and justice, and data protection is necessary to ensure that criminal, civil, copyright and privacy regimes operate as a coherent whole rather than as parallel and sometimes contradictory systems.

X. Conclusion

Deepfakes expose the limits of a legal architecture designed for an earlier technological era. India's response—an evolving combination of adapted criminal provisions, judicially expanded personality rights, data-protection rules and intensified platform duties—has achieved partial coverage of the most serious harms. Yet the absence of a unified conceptual foundation leaves the regime vulnerable to both under-enforcement and over-reach.

The path forward lies neither in wholesale criminalisation of synthetic media nor in continued reliance on interpretive stretch. It lies in the deliberate construction of a digital identity right, the careful calibration of criminal sanctions to the severity of harm, the professionalisation of forensic and investigative capacity, and the refinement of platform accountability mechanisms. If these elements can be brought into alignment, India will be better positioned to protect individual dignity and democratic discourse without sacrificing the innovative and expressive potential of artificial intelligence.

References

1. Ministry of Electronics and Information Technology, Advisories on Deepfakes and Synthetically Generated Information (2023–2026).

2. Information Technology Act, 2000 (as amended).

3. Bharatiya Nyaya Sanhita, 2023.

4. Digital Personal Data Protection Act, 2023.

5. Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 (as amended 2025–2026).

6. Anil Kapoor v. Simply Life India & Ors. (Delhi High Court, 2023); Arijit Singh v. Codible Ventures LLP (Bombay High Court, 2024); related personality-rights injunctions.

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

8. Copyright Act, 1957, especially ss. 2(d)(vi), 38, 38A and 57.

9. Eastern Book Company v. D.B. Modak, (2008) 1 SCC 1.

10. ANI Media Pvt. Ltd. v. OpenAI Inc. & Anr., CS(COMM) 1028/2024 (Delhi High Court).

11. R. Rajagopal v. State of Tamil Nadu, (1994) 6 SCC 632.

12. European Union Artificial Intelligence Act, Regulation (EU) 2024/1689.

13. Provisions on the Administration of Deep Synthesis of Internet Information Services (China, 2022).