

**Synthetic Media and the Law of Personal Identity in India: Deepfakes,
Platform Liability and the Constitutional Limits of the Information
Technology Amendment Rules, 2026**

Parul

Assistant Professor, Chaudhary Ranbir Singh University, Jind

Abstract

Generative artificial intelligence has made the fabrication of a person's face, voice and gestures inexpensive, rapid and convincing. India responded on 10 February 2026 by notifying the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026, which for the first time define "synthetically generated information", require prominent labelling and embedded provenance metadata, oblige significant social media intermediaries to obtain and technically verify user declarations about synthetic content, and compress the removal deadline for flagged unlawful material to three hours. This paper evaluates that framework against the constitutional standards governing intermediary liability and freedom of expression, and against the private law of personality that Indian courts have been constructing case by case since 2022. It argues that Indian law is now organised around the wrong regulatory object. The 2026 Rules treat synthetic media principally as a problem of deception, to be cured by disclosure; but the gravest harms — non-consensual intimate imagery, sexual humiliation, identity-based fraud and dignitary injury — are not cured by a label, because they are not diminished by the audience knowing that the depiction is fabricated. Simultaneously, the compressed timelines, the incentive to over-remove, the reliance on automated detection and the absence of independent review sit uneasily with the reading down of section 79(3)(b) of the Information Technology Act, 2000 in *Shreya Singhal*, with the proportionality standard in *K.S. Puttaswamy*, and with the Bombay High Court's reasoning in *Kunal Kamra*. The paper further shows that the judicial remedy that presently works — the wide John Doe injunction protecting personality rights — is available in practice only to claimants of celebrity, leaving ordinary victims with a criminal process ill-adapted to speed and anonymity. Drawing on the EU AI Act, China's labelling measures and the United States TAKE IT DOWN Act, the paper proposes a statutory right in one's likeness and voice, a harm-differentiated remedial structure, evidentiary reform under the Bharatiya Sakshya Adhinyam, 2023, and independent oversight of takedown power.

Keywords: Deepfakes; synthetically generated information; personality rights; intermediary liability; safe harbour; Information Technology Amendment Rules, 2026; non-consensual intimate imagery; freedom of expression; proportionality; India.

2. Introduction

The legal system has always assumed a rough correspondence between a recorded image and the event it records. Photographic and audio-visual evidence, defamation by publication, the tort of passing off, the offence of personation and the commercial law of endorsement all rest on the premise that fabricating a convincing likeness is difficult, expensive, and detectable. Generative models have dissolved that premise. A usable voice clone can be produced from

seconds of sample audio, a video substitution from a handful of photographs, and the output distributed to millions before any legal process can be initiated.

India's regulatory response has been rapid but uneven. Between 2022 and 2026 the High Courts at Delhi and Bombay assembled, largely through interim injunctions, a body of personality rights jurisprudence protecting name, image, voice, likeness and idiosyncratic mannerisms against artificial reproduction. The executive moved separately: advisories to intermediaries in 2023 and 2024, the India Artificial Intelligence Governance Guidelines released on 5 November 2025, the Digital Personal Data Protection Rules notified on 14 November 2025, and finally the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026, notified on 10 February 2026 and brought into force ten days later. Parliament has legislated nothing specific to synthetic media at all.

That distribution of institutional effort is itself the subject of this paper. India now regulates deepfakes through delegated legislation made under a statute enacted in 2000, supplemented by injunctions granted mostly at the instance of well-resourced public figures, and by criminal provisions drafted for analogue harms. The result is a framework with two structural defects. It misidentifies the principal harm, treating synthetic media as a deception problem soluble by labelling when its most serious manifestations are dignitary and sexual; and it allocates the power of suppression to platforms and the executive on timelines that make meaningful assessment impossible, without the procedural safeguards that Indian constitutional law has consistently demanded of restrictions on speech.

The paper proceeds in eight further parts. Part 2 states the research problem and method. Part 3 offers a taxonomy of synthetic media harms. Part 4 maps the pre-2026 legal position and its inadequacies. Part 5 sets out the architecture of the 2026 Rules. Part 6 assesses that architecture against constitutional doctrine. Part 7 compares the Indian approach with three foreign models. Part 8 examines the gendered distribution of harm and the access-to-justice deficit. Part 9 proposes reform.

3. Research Problem, Objectives and Methodology

The research problem is that India's emerging law of synthetic media protects the integrity of the information environment more effectively than it protects the individual whose identity is appropriated, while simultaneously creating a mechanism of content suppression that lacks the procedural safeguards required of restrictions on protected speech.

The objectives are, first, to construct a harm-based taxonomy of synthetic media and to test the existing legal apparatus against it; second, to analyse the Amendment Rules of 2026 in their statutory and constitutional setting; third, to evaluate the judicially developed law of personality rights as a remedial route for ordinary claimants; fourth, to draw comparative lessons from jurisdictions that have legislated on the same subject; and fifth, to propose a coherent statutory scheme.

The research questions follow: (i) Which harms of synthetic media does Indian law presently address, and which does it leave unremedied? (ii) Are the obligations created by the 2026 Rules — in particular the compressed takedown timeline, automated verification and the clarification concerning safe harbour — consistent with section 79 of the Information Technology Act, 2000 as construed in *Shreya Singhal* and with the proportionality standard applicable to Article 19(1)(a)? (iii) Is the personality rights jurisprudence a viable general remedy or a remedy of

privilege? (iv) What legislative design would address dignitary, deceptive and commercial harms distinctly?

The method is doctrinal, analytical and comparative. Primary sources comprise the Constitution of India, the Information Technology Act, 2000 and the rules made under it, the Bharatiya Nyaya Sanhita, 2023, the Bharatiya Sakshya Adhiniyam, 2023, the Digital Personal Data Protection Act, 2023 and Rules, 2025, the Copyright Act, 1957, and the judgments and interim orders of the Supreme Court and the High Courts. Comparative material is drawn from Regulation (EU) 2024/1689, the Chinese labelling measures effective from 1 September 2025, and the United States legislation of May 2025 concerning non-consensual intimate imagery. Secondary literature comprises peer-reviewed scholarship and institutional reports. The study is confined to civil, constitutional and regulatory questions; it does not address the technical evaluation of detection systems except as a matter of legal feasibility, and it does not undertake empirical measurement of prevalence.

4. A Taxonomy of Synthetic Media Harms

Regulatory instruments are frequently drafted as though “deepfake” denoted a single wrong. It does not, and the differences matter because they call for different remedies.

The first category is sexual and dignitary harm: non-consensual intimate imagery in which a real person’s face or body is synthesised into sexual material. Its injury is to dignity, psychological integrity and social standing, and it is inflicted at the moment of viewing. Disclosure that the image is artificial does not undo it; a labelled fabrication of a woman in a sexual act humiliates her as effectively as an unlabelled one. The only adequate remedies are removal, non-recirculation and punishment.

The second is fraud and impersonation: cloned voices used to obtain money from relatives, fabricated video of executives authorising transfers, synthetic endorsements of investment schemes. Here the harm is economic and the deception is constitutive, so disclosure obligations and provenance tracing are genuinely responsive.

The third is informational and democratic: fabricated statements attributed to public officials, synthetic audio released during election periods, manufactured evidence of events that did not occur. Labelling helps, but the temporal problem is acute, since a fabrication circulated shortly before a poll may achieve its object before any correction is issued.

The fourth is commercial appropriation: use of a person’s likeness or voice to sell goods or generate engagement. This is a proprietary injury for which the common law of passing off and the developing personality rights jurisprudence are reasonably well adapted.

The fifth is evidentiary and is the least discussed. The plausibility of fabrication supplies every litigant with a new defence: that authentic recordings are themselves synthetic. This “liar’s dividend” degrades the reliability of audio-visual proof generally, and its consequences fall on the law of evidence rather than on content regulation.

A framework that responds to all five with a single instrument — disclosure — will be overinclusive for expressive uses and under-inclusive for dignitary ones. That is precisely the defect identified below.

5. The Legal Position Before February 2026

Statutory fragments

No Indian statute addressed synthetic media as such. Prosecutors assembled charges from provisions drafted for other purposes. Under the Information Technology Act, 2000, section 66C penalises the fraudulent use of another's electronic identity features, section 66D the cheating by personation through a computer resource, section 66E the capture or publication of images of a private area, and sections 67, 67A and 67B the transmission of obscene, sexually explicit and child sexual abuse material. Under the Bharatiya Nyaya Sanhita, 2023, the relevant provisions include section 77 (voyeurism), section 78 (stalking), section 79 (insulting the modesty of a woman), sections 318 and 319 (cheating and cheating by personation), section 336 (forgery) and section 356 (defamation). Performers enjoy limited protection under sections 38A and 38B of the Copyright Act, 1957, though the ownership of the output of a generative model trained on a performance is unsettled.

Two difficulties recur. The provisions are keyed to conduct — capture, personation, transmission — by an identified actor, whereas synthetic content is typically produced anonymously and disseminated by intermediaries. And several are keyed to obscenity rather than to consent, so a non-obscene but humiliating fabrication may fall outside them altogether. Under the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, rule 3(2)(b) required intermediaries to remove, within twenty-four hours of a complaint, content depicting a person in a sexual act or in the nature of impersonation, including artificially morphed images. That was the closest thing to a deepfake remedy, and it depended entirely on the victim finding each copy and complaining about it.

The judicial construction of personality rights

The more effective response came from the High Courts. Building on *R. Rajagopal v. State of Tamil Nadu* (1994) 6 SCC 632 and on the recognition of informational privacy and dignity in *K.S. Puttaswamy v. Union of India* (2017) 10 SCC 1, courts granted expansive injunctions protecting attributes of persona. In *Amitabh Bachchan v. Rajat Nagi* (Delhi High Court, 2022) relief extended to name, voice and image; in *Anil Kapoor v. Simply Life India* (2023) the Court protected likeness, voice and a signature catchphrase while expressly preserving space for satire, parody and legitimate commentary; in *Jaikishan Kakubhai Saraf v. Peppy Store* (2024) similar protection was extended, again with a free speech carve-out. In *Arijit Singh v. Codible Ventures LLP* (Bombay High Court, 2024) the Court restrained artificial intelligence platforms from enabling the cloning of the plaintiff's voice, treating synthetic vocal reproduction as an actionable invasion of persona; comparable relief was granted to Asha Bhosle. In September 2025 the Delhi High Court granted an ex parte interim injunction in favour of Aishwarya Rai Bachchan against websites, e-commerce listings, chatbots and deepfake videos exploiting her identity, and granted analogous relief to Abhishek Bachchan; in August 2025 it restrained the circulation of fabricated videos of a religious leader endorsing health claims he had never made. This jurisprudence is doctrinally significant. It locates personality rights simultaneously in Article 21 dignity and privacy and in the commercial logic of passing off; it accepts that voice and mannerism are protectable attributes; and it deploys John Doe orders and dynamic injunctions to reach unidentified defendants and future mirror sites. But it has three limits. It is

interim, so the governing principles have rarely been settled after contest. It is uncodified, so its scope varies with the pleadings before each court. Most importantly, it is commercially framed: the reasoning turns on goodwill, endorsement value and the misappropriation of a marketable persona. A woman whose face has been synthesised into pornography has no goodwill to protect and no commercial loss to plead. She has a dignitary injury — and the remedy that in practice works, the wide injunction with platform-directed takedown, has been developed in a body of case law to which her claim does not naturally belong.

6. The Architecture of the Amendment Rules, 2026

The Amendment Rules notified on 10 February 2026 and effective from 20 February 2026 do not create a new statute; they extend the due diligence framework of the 2021 Rules to synthetic content. Five elements are central.

First, definition. Newly inserted rule 2(1)(wa) defines synthetically generated information as audio, visual or audio-visual information that appears reasonably authentic or true but is artificially or algorithmically created, generated, modified or altered using a computer resource, in a manner likely to be perceived as indistinguishable from a natural person or a real event. Routine automated enhancements that do not represent real-world events are excluded.

Second, labelling and provenance. Synthetic content must carry a prominent and, in the case of visual media, persistently visible label, together with embedded metadata or provenance markers containing a unique identifier permitting the content to be traced to the computer resource that produced it. Intermediaries are forbidden to enable the removal, suppression or alteration of such labels or metadata.

Third, verification by significant social media intermediaries. Under rule 4(1A), such platforms must require users to declare at the point of upload whether content is synthetically generated, and must deploy automated technical measures to verify that declaration rather than accept it at face value. Knowingly permitting or promoting unlabelled synthetic content constitutes a failure of due diligence.

Fourth, the takedown clock. For content flagged through a court order or an authorised governmental notification, the period for removal or disablement is reduced from thirty-six hours to approximately three hours, with correspondingly shortened grievance redressal timelines. Categories treated as most serious include non-consensual intimate depictions, deceptive impersonation and child sexual abuse material.

Fifth, the safe harbour clarification. Rule 2(1B) provides that removing or disabling access to synthetic content in compliance with the Rules, including through automated tools and including where the intermediary acts of its own motion on becoming aware of a violation, does not compromise the protection available under section 79 of the Information Technology Act, 2000.

Taken together the Rules mark a decisive shift from a complaint-driven model, in which the intermediary was a passive conduit obliged to act on notice, to an anticipatory model in which it must classify, mark, verify and monitor. That shift is the source of both the framework's promise and its constitutional difficulty.

7. Constitutional and Doctrinal Assessment

Safe harbour and the legacy of Shreya Singhal

In *Shreya Singhal v. Union of India* (2015) 5 SCC 1 the Supreme Court struck down section 66A and read down section 79(3)(b), holding that an intermediary's obligation to remove content arises only upon actual knowledge in the form of a court order or a notification by the appropriate government, and that the government notification must itself relate to a ground specified in Article 19(2). The reason was not formalism. The Court recognised that if private platforms were required to judge legality on private complaint, they would remove first and reason later, because removal is costless to them and refusal is not.

The 2026 Rules do not overrule that holding for ordinary content, but they alter its practical force in two ways. The safe harbour clarification in rule 2(1B) invites intermediaries to act suo motu against synthetic content without an order, and assures them that doing so is safe. Rule 4(1A) requires them to detect unlabelled synthetic content proactively. The combined incentive structure points in exactly the direction *Shreya Singhal* sought to prevent: platforms will resolve every doubt in favour of removal, because a wrongful removal exposes them to nothing while a wrongful retention imperils immunity. The asymmetry is not cured by the availability of grievance redressal, which is internal to the platform and operates after the speech has already been suppressed.

Proportionality and the three-hour window

Restrictions on expression must satisfy the proportionality standard articulated in *K.S. Puttaswamy* and applied to communications restrictions in *Anuradha Bhasin v. Union of India* (2020) 3 SCC 637: a legitimate aim, a rational connection, necessity in the sense that no less restrictive alternative suffices, and a balance between the restriction and the right. Preventing sexual fabrication and fraud is plainly legitimate, and speed is rationally connected to the aim, since the harm of viral distribution is done early. The difficulty lies in necessity and balance. A three-hour compliance window makes individualised legal assessment impossible at scale; compliance can be achieved only by automated matching and pre-emptive removal, technologies that cannot evaluate context, and therefore cannot distinguish a malicious fabrication from satire, parody, art, political caricature, academic illustration or journalism reporting on the very fabrication in issue. Less restrictive alternatives exist and are not obviously inadequate: differentiated timelines by harm category, with the shortest applying to non-consensual intimate imagery and child sexual abuse material, and longer periods for content whose unlawfulness turns on context.

The absence of independent review compounds the difficulty. The determination that content is unlawful is made by an executive authority or by the platform, and the affected speaker learns of it after the event. In *Kunal Kamra v. Union of India* (Bombay High Court, 2024) the constitutional validity of the amendment establishing a governmental fact check unit under the 2021 Rules was successfully challenged; the tie-breaking opinion held the provision incompatible with Articles 14, 19(1)(a) and 19(1)(g), reasoning that the State could not constitute itself the arbiter of truth in respect of speech concerning its own affairs, and that vague standards conferred effectively unreviewable discretion. The reasoning is not confined to the fact check unit. Any mechanism by which an executive authority determines that content

is unlawful and secures its suppression within hours, without notice to the speaker or independent adjudication, invites the same objection.

Definitional breadth

The definition of synthetically generated information turns on whether content “appears reasonably authentic” and is likely to be perceived as indistinguishable from a real person or event. This is a perception test without a specified audience, a specified degree of realism, or a requirement of intent to deceive. On a literal reading it captures a considerable volume of ordinary expression: restored historical footage, dubbed and lip-synchronised translations, reconstructions in documentary film, synthetic voice used as an accessibility measure, and computer-generated imagery in commercial cinema. Exclusions for routine enhancement mitigate but do not resolve this, because the boundary between enhancement and generation is a technical continuum. Where a regulatory definition is both broad and vague, and the sanction is loss of statutory immunity, the predictable consequence is conservative compliance by intermediaries and a corresponding contraction of lawful expression.

The mislocated regulatory object

The most fundamental objection is conceptual. The Rules are constructed around disclosure: the wrong to be prevented is that an audience should mistake artifice for reality, and the cure is to inform the audience. That analysis fits fraud, impersonation and political misinformation. It does not fit sexual and dignitary harm at all. A synthetic intimate image of an identifiable woman, correctly labelled as synthetic and carrying immaculate provenance metadata, remains a humiliation. Indeed labelling may aggravate it by authenticating the identification of the subject while disclaiming only the technique. Conversely, provenance requirements impose costs on the very producers most likely to comply and least likely to cause harm, while malicious actors operating outside the regulated ecosystem strip metadata as a matter of course. Regulation designed around the compliant producer will systematically miss the non-compliant one.

Evidence and the liar’s dividend

A consequence largely absent from the regulatory debate concerns proof. Section 63 of the Bharatiya Sakshya Adhiniyam, 2023 governs the admissibility of electronic records and requires a certificate as to the manner of production. That mechanism authenticates the process by which a record was retrieved from a device; it does not establish that the content of the record corresponds to a real event. As synthetic media becomes ubiquitous, contested audiovisual evidence in criminal and matrimonial proceedings will increasingly attract the defence of fabrication, and courts will be asked to resolve authenticity questions for which the statutory scheme provides no framework. Provenance metadata of the kind mandated by the 2026 Rules could, if standardised and preserved, become forensically valuable — but only if evidence law is amended to give it a defined status. At present the two regimes do not speak to each other.

8. Comparative Perspectives

Three foreign models illustrate different regulatory objects, and the contrast clarifies the Indian choice.

Regulation (EU) 2024/1689, the Artificial Intelligence Act, addresses synthetic media through transparency rather than prohibition. Article 50(2) obliges providers of generative systems to mark outputs in a machine-readable format detectable as artificially generated, and Article 50(4) obliges deployers of deepfakes to disclose that the content is artificially generated or manipulated; these obligations apply from 2 August 2026, supported by a Code of Practice on the transparency of AI-generated content. The European approach is deliberately upstream, imposing duties on those who build and deploy the systems rather than principally on those who host the output, and it preserves exceptions for evidently artistic, satirical and fictional works.

China's labelling measures, effective from 1 September 2025, impose both explicit labels perceptible to users and implicit labels embedded in file metadata, and extend obligations along the distribution chain to content dissemination platforms and application stores. The regime is comprehensive and covers text as well as audio-visual content, but it is embedded in a broader system of content control that renders it a poor model for a jurisdiction bound by Article 19(1) (a).

The United States legislation enacted in May 2025 on non-consensual intimate imagery takes the opposite route. It does not regulate synthetic media generally. It criminalises the publication of non-consensual intimate images, including computer-generated depictions of identifiable individuals, and requires covered platforms to establish a notice-and-removal process operating within forty-eight hours of a valid request. The instrument is narrow, harm-specific and victiminitiated, and it targets precisely the category that disclosure obligations cannot address.

Table 1. Regulatory object, principal obligation and residual gap in four synthetic media regimes

Regime	Primary regulatory object	Principal obligation	Residual gap
European Union — AI Act, Article 50 (applicable 2 August 2026)	Deception of the audience	Machine-readable marking by providers; disclosure of deepfakes by deployers; exceptions for artistic and satirical work	No dedicated victim-initiated removal right for nonconsensual intimate imagery
China — Labelling Measures (effective 1 September 2025)	Traceability across the distribution chain	Explicit user-facing labels and implicit metadata labels; duties on platforms and application stores	Embedded in a wider contentcontrol system; poor fit for a jurisdiction bound by Article 19(1)(a)
Regime	Primary regulatory object	Principal obligation	Residual gap

United States — TAKE IT DOWN Act, 2025	Sexual and dignitary harm to an identified person	Criminal liability for publication; platform notice-and-removal within forty-eight hours of a valid request	No general transparency or provenance regime for synthetic content
India — IT Amendment Rules, 2026	Deception of the audience, enforced through intermediaries	Labelling and provenance metadata; user declaration verified by automated tools; three-hour removal on authorised notice	No statutory right in likeness or voice; no victim-initiated removal right; no independent review of suppression

The comparison indicates that the Indian framework has adopted the European regulatory object — transparency — while implementing it through the Chinese institutional method of distribution-chain liability and short compliance deadlines, and while omitting the American harm-specific victim remedy altogether. That combination maximises the burden on intermediaries and on expression while leaving the most serious individual harm dependent on the general takedown machinery.

9. The Gendered Distribution of Harm and the Access-to-Justice Deficit

Available evidence from multiple jurisdictions indicates that the overwhelming majority of non-consensual synthetic media depicts women in sexual contexts. The consequences reported by victims — withdrawal from public and professional life, reputational injury within family and community, harassment following circulation — are of a kind that Indian law recognises when inflicted by other means, through the offences of voyeurism, stalking and outraging modesty, and through the constitutional protection of dignity affirmed in *Puttaswamy*.

The remedial reality is nevertheless stark. A public figure with counsel obtains, within days, a John Doe injunction extending to unnamed defendants, with directions to platforms and, increasingly, dynamic effect over future mirror copies. A woman without those resources has three options, each inadequate. She may complain to the platform under rule 3(2)(b) of the 2021 Rules as now amended, which addresses the copy she has found but not those she has not. She may report the matter to the police, whose capacity for the forensic tracing of anonymous synthetic content is uneven, and whose process is measured in months. Or she may sue, which requires identifying a defendant, funding litigation and accepting that the proceedings will publicise the very material she wishes suppressed.

This is an Article 14 problem as much as an Article 21 problem. The state has created a mechanism of rapid removal and placed its operation in the hands of executive notification and platform discretion, while the individual most acutely injured has no corresponding right to trigger it and no remedy if it is not triggered. A framework that delivers three-hour removal on governmental notice, but leaves an unrepresented victim to file complaints copy by copy, distributes the benefit of speed to the state and its burden to her.

10. Recommendations

First, a statutory right in likeness and voice. Personality rights should be codified, with the right vesting in every person rather than depending on demonstrable goodwill, with defined exceptions for news reporting, satire, parody, criticism, artistic and scholarly use, and with remedies including injunction, damages and mandatory de-indexing. This would place the woman in the deepfake pornography case and the film star in the endorsement case on the same statutory footing, which the present commercially framed jurisprudence does not.

Second, harm-differentiated remedies. The regulatory response should track the taxonomy in Part 3. Non-consensual intimate imagery and synthetic child sexual abuse material warrant the shortest removal timelines, hash-matching to prevent recirculation, and criminal liability for creation as well as distribution; labelling is irrelevant to them. Fraud and impersonation warrant provenance and disclosure duties. Political and informational fabrication warrants disclosure with heightened obligations during election periods. Commercial appropriation warrants the civil remedy proposed above. A single instrument applied uniformly to all five categories will misfire in both directions.

Third, a victim-initiated statutory removal right. Every person should have a right, exercisable directly and without litigation, to require an intermediary to remove synthetic intimate or impersonating depictions of themselves, with a defined response period, an obligation to prevent re-upload of matching content, and a penalty for non-compliance. A single-window national portal, with support for regional languages and legal aid referral, should be its operational vehicle.

Fourth, independent oversight of suppression. Takedown directions issued otherwise than by a court should be reviewable by an independent statutory body with power to restore content, subject to time limits and to publication of aggregate transparency data. This addresses the objection identified in Kunal Kamra without impairing the state's capacity to act quickly against genuinely unlawful material.

Fifth, differentiated compliance timelines with an obligation of preservation. Content removed under compressed timelines should be preserved by the intermediary for a defined period rather than deleted, so that evidence remains available both for the victim's prosecution and for the speaker's challenge. Removal and destruction should not be conflated.

Sixth, alignment of evidence law. The Bharatiya Sakshya Adhiniyam, 2023 should be amended to provide a framework for the authentication of audio-visual records where synthetic origin is alleged, including the evidentiary significance of provenance metadata, the allocation of the burden where a party asserts fabrication, and the role of accredited forensic examination. Without this, content regulation and adjudication will continue to develop on separate tracks.

Seventh, statutory footing for governance instruments. The India Artificial Intelligence Governance Guidelines of November 2025 and the various sectoral advisories are useful but non-binding. Whether through the anticipated digital legislation or a dedicated enactment, the definitional and provenance standards now contained in delegated legislation should be placed in primary law, so that the boundaries of an intermediary's duty and the citizen's corresponding right are determined by Parliament rather than by notification.

10. Conclusion

The Amendment Rules of 2026 are a serious regulatory effort and, in several respects, an advance. They give Indian law its first operative definition of synthetic content, they establish provenance as a legal requirement rather than an industry aspiration, and they remove the argument that a platform acting against deepfakes risks its statutory immunity. Those are real gains.

The framework nevertheless misidentifies its object and underprotects its most exposed subjects. Built around disclosure, it treats the wrong of synthetic media as deception, when the wrong that most urgently requires legal response is the appropriation of an identifiable person's body and voice for humiliation, extortion or profit — a wrong that a label does not diminish. Built around executive notification and platform detection on three-hour timelines, it creates a suppression mechanism whose incentives run consistently toward over-removal, without the independent review that Shreya Singhal, Anuradha Bhasin and Kunal Kamra together indicate is constitutionally required. And built as delegated legislation upon a statute of 2000, it leaves the fundamental questions — what a person owns in their own face and voice, and what they may demand when it is taken — to be answered by interim injunctions in commercial suits brought by those who can afford them.

The corrective is not to abandon transparency but to supplement it. A codified right in likeness and voice available to every person, remedies differentiated by the harm they address, a victiminitiated removal right with real consequences for non-compliance, independent review of executive suppression, and an evidence law equal to a world in which recordings can be manufactured: these are the elements of a framework that would protect both the individual and the expression that a constitutional democracy must tolerate. Until they exist, Indian law will continue to offer its most effective protection to those who need it least.

References

- Aishwarya Rai Bachchan v. Aishwaryaworld.com & Ors., CS(COMM) 956/2025, Delhi High Court, order dated 9 September 2025.
- Amitabh Bachchan v. Rajat Nagi & Ors., CS(COMM) 819/2022, Delhi High Court, order dated 25 November 2022.
- Anil Kapoor v. Simply Life India & Ors., 2023 SCC OnLine Del 6914.
- Anuradha Bhasin v. Union of India, (2020) 3 SCC 637.
- Arijit Singh v. Codible Ventures LLP & Ors., Bombay High Court, interim order dated 26 July 2024.
- Balkin, J. M. (2018). Free speech in the algorithmic society: Big data, private governance, and new school speech regulation. *UC Davis Law Review*, 51(3), 1149–1210.
- Bharatiya Nyaya Sanhita, 2023 (No. 45 of 2023). Government of India.
- Bharatiya Sakshya Adhiniyam, 2023 (No. 47 of 2023). Government of India.
- Chesney, R., & Citron, D. K. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review*, 107(6), 1753–1820.
- Citron, D. K. (2019). Sexual privacy. *Yale Law Journal*, 128(7), 1870–1960.
- Constitution of India, 1950, Articles 14, 19(1)(a), 19(2), 21, 32, 226.
- Copyright Act, 1957 (No. 14 of 1957), sections 38A and 38B. Government of India.

Cyberspace Administration of China. (2025). Measures for the labelling of artificial intelligence generated synthetic content (effective 1 September 2025).

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

DM Entertainment Pvt. Ltd. v. Baby Gift House, 2010 SCC OnLine Del 4790.

European Commission. (2025). Draft Code of Practice on transparency of AI-generated content. European AI Office.

European Union. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), Article 50. Official Journal of the European Union.

Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026, G.S.R. 120(E), notified 10 February 2026, effective 20 February 2026. Ministry of Electronics and Information Technology.

Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021. Ministry of Electronics and Information Technology.

Information Technology Act, 2000 (No. 21 of 2000), sections 66C, 66D, 66E, 67, 67A, 67B and 79. Government of India.

Jaikishan Kakubhai Saraf alias Jackie Shroff v. The Peppy Store & Ors., 2024 SCC OnLine Del 3664.

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

Kaye, D. (2019). Speech police: The global struggle to govern the internet. Columbia Global Reports.

Kunal Kamra v. Union of India, Bombay High Court, judgment dated 20 September 2024 (Writ Petition (L) No. 9792 of 2023).

Ministry of Electronics and Information Technology. (2025). Digital Personal Data Protection Rules, 2025, notified 14 November 2025. Government of India.

Ministry of Electronics and Information Technology. (2025). India Artificial Intelligence Governance Guidelines. Government of India.

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

Rajinikanth v. Varsha Productions, 2015 SCC OnLine Mad 158.

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

Solove, D. J. (2006). A taxonomy of privacy. University of Pennsylvania Law Review, 154(3), 477–560.

Sundara Rajan, M. T. (2011). Moral rights: Principles, practice and new technology. Oxford University Press.

Tata Sons Ltd. v. Greenpeace International, 2011 SCC OnLine Del 466.

United States Congress. (2025). Tools to Address Known Exploitation by Immobilizing Technological Deepfakes on Websites and Networks Act (TAKE IT DOWN Act), Public Law 119-12.

Westerlund, M. (2019). The emergence of deepfake technology: A review. Technology Innovation Management Review, 9(11), 39–52.