

SUPREMO AMICUS

INDIA'S FIRST AI INTEGRATED LAW JOURNAL

Peer Reviewed, Refereed and Open access Journal

- Available in 331+ International Libraries
- Indexed at 32 Databases



ISSN NO. 2456-9704
Volume 10 Issue 3
www.supremoamicus.org



DISCLAIMER

The information presented in this article is intended for general informational and educational purposes only. While every effort has been made to ensure that the content is accurate, up-to-date, and reliable at the time of publication, the editorial board and publisher make no representations or warranties of any kind, express or implied, regarding the completeness, accuracy, reliability, suitability, or availability of the information contained herein.

The views and opinions expressed in this article are those of the author and are based on personal research, experience, and interpretation. They do not necessarily reflect the official policy, position, or opinions of any affiliated organization, institution, or entity.

This article is not intended to serve as professional advice of any kind. The editorial board and publisher shall not be held liable for any errors or omissions in the content, nor for any losses, injuries, or damages arising from the use of or reliance on this information.



ABOUT THE JOURNAL

Supremo Amicus is an online, peer-reviewed international journal devoted to the interdisciplinary fields of law and science. In an era marked by rapid technological progress and evolving legal frameworks, the journal seeks to bridge the gap between these two dynamic domains by offering comprehensive and critical insights into their various aspects. The journal places a strong emphasis on contemporary advancements, emerging trends, and the complex challenges faced by both the legal community.

The primary objective of the journal is to encourage and promote original, high-quality research. It is committed to publishing well-researched, analytically sound, and thought-provoking articles that adhere to rigorous academic standards. Each submission undergoes a thorough peer-review process to ensure authenticity, relevance, and scholarly integrity. In doing so, the journal maintains its commitment to excellence and credibility.

In addition to fostering research, the journal aims to make complex ideas accessible and engaging for a diverse readership. It strives to present content that is not only intellectually enriching but also clearly written and reader friendly.

Furthermore, the journal is committed to promoting interdisciplinary collaboration and global engagement. It welcomes diverse perspectives from contributors across different regions and backgrounds, thereby enriching the quality and scope of discussions presented within its pages.

With this vision we proudly present Supremo Amicus to our readers.

**-Editorial Team
Supremo Amicus**



DEEPFAKES AND THE LAW IN INDIA: REGULATORY GAPS AND THE WAY FORWARD

By Anchal Preet Kaur

*From Model Institute of Engineering and
Technology, Jammu and Kashmir*

Abstract

The rapid proliferation of deepfake technology has exposed significant gaps in India's legal architecture, prompting a series of regulatory interventions between 2023 and 2026. This article examines whether India's existing legal framework comprising the Information Technology Act, 2000, the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2025–26, the Bharatiya Nyaya Sanhita, 2023, and the Digital Personal Data Protection Act, 2023 adequately addresses the harms caused by deepfakes and synthetically generated information, or whether a dedicated standalone legislative framework remains necessary. Through doctrinal analysis of the applicable statutory provisions, examination of recent judicial developments including the Delhi High Court's landmark personality rights order in *Nandamuri Taraka Rama Rao v. Ashok Kumar/John Doe & Ors.* (2025), and comparative reference to the European Union's AI Act and the United States' TAKE IT DOWN Act, this article concludes that while India's framework has advanced considerably, it remains substantively incomplete — particularly in its failure to provide a direct, victim-initiated statutory remedy accessible to ordinary, non-celebrity individuals. A consolidated Synthetic Media Act is proposed as the most durable solution.

Research and drafting for this article was conducted with the assistance of AI tools.

Keywords: Deepfakes; Synthetically Generated Information; IT Rules 2026; Personality Rights; Bharatiya Nyaya Sanhita; AI Regulation; India; Synthetic Media Act.

INTRODUCTION

In November 2023, a manipulated video of actor Rashmika Mandanna began circulating widely across social media platforms¹. The video had digitally superimposed her face onto the body of a British influencer, Zara Patel, using generative adversarial network (GAN) technology a manipulation so seamless that millions of viewers initially believed it to be authentic. Mandanna herself described the experience as "extremely scary," noting that the incident exposed how vulnerable individuals have become to the misuse of emerging technology². The case did not stand alone; celebrities including Kajol, Alia Bhatt, and cricketer Sachin Tendulkar have all faced similar manipulated content in recent years, signaling that deepfake technology has moved from a novelty to a genuine and recurring legal harm³.

A deepfake is synthetic media image, audio, or video generated or altered using artificial intelligence to convincingly depict a person saying or doing something they never actually said or did. What was once a niche technical curiosity is now widely accessible, requiring little more than a smartphone application and a handful of source images. The consequences extend well beyond celebrity reputation: deepfakes have been used to spread political disinformation, perpetrate financial fraud, fabricate non-consensual sexual content, and erode public trust in audio-visual evidence more broadly. This is not a marginal concern the Entrust Cybersecurity Institute's 2025 Identity Fraud Report found that a deepfake identity attack occurred somewhere in the world once every five minutes in 2024, with deepfakes now accounting for roughly 40% of all video-based biometric fraud attempts

¹ Main accused arrested in Rashmika Mandanna deepfake video case," The Quint, 20 January 2024.

² Ibid.

³ Ibid.



globally.⁴

India's legal response has evolved considerably since the Mandanna incident. At the time, law enforcement was forced to rely on a patchwork of general-purpose provisions for forgery and defamation provisions under the Indian Penal Code, alongside Sections 66C and 66E of the Information Technology Act, 2000, which were never drafted with AI-generated synthetic media in mind⁵. By early 2026, the regulatory landscape shifted markedly: the Ministry of Electronics and Information Technology notified the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026, formally defining "synthetically generated information," imposing mandatory labelling obligations on platforms, and shortening takedown timelines to as little as three hours for harmful synthetic content.⁶ Courts, too, have begun stepping in independently; the Delhi High Court's December 2025 order protecting actor NTR Jr.'s personality rights against AI-generated misuse signals a judiciary willing to extend protection even ahead of clear legislative guidance.⁷

Yet the question of whether this evolving framework is genuinely adequate or merely a reactive patchwork stitched together in response to high-profile incidents remains open. This article examines whether India's existing legal framework, comprising the Information Technology Act, 2000, the Information Technology Rules (as amended in 2025–26), the Bharatiya Nyaya Sanhita, 2023, and the Digital Personal Data Protection Act, 2023, adequately addresses the harms caused by deepfakes, or whether the persistence of significant gaps in enforcement capacity, in the balance with free

speech, and in protection for victims who are not public figures necessitates the enactment of a dedicated, standalone legislative framework.

II The Existing Legal Framework

India does not yet have a single, dedicated statute governing deepfakes or synthetically generated content. Instead, protection is assembled from four distinct legal sources, each addressing a different facet of the harm. This section examines each in turn.

A. The Information Technology Act, 2000

The IT Act remains the primary statute invoked in deepfake-related prosecutions, despite having been drafted nearly two decades before generative AI became accessible to the public. Three provisions are routinely relied upon.

Section 66C (Identity Theft) penalizes the fraudulent or dishonest use of another person's electronic signature, password, or any other unique identification feature. This provision was explicitly invoked in the Rashmika Mandana FIR, where investigators applied it to the unauthorized digital replication of her likeness.⁸

Section 66E (Violation of Privacy) penalizes the capturing, publishing, or transmission of images of a person's private area without consent and carries a sentence of up to three years along with a fine of up to ₹2 lakh. This provision was also invoked in the Mandanna case, though its language is narrower than the harm deepfakes typically cause — it was designed for non-consensual *capture* of real images, not the

⁴ Entrust Cybersecurity Institute, 2025 Identity Fraud Report (with Onfido), November 2024.

⁵ Information Technology Act, 2000, Sections 66C, 66E; Indian Penal Code, 1860, Sections 465, 469 (as invoked in the Mandanna FIR; the IPC has since been replaced by the Bharatiya Nyaya Sanhita, 2023).

⁶ Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules,

2026, G.S.R. 120(E) (10 February 2026), Rule 2(1)(wa), Rule 3(1)(d), Rule 3(3)(a)(ii).

⁷ Nandamuri Taraka Rama Rao v. Ashok Kumar/John Doe & Ors., CS(COMM) 1305/2025 (Delhi HC, order dated 22 December 2025).

⁸ Information Technology Act, 2000, Section 66C; "Main accused arrested in Rashmika Mandanna deepfake video case," The Quint, 20 January 2024.



generation of entirely fabricated, AI-synthesized content where no actual image of the victim's body was ever taken.⁹

Section 66D (Cheating by Personation) penalizes cheating by impersonation using a computer resource, carrying a three-year sentence and a fine of up to ₹1 lakh.¹⁰

Sections 67, 67A, and 67B criminalize the publication or transmission of obscene and sexually explicit material in electronic form, including child sexual abuse material. These provisions become relevant where deepfakes are sexual or pornographic in nature, which empirically constitutes a significant share of non-consensual deepfake content globally.¹¹ The structural problem is evident: none of these provisions were drafted with synthetic media in mind. They address unauthorized use of *real* personal data or *real* captured images, not the creation of entirely fabricated content that never existed in the first place. Prosecutors have therefore been forced to apply these provisions through analogy rather than direct application.

B. The Information Technology Rules 2025 Draft and 2026 Amendment

Recognizing the gap left by the IT Act's outdated provisions, the Ministry of Electronics and Information Technology (**MeitY**) moved to directly regulate synthetic media through amendments to the IT (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021. This is the most significant and most current part of India's regulatory response, and it evolved through two distinct stages.

In October 2025, **MeitY** released a **draft amendment** to the IT Rules for public consultation. This draft proposed a fixed, quantitative labelling standard a watermark covering at least 10% of the surface area for images and video along with staggered takedown windows of 24 hours for sexual or explicit deepfakes and 36 hours for other categories of deepfake-driven misinformation.¹²

Following stakeholder feedback, the **Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026** were finally notified on 10 February 2026,¹³ vide Gazette Notification G.S.R. 120(E), and came into force from 20 February 2026. The final version differs from the draft in several material respects, generally tightening rather than loosening the earlier proposal:

First, a new provision inserted into the Rules defines "synthetically generated information" (SGI) as audio, visual, or audio-visual content that is artificially or algorithmically created or altered so as to appear authentic and be indistinguishable from a real person or event. This definition is technology-neutral; it turns on the *effect* of appearing real, not on the specific method or tool used to create the content.¹⁴

Second, on labelling, the fixed 10% watermarking rule proposed in the draft was dropped entirely. The final Rules instead require visual SGI to carry labels that are "prominent, easily noticeable, and adequately perceivable," while audio SGI must carry a prominently prefixed audio disclosure a qualitative standard rather than a fixed numerical one. Platforms are further barred from permitting the removal or

⁹ Information Technology Act, 2000, Section 66E.

¹⁰ Information Technology Act, 2000, Section 66D.

¹¹ Information Technology Act, 2000, Sections 67, 67A, and 67B.

¹² Ministry of Electronics and Information Technology, Draft Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, October 2025.

¹³ Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026, Gazette Notification G.S.R. 120(E), 10 February 2026.

¹⁴ Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026, Rule 2(1)(wa).



suppression of such labels once applied¹⁵.

Third, the takedown timelines were tightened, not loosened, between draft and final version. The 36-hour window proposed for general deepfake misinformation was shortened to **three hours** in the final Rules, with certain especially sensitive categories such as non-consensual intimate deepfakes subject to windows as short as two hours.¹⁶

Fourth, the final Rules introduced a specific safe-harbor clarification: removing or disabling synthetic content in compliance with these obligations including through the use of automated detection tools does not itself amount to a violation of an intermediary's safe harbor protection under Section 79(2) of the IT Act, 2000. This is a protective provision for platforms that act in compliance, rather than a penalty provision aimed at platforms that fail to act, although the broader due diligence framework under the Rules continues to expose non-compliant intermediaries to liability through other means.¹⁷

Finally, references to the Indian Penal Code within the Rules were updated to reflect the Bharatiya Nyaya Sanhita, 2023, bringing the Rules formally in line with India's current criminal code.¹⁸

Taken as a whole, this amendment represents one of the more aggressive platform-accountability regimes globally on this issue arguably exceeding the obligations imposed by the European Union's AI Act, which contains no deepfake-specific takedown timeline at all.

C. The Bharatiya Nyaya Sanhita, 2023

The Bharatiya Nyaya Sanhita, 2023, which replaced the Indian Penal Code with effect from 1 July 2024, retains and consolidates several provisions relevant to deepfakes, though as with the IT Act none were drafted with AI-generated synthetic media specifically in mind.

Section 319 (Cheating Personation) is the BNS's direct replacement for Section 416 of the former IPC. It penalizes deceiving another person by pretending to be someone else whether that person is real, deceased, or even imaginary and carries a sentence of imprisonment extending to five years, a fine, or both. Of the BNS provisions, this is the most directly relevant to deepfakes, since it squarely addresses the core harm of a synthetic depiction being used to impersonate a specific, identifiable individual.¹⁹

Section 353 (Statements Conducing to Public Mischief) replaces the former Section 505 IPC. It penalizes making, publishing, or circulating false statements, rumors, or reports explicitly including "through electronic means" with intent to cause public alarm, incite disregard of military duty, or provoke enmity between different groups, with punishment extending to three years (or five years where the offence is committed at a place of worship). This provision is relevant to politically disruptive or communally inflammatory deepfakes, though its scope is narrower than a general anti-misinformation provision would be it requires specific intent or likelihood tied to public order, mutiny, or communal enmity, rather than mere falsity of the content.²⁰

Separately, the BNS's general provisions on **criminal**

¹⁵ Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026, Rule 3(3)(a)(ii), Rule 3(3)(b).

¹⁶ Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026, Rule 3(1)(d).

¹⁷ Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026, Rule 2(1B).

¹⁸ Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026, Rule 7 (as amended).

¹⁹ Bharatiya Nyaya Sanhita, 2023, Section 319.

²⁰ Bharatiya Nyaya Sanhita, 2023, Section 353.



intimidation and defamation remain available where a deepfake is used to threaten an individual or damage their reputation, though these too apply only by extension rather than by direct design.²¹

D. The Digital Personal Data Protection Act, 2023

The DPDP Act becomes relevant only indirectly, where deepfake creation involves the unauthorized processing of biometric data.²² The often-cited ₹250 crore figure is the **maximum penalty under Section 8(5), enforced via Section 33**, for a Data Fiduciary's failure to implement *reasonable security safeguards* against a personal data breach²³ not a deepfake-specific penalty at all. It would apply only incidentally: for instance, where an entity holding someone's facial or voice biometric data negligently allowed it to be breached and subsequently used to train a deepfake model.²⁴ The DPDP Act does not penalize the *creation* of deepfakes themselves.

III. Judicial Response

While the legislature has only recently moved to address synthetic media directly, Indian courts particularly the Delhi High Court had already begun fashioning protection through existing common law and constitutional principles, often ahead of the 2026 IT Rules amendment discussed above. This judicial activity demonstrates both the judiciary's willingness to act swiftly in the face of legislative gaps and the practical limitations of case-by-case relief as a substitute for systemic regulation.

A. The NTR Jr. Case

On 22 December 2025, the Delhi High Court, in *Nandamuri Taraka Rama Rao v. Ashok Kumar/John Doe & Ors.*, granted an ex-parte ad-interim injunction in favor of Telugu actor Nandamuri Taraka Rama Rao Jr. ("NTR Jr."), restraining seventeen defendants including unidentified "John Doe" infringers, specific online sellers, and e-commerce platforms such as Amazon, Flipkart, and Meesho²⁵ from the unauthorized commercial exploitation of his name, image, likeness, voice, and associated identifiers, including the sobriquet "Man of Masses."

Justice Manmeet Pritam Singh Arora held that an individual's personality traits including name, likeness, and image are protectable elements of that individual's personality rights, entitling the plaintiff to seek an injunction against third parties using those attributes for commercial gain without authorization.²⁶ Significantly, the order was not confined to traditional merchandise infringement; the Court explicitly extended its protection to exploitation through artificial intelligence and deepfake technology, holding that such misuse whether for advertisements, promotional content, or AI-generated material would be impermissible regardless of the medium used to create it.²⁷ The plaintiff's case rested not only on common law personality rights under Articles 19(1)(a) and 21 of the Constitution, but also on registered trademark protections under the Trademarks Act, 1999, covering variations of his name and stage identifiers.²⁸

The Court directed the immediate takedown of identified infringing listings and content and granted relief operative until the next date of hearing, listed

²¹ Bharatiya Nyaya Sanhita, 2023 (general provisions on criminal intimidation and defamation).

²² Digital Personal Data Protection Act, 2023, Section 7 (consent requirements for processing personal data, including biometric data).

²³ Digital Personal Data Protection Act, 2023, Section 8(5), Section 33; The Schedule to the Digital Personal Data Protection Act, 2023, Item 1.

²⁴ Digital Personal Data Protection Act, 2023, Section 8(5).

²⁵ *Nandamuri Taraka Rama Rao v. Ashok Kumar/John Doe & Ors.*, CS(COMM) 1305/2025 (Delhi HC, order dated 22 December 2025).

²⁶ *Ibid.*

²⁷ *Ibid.*

²⁸ *Ibid.*



for 19 May 2026.²⁹

B. Significance and Limitations of the Judicial Approach

The NTR Jr. order is significant for two reasons. First, it confirms that Indian courts are willing to extend personality rights jurisprudence to explicitly cover AI-generated and deepfake misuse, even in the absence of a specific statutory provision addressing synthetic media at the time the order was passed, the 2026 IT Rules amendment had not yet been notified. Second, it forms part of a broader pattern: in the months surrounding this order, several other public figures including Hrithik Roshan, Aishwarya Rai Bachchan, Abhishek Bachchan, Karan Johar, and Sri Ravi Shankar independently approached the Delhi High Court seeking similar protection³⁰, suggesting that courts, rather than statute, were the primary forum of first resort for high-profile deepfake harms through late 2025.

This pattern, however, exposes the central limitation of relying on judicial intervention as a substitute for legislative or regulatory protection. Each of these orders required a plaintiff with the resources, recognition, and legal access necessary to mount a civil suit typically a public figure represented by senior counsel, capable of demonstrating significant commercial goodwill in their identity. An ordinary citizen victimized by a non-consensual deepfake, lacking comparable trademark registrations, commercial endorsements, or the means to retain comparable legal representation, would face a substantially higher barrier to obtaining equivalent relief through the same route. The judiciary's protection, in other words, has so far developed primarily as a shield for commercial personality rights rather than as a general remedy available to all victims

of synthetic media — a gap examined further in Section IV below.

IV. Critical Gaps

While the 2025–26 IT Rules amendment represents a significant regulatory advance, several structural gaps remain in enforcement capacity, in the constitutional balance with free speech, in technical reliability, and in protection for victims who are not public figures. Each is examined below.

A. Enforcement Capacity and the Burden on Smaller Platforms

The three-hour takedown window, while protective of victims, imposes a compliance burden that larger platforms are far better positioned to absorb than smaller ones. Meeting this deadline requires round-the-clock legal and moderation teams capable of acting within 180 minutes of a complaint, including during holidays and high-traffic periods a standard that industry commentary suggests may be genuinely difficult for smaller intermediaries to sustain.³¹ Regional and startup platforms, in particular, have flagged the cost of real-time content moderation as a serious operational burden, raising the prospect that compliance costs could be absorbed only by large, well-resourced technology companies, potentially leading to market consolidation and reduced competition within India's digital ecosystem.³²

B. Technical Limitations of Detection

The Rules presuppose those platforms can reliably distinguish synthetic content from genuine content, but detection technology has not kept pace with this assumption. Automated detection tools continue to struggle against high-quality deepfakes, and the same

²⁹ Ibid.

³⁰ Delhi High Court Protects NTR Jr's Personality Rights, Bans Unauthorized Use of Name, Image and 'Man of Masses' Identity," LawChakra, 29 December 2025.

³¹ "Three-Hour Deepfake Deadline Under IT Rules 2026," AI CERTs News, 2 March 2026.

³² India's new IT Rules 2026: Deepfake crackdown, mandatory AI labels, 3-hour takedown deadline explained," PTC News, 13 February 2026.



systems are prone to false positives flagging genuine satire, parody, and journalistic content as synthetic. This creates a structural tension: platforms facing the loss of safe harbor for under-removal have a strong incentive to over-remove instead, since erring on the side of caution carries no equivalent penalty under the Rules.³³

C. The Free Speech Tension

The compressed takedown timeline raises a deeper constitutional concern rooted in the Supreme Court's reasoning in *Shreya Singhal v. Union of India* (2015). In that case, the Court read down Section 79 of the IT Act to hold that an intermediary's safe-harbor protection turns on "actual knowledge" received through a *court order or government notification* not through private complaints precisely because requiring intermediaries to adjudicate the legality of content themselves would push them toward excessive, risk-averse takedowns to preserve their immunity.³⁴ The 2026 Rules' three-hour deadline, triggered in some circumstances by government or court directive but reliant heavily on automated detection and platform self-assessment of whether content "appears real," risks reintroducing exactly the dynamic *Shreya Singhal* sought to avoid: a structure where platforms, fearing loss of safe harbor, are incentivized to remove content first and assess its legality later. This tension is not merely theoretical it is presently the subject of live litigation, with the Karnataka High Court hearing arguments in *X Corp v. Union of India* on the broader question of whether the government's takedown powers under the current framework are being exercised more broadly than *Shreya Singhal* permits.³⁵

D. The Non-Celebrity Gap

As Section III demonstrated, the most well-

established avenue of protection against deepfakes personality rights litigation before the Delhi High Court has so far developed almost exclusively around public figures with the recognition, trademark portfolios, and financial resources necessary to mount a civil suit.³⁶ The 2026 Rules improve this position by creating direct regulatory obligations on platforms that do not depend on a victim filing suit at all. However, an ordinary citizen victimized by a non-consensual deepfake still depends entirely on the platform's own compliance there is no statutory mechanism granting such a victim a direct, independent right to compensation or civil damages comparable to what a celebrity plaintiff might secure through litigation. The structural protection has shifted from "litigate if you can afford to" toward "wait for the platform to comply," which is a meaningful improvement, but still falls short of a victim-centered remedy.

V. Comparative Perspective

A brief comparison with two major jurisdictions clarifies what is distinctive and what is missing from India's approach.

A. The European Union

The EU's AI Act takes a fundamentally different regulatory strategy from India's. Rather than mandating *removal* of deepfake content, Article 50(4) of the Act imposes a **disclosure obligation**: any deployer using AI to generate or manipulate image, audio, or video content that constitutes a deepfake must disclose that the content is artificially generated or manipulated, through measures such as persistent visible labels, opening disclaimers for video, or audible warnings for audio.³⁷ Crucially, the EU framework contains a built-in carve-out for evidently artistic, satirical, or fictional content, where the disclosure obligation is reduced to avoid hampering

³³ "IT Rules Amendment 2026: Deepfake Regulation Explained," Insights on India, 11 February 2026.

³⁴ *Shreya Singhal v. Union of India*, (2015) 5 SCC 1.

³⁵ *X Corp v. Union of India*, Karnataka High Court (pending).

³⁶ See Section III above; Nandamuri Taraka Rama Rao v. Ashok Kumar/John Doe & Ors., CS(COMM) 1305/2025.

³⁷ Regulation (EU) 2024/1689 (Artificial Intelligence Act), Article 50(4).



the work's display or enjoyment, and a separate exemption for law-enforcement use.³⁸ These transparency obligations apply from 2 August 2026, with non-compliance attracting fines of up to €7.5 million or 1.5% of global turnover, whichever is higher.³⁹

The contrast with India is instructive: the EU model assumes deepfakes may legitimately exist on platforms so long as they are honestly labelled, whereas India's three-hour and two-hour windows assume the most harmful categories must be removed outright, regardless of labelling. The EU, in other words, regulates *disclosure*; India regulates *removal*. Neither model squarely addresses the other's primary concern the EU framework does little to protect a victim from a labelled-but-still-circulating non-consensual deepfake, while India's removal-first approach, as discussed in Section IV, risks over-removal of legitimate satire or commentary that a disclosure-based model would simply allow to remain, clearly marked.

B. The United States

The United States lacks a single federal framework comparable to India's IT Rules but has converged on similar protections through a combination of federal and state law specifically targeting non-consensual intimate deepfakes. The federal **TAKE IT DOWN Act**, signed into law in May 2025, criminalizes the knowing publication of non-consensual intimate imagery including AI-generated "digital forgeries" and requires covered platforms to establish a notice-and-removal process, with a **48-hour** takedown window from the date a valid request is received; platform compliance under this provision came into effect on 19 May 2026.⁴⁰ Separately, by early 2026,

46 states had enacted their own laws addressing sexually explicit deepfakes, several of which provide civil causes of action allowing victims to sue creators directly for damages — a remedy presently unavailable to an ordinary, non-celebrity victim under Indian law.⁴¹

Two points of contrast with India stand out. First, the US 48-hour window for intimate-imagery takedowns is markedly longer than India's two-to-three-hour window, suggesting India has prioritized speed over the kind of verification process the US notice-and-removal mechanism builds in. Second, and more significantly, the US framework through pending legislation such as the DEFIANCE Act, which passed the Senate in January 2026 and proposes statutory damages of up to \$150,000 (or \$250,000 were linked to sexual assault, stalking, or harassment)⁴² is moving toward a direct, victim-initiated civil remedy independent of any platform compliance failure. India, by contrast, currently offers no equivalent statutory civil remedy to an ordinary deepfake victim; the closest analogue, personality rights litigation, remains realistically available only to plaintiffs with the public profile and resources to litigate, as discussed in Section IV(D).

C. Takeaways for India

Read together, the EU and US examples suggest two reform directions India has not yet taken. From the EU, India could draw the idea of a calibrated, content-sensitive approach distinguishing satire and clearly fictional work from harmful impersonation, rather than relying on a uniform fixed-hour countdown for all "synthetically generated information." From the US, India could draw the idea of a direct, statutory civil remedy for victims, modelled on the proposed

³⁸ Ibid., Article 50(4); Recital 135.

³⁹ Ibid., Article 99(4)(g).

⁴⁰ Tools to Address Known Exploitation by Immobilizing Technological Deepfakes on Websites and Networks Act (TAKE IT DOWN Act), Pub. L. No. 119-12, 139 Stat. (2025).

⁴¹ "Deepfake Legislation Tracker: Federal, State Laws," StackCyber, 14 May 2026.

⁴² Disrupt Explicit Forged Images and Non-Consensual Edits Act of 2025 (DEFIANCE Act), S.3696, 119th Cong. (as passed by the Senate, January 2026).



DEFIANCE Act, that does not depend on a platform's voluntary or compelled compliance.

VI. Recommendations

Based on the gaps identified in Section IV and the comparative lessons drawn in Section V, this section proposes a set of targeted reforms rather than a wholesale rejection of India's current approach, which as demonstrated throughout this article has already evolved considerably since 2023.

A. A Calibrated, Content-Sensitive Labelling Standard

India's regulatory trajectory has so far moved in one direction only: toward stricter labelling. The October 2025 draft proposed a fixed 10% watermark; the February 2026 final Rules replaced this with a qualitative "prominent and easily noticeable" standard; and a further Draft Second Amendment, released on 30 March 2026 and presently under public consultation, proposes tightening this further still, to require labels that are continuous and visible throughout the entire duration of visual content.⁴³ This trajectory, while well-intentioned, has not been matched by an equivalent evolution in *content-sensitive* exemptions. Unlike the EU's AI Act, which calibrates its disclosure obligation to reduce the burden on evidently artistic, satirical, or fictional content, India's Rules apply a largely uniform standard regardless of context. As the consultation on the Second Amendment Rules remains open, this is a live opportunity: India should adopt an explicit, EU-style carve-out for clearly marked satire, parody, and creative work, distinguishing it from impersonation intended to deceive reducing the risk of over-removal

identified in Section IV(B) without weakening protection against genuinely harmful content.

B. A Direct Statutory Civil Remedy for All Victims

The most significant gap identified in this article is not regulatory design but victim access: India's framework, even after the 2026 amendments, still channels the strongest remedies personality rights injunctions of the kind secured in the *NTR Jr.* case toward plaintiffs with the recognition and resources to litigate. Drawing on the model of the proposed US DEFIANCE Act, India should consider a statutory civil cause of action, available to any identifiable victim of a non-consensual deepfake regardless of public profile, providing for fixed statutory damages without requiring proof of quantifiable financial loss a threshold that is often difficult for non-celebrity victims to meet under ordinary tort principles.⁴⁴

C. A National Deepfake Forensics and Verification Capacity

Given the technical detection limitations discussed in Section IV(B), platforms particularly smaller intermediaries cannot be expected to single-handedly develop reliable deepfake-detection capacity within the compressed timelines the Rules now impose. A centrally maintained forensic verification facility, potentially housed within India's existing CERT-In structure or the newly announced AI Safety Institute under the India AI Governance Guidelines, 2025,⁴⁵ could provide platforms especially smaller ones lacking in-house resources a shared verification resource, addressing the enforcement-capacity gap identified in Section IV(A) without requiring each

⁴³ MEITY, Draft Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Second Amendment Rules, 2026 (released 30 March 2026, public consultation closed 7 May 2026); see also Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026, Gazette Notification G.S.R. 120(E) (10 February 2026)

⁴⁴ Disrupt Explicit Forged Images and Non-Consensual Edits Act of 2025 (DEFIANCE Act), S.3696, 119th Cong. (as passed by the Senate, January 2026).

⁴⁵ Ministry of Electronics and Information Technology, India AI Governance Guidelines, 2025 (released at the AI Impact Summit, February 2026).



platform to build duplicate detection infrastructure independently⁴⁶. India already has the beginnings of such capacity: C-DAC, the government's own computing research body, has reportedly developed a prototype deepfake-detection tool claiming approximately 89% accuracy, though independent validation of this figure remains pending.

D. Monitoring the Digital India Act Process

Discussions around a comprehensive Digital India Act, intended eventually to replace the IT Act, 2000 in its entirety, have continued through 2025 and into 2026, with indications that it may introduce risk-based platform classifications and dedicated AI-specific provisions. Separately, a private member's Artificial Intelligence (Ethics and Accountability) Bill, 2025, introduced in the Lok Sabha in December 2025, proposes a statutory AI Ethics Committee and penalties of up to ₹5 crore for certain violations, though it remains unenacted.⁴⁷ It is worth noting that the Government's own official position, articulated through the India AI Governance Guidelines, 2025, is that a standalone AI statute is not presently necessary, on the basis that existing laws,

applied through a risk-based lens, can address most AI-related harms.⁴⁸ This article respectfully takes a different view: while India's current patchwork has proven more responsive than is often assumed as the rapid sequence of amendments between October 2025 and March 2026 demonstrates responsiveness through repeated, fast-moving delegated rulemaking is not a substitute for the durability, parliamentary scrutiny, and victim-centered remedies that a dedicated statute would provide.

VII. Conclusion

This article set out to examine whether India's existing legal framework comprising the Information Technology Act, 2000, the Information Technology Rules as amended in 2025–26, the Bharatiya Nyaya Sanhita, 2023, and the Digital Personal Data Protection Act, 2023 adequately addresses the harms caused by deepfakes, or whether a dedicated, standalone legislative framework is required.

The analysis suggests an answer more nuanced than extreme. India's framework is no longer the empty field it was at the time of the Rashmika Mandanna incident in November 2023, when investigators were forced to stretch identity-theft and privacy provisions never designed for AI-generated content. By early 2026, India had assembled through the IT Act, the BNS, the DPDP Act, and most significantly through the rapidly evolving IT Rules amendments of October 2025, February 2026, and the draft Second Amendment of March 2026 one of the more aggressive platform-accountability regimes for synthetic media anywhere in the world, compressing takedown timelines to as little as two to three hours and embedding mandatory labelling obligations directly into intermediary due diligence requirements. The judiciary, too, has not waited on the legislature: the Delhi High Court's intervention in the *NTR Jr.* case demonstrates a willingness to extend personality rights doctrine to cover AI-generated misuse even ahead of statutory guidance.

Yet adequacy cannot be measured by regulatory speed and stringency alone. As Section IV demonstrated, the current framework remains structurally uneven in three respects. First, its enforcement burden falls disproportionately on smaller intermediaries lacking

⁴⁶ See C-DAC's prototype deepfake-detection tool, reported to score approximately 89% accuracy across global benchmarks, though independent validation remains ongoing; "IT Rules 2026: India's Rapid Deepfake Crackdown," AI CERTs News, 28 March 2026; CERT-In Advisory CIAD-2024-0060 (27 November 2024).

⁴⁷ Artificial Intelligence (Ethics and Accountability) Bill, 2025 (Private Member's Bill, Lok Sabha, introduced December 2025).

⁴⁸ Ministry of Electronics and Information Technology, India AI Governance Guidelines, 2025 (released at the AI Impact Summit, February 2026).



the resources of dominant platforms, risking a level of market consolidation that itself raises competition concerns. Second, its reliance on compressed, complaint-or-detection-triggered takedowns sits uneasily against the constitutional reasoning in *Shreya Singhal v. Union of India*, creating a structural incentive toward over-removal that India's framework has not yet squarely confronted. Third, and most significantly, the strongest available remedy personality rights litigation has so far developed almost exclusively as a shield for public figures with the resources and recognition to litigate, leaving the ordinary, non-celebrity victim dependent on platform compliance rather than possessing any direct, independent statutory remedy of their own.

The comparative review in Section V reinforced this conclusion. The European Union's disclosure-based model and the United States' emerging victim-centered civil remedies each address dimensions of the problem that India's removal-first, platform-centered approach does not. India has, in effect, optimized for speed and platform accountability while leaving open the question of direct victim redress precisely the gap a dedicated standalone statute, rather than further delegated rulemaking under the IT Act, could most durably close.

This article therefore concludes that while India's existing legal framework, taken cumulatively, goes considerably further toward addressing deepfake-related harms than is often assumed, it remains *substantively incomplete* rather than substantively absent. The repeated, rapid cycle of amendments between 2023 and 2026 three major regulatory interventions in under eighteen months is itself evidence that delegated rulemaking under the IT Act, however responsive, is an unstable foundation for a harm of this magnitude and durability. What India requires is not a wholesale abandonment of its current approach, but its consolidation and completion: a dedicated Synthetic Media Act that preserves the platform-accountability gains of the 2025–26 IT Rules while adding the element conspicuously missing from them a direct, victim-initiated statutory

remedy, available to every individual regardless of public profile, that does not depend on a platform's compliance, a court's discretion, or the resources required to litigate.
