INTERNATIONAL JOURNAL OF HUMAN RIGHTS LAW REVIEW
An International Open Access Double Blind Peer Reviewed, Referred Journal

Volume 5 | Issue 6 | 2026 Art. 2

Beyond Labels: A Victim-Centric Legislative Framework for Harmful Synthetic Impersonation in India

Devika Singhal
Law Student,
5th Year, BBA.LL.B.,
O.P. Jindal Global University, Sonipat, Haryana

Recommended Citation
Devika Singhal, *Beyond Labels: A Victim-Centric Legislative Framework for Harmful Synthetic Impersonation in India,* 5 IJHRLR 12-23 (2026).
Available at www.ijhrlr.in/current-issues/.

This Article is brought to you for free and open access by the International Journal of Human Rights Law Review by an authorized Lex Assisto & Co. administrator.
For more information,
please contact submission.ijhrlr@gmail.com

Beyond Labels: A Victim-Centric Legislative Framework for Harmful Synthetic Impersonation in India

EXECUTIVE SUMMARY

India entered a new age of deepfake regulation in February 2026. The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026 now define "synthetically generated information" (SGI), require intermediaries providing SGI-capable resources to take reasonable technical measures to address unlawful SGI, require prominent labeling and technically feasible provenance mechanisms for permissible SGI, and impose declaration and verification obligations on significant social media intermediaries.[1] These rules materially strengthen accountability for platforms, but they remain fundamentally a due diligence framework. They do not establish a coherent primary-law claim against synthetic identity appropriation, a SOCiTandard civil remedy for ordinary victims, or carefully calibrated offenses for harms such as non-consensual intimate deepfakes, malicious voice cloning, fraudulent impersonation and fabricated digital evidence.

This policy brief proposes amendments to the Information Technology Act, 2000, to introduce a targeted statutory framework for "harmful synthetic impersonation", along with consequential criminal and procedural reforms. The framework should differentiate between harmful conduct and legitimate synthetic media; create a specific offense for non-consensual intimate synthetic imagery and mens rea-based offenses for fraudulent or malicious impersonation; provide every natural person with a civil cause of action; increase victim-triggered takedown, preservation and prevention of re-upload; establish interoperable provenance standards; place graduated safety-by-design duties on high-risk generative services; and create specialized cybercrime and forensic capacity. Safeguards should be provided to protect journalism, satire, parody, art, research and other lawful uses for the public interest.

The goal is not to ban the technology of deepfakes. It aims to regulate identifiable harms in a technologically neutral, victim centered and constitutionally proportionate way. The proposed framework would complement, not replace, India's 2026 intermediary regime while boosting legal certainty, dignity and privacy, regulatory efficiency and

[1] The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026, G.S.R. 120(E), Gazette of India, Extraordinary, Part II, sec. 3(i), Feb. 10, 2026, rr. 1-4.

public trust.

BACKGROUND

Synthetic media are increasingly able to reproduce or simulate a person's face, voice, conduct or speech in a realistic way. Legitimate applications include dubbing, accessibility, education, entertainment and creative production. These same capabilities can generate sexual imagery without consent, impersonate people for fraud, generate defamatory content, or fabricate apparently credible evidence. Academic literature has thus warned against the idea that digital manipulation per se is unlawful and prefers responses targeting harmful uses.[2]

The problem is not anymore a matter of speculation. The OECD in 2026 reported that the share of media-reported AI incidents and hazards related to synthetic media increased 2.5-fold between 2022 and 2025 and accounted for more than 14 per cent in the third quarter of 2025.[3] This is not a prevalence estimate, but it does show increasing visibility. The gendered aspect is especially serious. According to a 2024 report by the UN Secretary-General, citing Sensity AI, estimates indicate that 90 to 95 per cent of online deepfakes were non-consensual pornographic images and about 90 per cent of these images involved women.[4]

India has already replied. The 2026 amendment to the IT Rules, coming into force on 20 February 2026, defines SGI as "audio, visual or audio-visual information that is artificially or algorithmically manipulated or generated, that appears to be real, authentic or truthful and depicts a person or event that is likely to be perceived as indistinguishable from reality. Routine good faith editing, accessibility improvements, and standard educational, research and document preparation are excluded.[5] The question is whether intermediary due diligence, older offenses and private-law doctrines together provide a complete rights-and-remedies framework for the person whose identity is synthetically appropriated.

STATEMENT OF PROBLEM

The existing framework is fragmented in four ways.

[2] Robert Chesney and Danielle Citron, "Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security" 107 California Law Review 1753 (2019).
[3] Organisation for Economic Co-operation and Development, "Trends in AI Incidents and Hazards Reported by the Media" 17 (2026).
[4] UN Secretary-General, Intensification of Efforts to Eliminate All Forms of Violence against Women and Girls: Technology-Facilitated Violence against Women and Girls, UN Doc A/79/500 (Oct. 8, 2024), para. 23.
[5] *Supra* note 1, rr. 1-2.

First, the wrong underlying is typically found indirectly. Liability is contingent on the presence of some other offense or cause of action such as cheating, personation, breach of privacy, obscenity, defamation or passing off. This leaves cases in which the core injury is non-consensual synthetic appropriation of identity, voice, likeness, or sexual autonomy.

Second, the 2026 Rules are broad as an intermediary regime but are still delegated legislation based on due diligence. Rule 3(3) already addresses unlawful SGI, which includes non-consensual intimate imagery, privacy-invasive content, false documents and deceptive portrayals of a person's identity, voice, conduct, action or statement. It also mandates labels and technically feasible provenance for permissible SGI.[6] Still lacking is a generally applicable substantive right against harmful synthetic impersonation and clear remedies against originators.

Third, court remedies are uneven. Indian courts have responded with doctrines of personality, publicity, privacy and intellectual-property, especially in celebrity litigation. No one should have to prove celebrity-style commercial personality rights to get quick relief from a sexualized deepfake, cloned-voice scam or humiliating synthetic video.

Fourth, labeling and takedown do not address circulation and evidence problems. Labels may be cropped, recordings re-encoded and copies transferred between services. Victims require evidence preservation, mechanisms for materially identical re-uploads, specialized forensic support and a clear path to damages or injunctions.

LEGAL FRAMEWORK IN PLACE

Constitutionally, dignity, privacy and freedom of expression are the basis for targeted reform. In Justice K.S. Puttaswamy (Retd.) v. Union of India, the Supreme Court recognized privacy as inherent to liberty, dignity and autonomy.[7] These interests can be engaged by synthetic appropriation of a person's intimate identity, face or voice. But any prohibition must be consistent with article 19(1)(a), so precision and proportionality are crucial.[8]

The Information Technology Act, 2000 addresses a number of related harms. Sections 66C and 66D deal with identity theft and cheating by personation through the use of computer resources. Section 66E deals with the intentional or knowing capture, publication or transmission of the image of a private area of any person without his or her consent, under circumstances violating privacy. Sections 67 and 67A deal with

[6] *Supra* note 1, r. 3(3).
[7] *Justice K.S. Puttaswamy (Retd.) v. Union of India*, (2017) 10 SCC 1.
[8] The Constitution of India, arts. 19(1)(a), 21.

obscene and sexually explicit electronic material.[9] While these provisions might be relevant to some deepfake fact patterns, they do not establish a general offense of harmful synthetic impersonation. Thus, section 66E is textually linked to a 'image of a private area' and its applicability is more complicated where the intimate image is wholly fictitious.

Relevant offenses within the framework of the abuse of synthetic media are covered in the Bharatiya Nyaya Sanhita, 2023, such as cheating, cheating by personation, forgery and defamation.[10] Liability again turns on the individual ingredients of those offenses. A fabricated sexual image may violate dignity and sexual autonomy even if no actual intimate photo exists, and a malicious voice clone may inflict serious harm even if no property fraud is involved.

The February 2026 IT Rules is the most direct intervention. Rule 3(3) requires reasonable and appropriate technical measures with regard to a specified unlawful SGI and labels and technically feasible provenance with regard to a permissible SGI. The amendment also reduced the rule 3(2)(b) grievance timeline for specific intimate and impersonation-related content from 24 hours to two hours, reduced the relevant rule 3(1)(d) timeline from 36 hours to three hours, and requires significant social media intermediaries to obtain and technically verify user declarations about SGI and prominently label confirmed SGI. Rule 4(4) further strengthens the obligation to deploy appropriate technical measures against certain content removed previously.[11] These reforms should form the basis for the next phase.

Data-protection law is merely a partial and transitional answer. 18 months from the date of publication in accordance with G.S.R. 843(E) dated 13th November, 2025 Sections 3 to 5, Section 6 (1) to (8) and (10) and Sections 7 to 17 of Digital Personal Data Protection Act, 2023. 1 year from the date of publication in accordance with G.S.R. 843(E) dated 13th November, 2025 Section 6(9) of Digital Personal Data Protection Act, 2023.[12] The main processing provisions relevant here are not yet in operation as at 29 September 2026. Further, once section 3 applies, section 3(c)(ii)(A) excludes personal data made or caused to be made available to the public by the Data Principal.[13] An example is provided by a photo that was posted publicly and then used for synthetic impersonation,

[9] The Information Technology Act, 2000 (Act 21 of 2000), ss. 66C, 66D, 66E, 67, 67A.
[10] The Bharatiya Nyaya Sanhita, 2023 (Act 45 of 2023), ss. 318, 319, 336, 356.
[11] *Supra* note 1, rr. 3(1)(d), 3(2)(b), 3(3), 4(1A), 4(4).
[12] Ministry of Electronics and Information Technology, Notification G.S.R. 843(E), Gazette of India, Extraordinary, Part II, sec. 3(i), Nov. 13, 2025.
[13] The Digital Personal Data Protection Act, 2023 (Act 22 of 2023), s. 3(c)(ii)(A).

showing the inadequacy of data-protection law.

The Indian cases on personality rights exhibit both judicial responsiveness and doctrinal limits. The Delhi High Court, in Anil Kapoor v. Simply Life India, granted an ex parte interim injunction restraining specified misuse of the plaintiff's name, likeness, image, voice and persona through technologies including AI and deepfakes.[14] Arijit Singh v. Codible Ventures LLP – Bombay High Court – Ex parte ad-interim order passed on a prima facie assessment relating to AI voice models and synthetic use of personality attributes.[15] These interlocutory orders were made in celebrity-centred litigation and are not a general statutory remedy for all natural persons.

Finally, Shreya Singhal v. Union of India established constitutional boundaries for "actual knowledge" under section 79(3)(b).[16] Instead, new duties should employ narrow statutory categories and procedural safeguards, not transform platforms into unregulated private adjudicators of lawful expression.

COMPARATIVE VIEWPOINT

Three lessons learned, three comparative models.

Under Article 50 of the European Union Artificial Intelligence Act, providers that generate synthetic audio, image, video or text must ensure machine-readable marking and detectability. Deployers of systems that generate or manipulate deepfake image, audio or video must disclose its artificial nature. This provision applies to obviously artistic, creative, satirical, fictional or similar works.[17] India's lesson is to pair visible labels with interoperable technical standards and expressive safeguards.

In May 2025, the United States passed the TAKE IT DOWN Act, which criminalizes certain nonconsensual intimate visual depictions and digital forgeries, and provides for takedowns initiated by victims. Platforms covered must remove qualifying material within forty-eight hours of a valid request and must make reasonable efforts in relation to known identical copies.[18] This links a specified underlying wrong to a structured remedy for the victim.

In England and Wales section 188 of the Online Safety Act 2023 concerns

[14] *Anil Kapoor v. Simply Life India*, 2023 SCC OnLine Del 6914.
[15] *Arijit Singh v. Codible Ventures LLP*, 2024 SCC OnLine Bom 2445.
[16] *Shreya Singhal v. Union of India*, (2015) 5 SCC 1.
[17] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 (Artificial Intelligence Act), art. 50(2), (4).
[18] Tools to Address Known Exploitation by Immobilizing Technological Deepfakes on Websites and Networks Act, 2025 (U.S.), Pub. L. No. 119-12, ss. 2-3, 139 Stat. 55.

sharing or threatening to share an intimate photograph or film that "shows, or appears to show" another person in an intimate state, the definition including computer-generated or altered imagery.[19] The Crime and Policing Act 2026 makes provision about purported intimate-image generators and requires reported intimate-image content, including the same or substantially the same material, to be taken down as soon as is reasonably practicable and no later than forty-eight hours, subject to statutory qualifications.[20]

India should adopt these ideas selectively: EU-style provenance, victim-triggered intimate-image removal, and proportionate action against substantially identical re-uploads, but not a blanket ban on synthetic media.

STAKEHOLDER ANALYSIS

Victims need speed, confidentiality, preservation of evidence and remedies that do not depend on fame or proof of commercial loss. Synthetic intimate abuse is particularly harmful to women and girls, while cloned voice fraud and reputation manipulation is more widespread. Law enforcement agencies need reliable preservation procedures, trained forensic personnel and legal access to relevant account and provenance information.

We need objective legal thresholds for platforms. Both over-removal and under-removal tend to increase with vague duties, and may make compliance with safe-harbours uncertain. AI developers require technically feasible, risk-calibrated obligations so that everyday creative and productivity tools are not regulated as if they were services designed for sexualized impersonation. Clearer rules against fraudulent synthetic identity and reliable indicators of provenance benefit businesses and consumers.

Civil society organizations, journalists, artists, comedians, educators and researchers also have a legitimate stake in preventing censorship. Satire, parody, fictionalization and transformative creative uses can include realistic synthetic media without producing the harms that justify legal intervention. That is to say, the regulatory unit should be harmful conduct and context, not synthetic generation in the abstract. The recent model framework of UN Women also emphasizes the need for thorough, survivor-centered legislation and effective remedies for technology-

[19] The Online Safety Act 2023 (UK), c. 50, s. 188.
[20] The Crime and Policing Act 2026 (UK), c. 20, ss. 99-100.

facilitated violence, rather than relying on isolated offenses only.[21]

PROPOSED LAW CHANGE

India needs to add a specific chapter on 'Harmful Synthetic Impersonation' to the Information Technology Act, 2000, along with consequential criminal and procedural amendments. It should be explicitly complementary to the 2026 Rules.

A. Definition of Harmful Synthetic Impersonation

The law should apply to materially synthesized or manipulated audio, visual or audio-visual content that falsely depicts an identifiable natural person saying, doing or appearing in a manner the person did not, where the depiction falls under defined categories of harm. These may include: non-consensual intimate representation; fraud or extortion; targeted harassment or intimidation; fabrication of evidence; and malicious impersonation causing serious reputational, psychological or economic harm.

We should exclude routine editing, accessibility tools, clearly fictional or disclosed creative works, and bona fide research. This indicates academic concern that a total ban on technology might also cover good uses and uses connected to expression.[22]

B. Develop Calibrated Criminal Offenses

Deliberate creation of non-consensual intimate synthetic imagery for distribution . Publishing it . Intentionally enabling distribution or threatening publication without the consent of the person depicted . This should be a specific offense even if the sexual act depicted never existed . Consent to creation should not mean consent to publication without further thought.

For non-sexual synthetic impersonation liability should require knowledge or recklessness as to falsity and intent to defraud, extort, harass, intimidate or cause serious harm. This mens rea distinguishes malicious impersonation from mistake, satire, and legitimate experimentation. The consequences may be aggravated if the victim is a child, the conduct is organized or repeated or the impersonation supports extortion or large-scale fraud.

C. Create a Civil Right Against Synthetic Misappropriation

[21] UN Women, "Model Framework for Legislation on Technology-Facilitated Violence against Women and Girls" (2025).
[22] *Supra* note 2.

Every natural person should have a civil cause of action where his or her name, face, likeness, voice or other identifying attribute is synthetically appropriated without lawful justification and causes, or is likely to cause, serious harm. Relief should include urgent interlocutory and permanent injunctions, deletion orders, compensatory damages, disgorgement of profits where commercial exploitation is established and reasonable costs.

The right should cover noncommercial damage. Courts should also be able to authorize pseudonymous filing and sealed sensitive exhibits in intimate-image cases and to issue carefully tailored dynamic injunctions against materially identical copies when necessary.

D. Prevention of Victim-Triggered Bolster Removal, Preservation and Re-Upload

The two-hour grievance mechanism put in place in 2026 should be maintained, with clearer statutory backing, for verified high-risk complaints. Emergency takedown of clearly non-consensual intimate synthetic imagery should not require a prerequisite FIR or civil order, as viral spread can render subsequent relief ineffective.

Major platforms should take reasonable and proportionate efforts to prevent the re-upload of identical or substantially similar copies of confirmed non-consensual intimate synthetic content through the use of perceptual hashing or similar technologies. Original material and relevant account information and available provenance data should be retained for a prescribed period. Disputed non-intimate material should be subject to notice to the uploader, meaningful human review, recorded reasons and appeal. Models of comparative "identical copy" and "substantially the same" show how re-upload prevention can work with defined triggers, rather than indiscriminate monitoring.[23] [24]

E. Make Provenance Interoprable and Evidentially Modest

The 2026 Rules already require "permanent metadata or other suitable provenance mechanisms" to the extent technically feasible. MeitY, the Bureau of Indian Standards, technology companies, independent researchers and civil society should now open standards for durable content credentials, provenance verification and machine readable disclosure.

The legislation should make clear that provenance is an evidentiary indication, not a proof of truth. The Coalition for Content Provenance

[23] *Supra* note 18, s. 3.
[24] *Supra* note 20, s. 100.

and Authenticity warns that provenance does not prove factual truth and can be removed from media.[25] Lack of metadata shouldn't be enough to prove falsity, nor should credentials be enough to prove authenticity. Legally required provenance should only be removed, or falsified, intentionally to facilitate a defined unlawful act, with safeguards for forensic, security and interoperability research.

F. Impose Increasing Safety-by-Design Duties

Specialized tools, designed or marketed for realistic face swapping, voice cloning, or "nudification," should have a higher preventative duty than normal general-purpose tools. Obligations should cover misuse-risk assessments, reasonable default safeguards, accessible abuse reporting and proportionate record retention. The regulatory burden should be a function of capability, scale, foreseeable misuse, and the provider's ability to mitigate harm.

This upstream approach is in line with emerging international policy that prefers prevention and safety by design to placing the entire burden on survivors after circulation.[26]

G. Build Institutional Capacity for Specialization

A Synthetic Media Response Cell should be established within the national cybercrime architecture and not as a new independent regulator. It should be able to coordinate victim referrals, inter-state investigations, digital-forensic assistance and standard operating procedures. Build SGI-specialist capacity in the existing Grievance Appellate Committee mechanism.

Investigators, prosecutors and judicial officers need to be trained on voice clone fraud, synthetic intimate imagery, provenance, preservation of evidence and the limits of automated detectors. The government should publish anonymized statistics each year on categories of complaint, response times, re-upload rates and enforcement outcomes.

H. Safeguard Legitimate Expression

The law should explicitly protect satire, parody, artistic and fictional works, bona fide journalism, academic and scientific research, education, authorized law-enforcement activity and use in judicial proceedings, where these are clearly disclosed. These safeguards should be context

[25] Coalition for Content Provenance and Authenticity, "C2PA Explainer", available at: https://c2pa.org/specifications/specifications/2.2/explainer/Explainer.html (last visited on Sep. 29, 2026).
[26] UN Women, "UN Women Strategy on Preventing and Eliminating Technology-Facilitated Violence against Women and Girls" 33 (2025).

sensitive: a perpetrator should not be able to turn non-consensual intimate abuse into protected "parody" by simply adding a label.

The goal should be a harm-based framework that targets lack of consent, deception and serious injury, while protecting legitimate innovation.

ROADMAP FOR IMPLEMENTATION

Phase I – Immediate Reforms (0-6 months) The Central Government should form a multidisciplinary drafting group with MeitY, Ministry of Home Affairs, digital-forensics experts, women's rights organizations, technology companies, academics and civil society. It must issue a consultation paper, a model victim notice form and a protocol for the preservation of evidence MeitY and Bureau of Indian Standards should start work on interoperable provenance standards.

Phase II – Medium-Term Reforms (6-18 months): Parliament should pass amendments and Government should notify graduated compliance rules. Major platforms should add verified synthetic-harm reporting and re-upload blocking for confirmed non-consensual intimate imagery. The Synthetic Media Response Cell, forensic support, transparency reporting and training for investigators, prosecutors and judicial officers should be operational.

Phase III – Long-Term Reforms (18-36 months): An independent review should evaluate victim outcomes, false-positive removals, compliance costs, technological circumvention, provenance effectiveness and impacts on lawful expression. Offenses, individual rights, defenses and core safeguards should be in primary legislation, with technical standards open to updating in subordinate legislation.

EXPECTED ADVANTAGE

The framework would strengthen governance by assigning responsibility across the synthetic-media lifecycle: to originators for intentional harmful conduct, to high-risk providers for proportionate preventive design, to intermediaries for defined response and preservation duties, and to public institutions for enforcement.

Legal certainty would be improved, since harmful synthetic impersonation would be separated from legitimate synthetic creation, rather than pushed into older offenses. Increased regulatory efficiency by building on the 2026 Rules rather than creating a parallel platform regime. Victims would have a simpler route to quick relief and intermediaries and developers would have more predictable thresholds

Better provenance, better preservation of evidence and redress would

help restore public confidence. Instead of general licensing for generative AI, the way to support economic development would be risk-calibrated obligations. Rights protection would also be better balanced: dignity, privacy and sexual autonomy for the persons depicted, and freedom of expression for journalists, artists, researchers and other legitimate speakers.

CONCLUSION

India's 2026 SGI amendments are a major regulatory step forward. They put in place intermediary obligations on illicit synthetic content, labeling, provenance, verification and fast-track grievance redressal. The remaining gap is a coherent primary-law framework that defines harmful synthetic impersonation as a substantive wrong and provides clear remedies against the underlying misconduct for each affected person.

So the next stage should be focused. Parliament should define harmful synthetic impersonation, create a specific consent-based offense for intimate synthetic abuse, calibrate other offenses through clear mens rea, establish a civil right for ordinary people, strengthen victim-triggered removal and evidence preservation and embed interoperability and safety-by-design into the technological chain. Safeguards should maintain satire, art, journalism, research and other lawful expression.

Deepfakes shouldn't be regulated simply because they're artificial. Regulation is needed when synthetic representation vitiates consent, appropriates identity, deceives, intimidates or causes serious harm. A framework based on that principle would shift Indian law from a platform-focused response to integrated prevention, accountability and redress, while allowing for legitimate innovation.