



This is an Open Access article distributed under the terms of the Creative Commons Attribution-Non-Commercial-Share Alike 4.0 International (CC-BY-NC-SA 4.0) License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium provided the original work is properly cited.

Deepfakes, Synthetic Media, and the IT Rules: Does India Need an AI-Specific Liability Standard?

Himanshu Dhar^a

^aLloyd Law College, Greater Noida, India

Received 12 August 2026; Accepted 12 September 2026; Published 16 September 2026

India regulated deepfakes for three years the way it regulated any other unlawful post: through post-publication complaints, platform takedown timelines, and scattered offences under the Information Technology Act, 2000. That model assumed a human publisher who could be identified, warned, and prosecuted; generative AI removed that assumption by letting anyone manufacture a stranger's face, voice, and words at a scale no complaint-driven system was built to absorb. This paper traces how that gap was closed first through an improvised line of Delhi High Court personality-rights injunctions, then through the Ministry of Electronics and Information Technology's Draft IT Rules of October 2025, and finally through the amendment that came into force on 20 February 2026, bringing 'synthetically generated information' within the due-diligence framework of the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, and compressing platform response timelines to as little as two hours. Reading this amendment against Article 50 of the European Union's Artificial Intelligence Act, the paper argues that India's model remains structurally downstream: it obliges platforms to detect and label synthetic content after it is generated, rather than obliging the tools that generate it to mark their own output at the point of creation. It proposes a three-tier statutory standard: generation-point provenance, platform-level verification, and a codified civil remedy for ordinary victims as a more coherent alternative to the current combination of criminal offence, platform rule, and judge-made injunction.

Keywords: *deepfakes, synthetic media, due diligence, personality rights, digital impersonation.*

INTRODUCTION

On 6 November 2023, a fabricated video superimposing the face of actress Rashmika Mandanna onto another woman's body went viral on X, prompting the Ministry of Electronics and Information Technology to remind platforms, within days, that Rule 3(2)(b) of the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 already obliged them to remove 'artificially morphed images' within twenty-four hours of a complaint.¹ Two years later, the Ministry conceded to the Delhi High Court that this complaint-driven model had not kept pace: deepfakes were increasingly used to target women during state elections, no uniform statutory definition of 'deepfake' existed, and enforcement, not fresh legislation, was framed as the more urgent problem.² What changed between those two moments was not the underlying law so much as the scale of the technology outrunning it: generative tools that once required specialist skill could, by 2025, manufacture a convincing likeness of any identifiable person's face or voice within minutes, at negligible cost, and often through a foreign-hosted tool entirely outside the reach of an Indian takedown notice.

STATEMENT OF THE PROBLEM

India's legal response to synthetic media, until February 2026, rested on three separate and uncoordinated instruments: criminal offences under the Information Technology Act, 2000 that punish impersonation and privacy violation after the fact; platform due-diligence obligations under the 2021 IT Rules that require takedown on complaint rather than prevention at source; and an emerging body of Delhi High Court injunctions recognising personality and publicity rights as a civil remedy for identifiable public figures. None of the three obliges the entity that actually generates synthetic content the AI tool itself to mark its output as artificial at the moment of creation. This paper argues that the due-diligence amendment which took effect on 20 February 2026, while a considerable tightening of

¹ 'Union Government issues advisory to social media intermediaries to identify misinformation and deepfakes' (*Press Information Bureau*, 07 November 2023) <<https://www.pib.gov.in/PressReleaseIframePage.aspx?PRID=1975445®=48&lang=2>> accessed 12 August 2026; Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules 2021, r 3(2)(b)

² *Rajat Sharma v Union of India* WP(C) 6560/2024

platform obligations, preserves this structural gap, and that a coherent standard requires shifting at least part of the compliance burden upstream, to the point of generation.

RESEARCH QUESTIONS

1. Did the pre-2026 combination of IT Act offences, IT Rules due diligence, and judge-made personality-rights injunctions provide an adequate remedy against synthetic-media harms, particularly for claimants who are not public figures?
2. Does the February 2026 amendment to the IT Rules close that gap, or does it merely compress timelines within the same downstream, platform-facing model?
3. What would a statutory standard look like that assigns responsibility across the point of generation, the point of hosting, and the point of civil remedy, rather than leaving each to a different, uncoordinated instrument?

SCOPE AND METHODOLOGY

This is a doctrinal paper grounded in primary Indian sources – the Information Technology Act 2000, the IT Rules 2021 as amended in October 2022, April 2023, and February 2026, the Bharatiya Nyaya Sanhita 2023, and the relevant Delhi High Court orders – read alongside Article 50 of the European Union's Artificial Intelligence Act as a comparative reference point. It does not evaluate the technical feasibility of provenance-marking or watermarking tools; it takes the existing, publicly disclosed state of that technology as a fixed input around which a workable legal standard is proposed.

THE PRE-2026 REGULATORY PATCHWORK

Criminal Law: Reactive, Person-Specific, and Slow: Section 66D of the Information Technology Act, 2000 punishes ‘cheating by personation by using computer resource’ with imprisonment of up to three years and a fine of up to one lakh rupees.³ Sections 66C and 66E similarly address identity theft and violation of bodily privacy, but each requires a complainant to identify a specific offender, initiate a criminal complaint, and await investigation a process unsuited to content that can be re-uploaded from a new anonymous account within minutes of takedown. The Bharatiya Nyaya Sanhita 2023 layers on offences of forgery, cheating, and defamation, but these too are calibrated to a single identifiable

³ Information Technology Act 2000, s 66D

publisher, not to synthetic content that can be regenerated and redistributed by anyone, anywhere, at negligible cost.

Platform Due Diligence Under the 2021 IT Rules: The more operative layer, in practice, was Rule 3 of the IT Rules 2021. Rule 3(1)(b)(vii) required intermediaries to inform users, through their terms of service, not to host content that impersonates another person, while Rule 3(2)(b) obliged intermediaries to remove content ‘in the nature of impersonation in an electronic form, including artificially morphed images,’ within twenty-four hours of a complaint from the affected individual.⁴ A parallel, broader takedown obligation under Rule 3(1)(d) gave intermediaries thirty-six hours to act on a government or court order concerning unlawful content generally.⁵ Rule 7 stripped an intermediary of its safe-harbour immunity under section 79 of the IT Act if it failed to observe these obligations.⁶

The Structural Gap: The result, before 2026, was a regime built entirely around the point of hosting and the point of complaint, never the point of generation. A synthetic video could be produced using a tool hosted outside Indian jurisdiction, uploaded through a series of disposable accounts, and regenerated in a materially altered form as soon as one copy was taken down, while the only accountable actor under Indian law remained the platform hosting the final upload, and even that obligation triggered only after a human complainant noticed and reported it.

THE JUDICIARY FILLS THE GAP: THE DELHI HIGH COURT'S PERSONALITY-RIGHTS LINE

In the absence of a statutory remedy calibrated to synthetic media, the Delhi High Court developed, case by case, a common-law personality and publicity rights doctrine to fill the space. In *Rajat Sharma v Tamara Doc*, the Court restrained several online entities from misusing the personality and publicity rights of journalist Rajat Sharma and the trademarks of India TV, following an ex parte interim order in December 2024 that directed Meta to block the infringing content, made permanent in August 2025.⁷ A similar suit brought by spiritual leader Sri Sri Ravi Shankar resulted in an order restraining circulation of deepfake content

⁴ Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules 2021, rr 3(1)(b)(vii), 3(2)(b)

⁵ *Ibid* r 3(1)(d)

⁶ *Ibid* r 7; Information Technology Act 2000, s 79

⁷ *Rajat Sharma and Ors v Tamara Doc and Ors* CS(COMM) 1147/2024

using his voice, face, and persona, and directing suspension of the offending domains.⁸ More recently, in *Ranganathan Madhavan v G Filmz Studioz*, the Court granted interim relief against AI-generated deepfake film trailers and morphed content using actor R. Madhavan's likeness,⁹ and in *Akira Nandan v Sambhawaami Studios LLP*, against an AI-generated film portraying the plaintiff in fabricated romantic scenes without consent.¹⁰

These orders share a structure: each was a 'John Doe' injunction against unknown defendants, treated the conduct as prima facie misappropriation of identity rather than protected expression, and directed named intermediaries to take down the content and disclose the uploaders' account details. The underlying doctrine that personality and publicity rights, though uncoded in Indian statute, are protected under Article 21 and the common law of privacy and passing-off traces back to the constitutional vocabulary the Supreme Court supplied in *K.S. Puttaswamy v Union of India*.¹¹

FROM DRAFT TO FORCE: THE FEBRUARY 2026 AMENDMENT

The Committee Process: The statutory response began not with a ministry initiative but with a court direction. Hearing petitions filed by Rajat Sharma and advocate Chaitanya Rohilla, the Delhi High Court, on 21 November 2024, directed the Central Government to nominate members to a dedicated committee formed the previous day to examine deepfake regulation, drawing on the experience of intermediaries, telecom providers, and victims, and taking into account foreign frameworks including the European Union's.¹² By March 2025, the Ministry's status report to the Court flagged, among other concerns, the absence of any uniform statutory definition of 'deepfake' and the disproportionate targeting of women during state elections, while cautioning that better enforcement of existing law, not necessarily new legislation, might be the more urgent need.

The October 2025 Draft and Its Critics: On 22 October 2025, the Ministry released the Draft Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2025, introducing the term 'synthetically generated information' and proposing a mandatory labelling requirement covering at least ten per cent of a visual

⁸ *Ravi Shankar v John Doe(S) / Ashok Kumar(S) and Ors* CS(COMM) 889/2025

⁹ *Ranganathan Madhavan v G Filmz Studioz and Ors* (2025) CS(COMM) 1392/2025

¹⁰ *Mr Akira Desai Alias Akira Nandan v Sambhawaami Studios LLP & Ors* CS(COMM) 68/2026

¹¹ *Justice KS Puttaswamy (Retd) and Anr v Union of India and Ors* (2017) 10 SCC 1

¹² *Rajat Sharma v Union of India* WP(C) 6560/2024

display, or the first ten per cent of an audio clip's duration.¹³ The Internet Freedom Foundation, among other stakeholders, urged the Ministry to withdraw the draft, arguing that its definition of synthetic information was overbroad enough to capture routine AI-assisted editing, and that a labelling-and-takedown model built on intermediary due diligence risked functioning as a disproportionate censorship mechanism rather than a narrow response to demonstrated harms.¹⁴

What Changed by February 2026: The final notification, issued on 10 February 2026 and in force from 20 February 2026, responded to at least some of this criticism. It narrowed the definition of 'synthetically generated information' to audio, visual, or audio-visual content that is artificially generated and likely to be perceived as indistinguishable from a real person or event, and dropped the fixed ten-per-cent labelling threshold in favour of a requirement that labels be 'clearly and prominently' displayed, without a prescribed numerical minimum.¹⁵ It also excluded routine, good-faith editing, formatting, colour correction, transcription, and similar operations that do not materially alter meaning from the definition altogether.

Where the amendment tightened rather than relaxed the regime was in compliance speed:

Obligation	Pre-Amendment (2021 Rules)	Post-Amendment (enforced 20 Feb 2026)
Takedown on Govt./Court Order – Rule 3(1)(d)	36 hours	3 hours

¹³ 'EXPLANATORY NOTE [Proposed Amendments to the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 in relation to synthetically generated information]' (*Ministry of Electronics and Information Technology*, 22 October 2025) <<https://www.meity.gov.in/static/uploads/2025/10/8e40cdd134cd92dd783a37556428c370.pdf>> accessed 08 August 2026

¹⁴ Apar Gupta and Naman Kumar, 'Withdraw the Draft Synthetic Information IT Rules, 2025' (*Internet Freedom Foundation*, 15 November 2025) <<https://internetfreedom.in/withdraw-the-draft-synthetic-information-it-rules-2025/>> accessed 08 August 2026

¹⁵ Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules 2026

Grievance Acknowledgment – Rule 3(2)(a)(i)	15 days	7 days
First Proviso to Rule 3(2)(a)(i)	72 hours	36 hours
Impersonation/Morphed Content on Individual Complaint – Rule 3(2)(b)	24 hours	2 hours

Significant social media intermediaries are additionally required to obtain a user declaration on upload as to whether content is synthetically generated, apply ‘reasonable technical measures’ to verify that declaration, and embed permanent metadata or a traceable identifier where feasible. The Election Commission of India, following release of the October 2025 draft, separately issued an advisory on the responsible use of synthetically generated information during elections.¹⁶

A COMPARATIVE LOOK: THE EU AI ACT'S UPSTREAM MODEL

Article 50 of the European Union's Artificial Intelligence Act takes a materially different approach. It places a transparency obligation directly on the provider of an AI system that generates synthetic audio, image, video, or text content, requiring that the system's output be marked in a machine-readable format and detectable as artificially generated or manipulated.¹⁷ The obligation attaches to the generation tool itself at the moment content is created rather than solely to whichever platform later happens to host it.

This is not merely a difference in drafting style. A generation-point obligation reaches content the very first time it is produced, before it has been uploaded to any Indian platform, regardless of where the tool's developer is based, so long as the system is offered into the EU market. A hosting-point obligation, by contrast, can only ever regulate what happens after

¹⁶ ‘EC Issues Fresh Guidelines on AI and Synthetic Content Use in Bihar Poll Campaigns’ (*News on AIR*, 24 October 2025) <<https://newsonair.gov.in/ec-issues-fresh-guidelines-on-ai-and-synthetic-content-use-in-bihar-poll-campaigns/>> accessed 08 August 2026

¹⁷ EU Artificial Intelligence Act 2024, art 50

the fact, and depends entirely on an Indian-registered ‘significant social media intermediary’ being the one to host the content a dependency that generative-AI distribution, which routinely moves content across messaging apps, private groups, and foreign-hosted tools before it ever reaches a regulated platform, is well placed to exploit.

DOES INDIA NEED AN AI-SPECIFIC LIABILITY STANDARD? A THREE-TIER PROPOSAL?

The central argument of this paper is that India's February 2026 amendment, though a genuine improvement, completes only one of three tiers a coherent standard requires.

Tier One – Generation-point provenance: Providers of AI tools capable of generating a realistic likeness of an identifiable natural person should be placed under a statutory obligation, modelled on Article 50 of the EU AI Act, to embed a machine-readable, tamper-resistant provenance marker in their output at the point of generation, enforceable against any provider offering such a tool into the Indian market, regardless of where it is hosted.

Tier Two – Platform verification and labelling: The obligations introduced in February 2026 user declaration, technical verification, and prominent labelling should be retained, but strengthened with a published minimum technical standard for what counts as an acceptable ‘reasonable technical measure,’ so that compliance cannot collapse into a checkbox exercise.

Tier Three – A Codified Civil Remedy: Rather than leaving ordinary victims dependent on a Delhi High Court ‘John Doe’ injunction that presupposes the resources to litigate, Parliament should codify a statutory tort of digital impersonation, available to any identifiable individual, not only public figures, through a simplified, low-cost forum such as the Consumer Protection Act's district commissions or a dedicated grievance-appellate mechanism under the IT Rules themselves.

These tiers assign responsibility to the three points at which synthetic media actually causes harm: creation, distribution, and injury to the individual rather than concentrating the entire compliance burden, as the current framework does, on the single point of hosting.

CONCLUSION

India regulated deepfakes, until February 2026, largely as an afterthought to a content-moderation regime designed for an earlier, human-authored internet. The Delhi High Court's personality-rights injunctions supplied real but selectively available relief; the February 2026 amendment supplied real but structurally downstream reform. Neither reaches the point at which synthetic content is actually made. A standard that assigns statutory responsibility to the generation tool itself, alongside a tightened platform obligation and an accessible civil remedy for ordinary victims, would close a gap that faster takedown timelines, however welcome, cannot close on their own.