International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Regulation of Deepfakes, Identity Fraud and Ai-Created Harms Under New Penal Provisions

Ana Zehra[1], Dr Jyoti Yadav[2]

[1]Student, LLM (CL &CS), Amity Law School, Lucknow
[2]Assit. Professor, Amity Law School, Lucknow

Abstract
Deepfakes, AI voice cloning and synthetic identities now sit at centre of digital risk landscape in India. These tools allow hyper realistic alteration of faces, bodies and voices, so they enable intimate image abuse, financial fraud and targeted political manipulation on huge scale. Recent Indian and comparative scholarship shows that such AI created harms strike directly at privacy, autonomy, reputation and even public trust in democratic processes. At same time, legal responses stay fragmented across Information Technology Act 2000, Bharatiya Nyaya Sanhita 2023, Bharatiya Sakshya Adhiniyam 2023 and Digital Personal Data Protection Act 2023, none of which yet define deepfakes or synthetic media in express terms. This paper studies how new penal architecture should address deepfakes, identity fraud and wider AI created harms. It maps provisions on cheating, impersonation, obscenity, defamation, cyber offences and digital evidence under BNS and BSA and then tests their capacity against deepfake related misuse such as non-consensual pornography, fraud and political disinformation. The analysis is placed against proposals for dedicated Deepfake Prevention and Regulation framework in India and recent Deepfake Prevention and Criminalisation Bill introduced in Parliament. Doctrinal and comparative method is used to conceptualise "AI created harms" and "identity fraud," to link them with victim consent, expectation of privacy and reputational interests. The paper finally argues that Indian penal policy must become victim centred and technologically informed while still respecting data protection guarantees and due process in investigations and trials.
This paper argues that the current fragmented approach is inadequate. It proposes a unified, victim oriented regulatory framework. That integrates criminal law, data protection and intermediary responsibility. Such an approach is necessary to ensure both effective enforcement and protection of fundamental rights.

Keywords: Deepfakes, Artificial Intelligence, Identity Fraud, Cyber Law, Digital Evidence, Data Protection

INTRODUCTION AND BACKGROUND OF THE STUDY

A. Evolution of AI, Deepfakes, and Synthetic Media

Deepfakes started as experimental outputs of deep learning models trained on facial images and short clips.[1] Very soon, generative adversarial networks and diffusion models made face swapping, lip syncing and style transfer accessible to ordinary users, not only expert coders.[2] The ecosystem now covers video

[1] Shinu Vig, 'Regulating Deepfakes: An Indian Perspective' (2024) 17(3) Journal of Strategic Security 70.
[2] *Ibid.*

face swaps, image editing, voice cloning and full body synthesis that create completely synthetic persons. Through cheap online tools, user can clone voice of another person from few seconds of audio and paste it on scripted phone call or ransom demand.[3] Because content looks and sounds real, harms spread across non consensual intimate deepfakes, romance scams, investment frauds, and micro targeted political propaganda.[4] Scholars also emphasise that these harms erode baseline assumption that "seeing is believing" in law, media and everyday communication, which is extremely serious.[5]

B. Indian Digital Ecosystem and Rise of AI Harms

India now has huge and still growing internet and smartphone user base, with dense use of social media, messaging apps and short video platforms. In such crowded digital ecosystem, manipulated audio visual content spreads very fast and becomes viral much before any formal takedown process finishes.[6] Documented Indian incidents already show deepfake pornography against women, impersonation of executives for fund transfers, and fake political clips circulated during elections.[7] Law review literature also notes coordinated trolling and harassment where intimate deepfakes are used to shame, silence or extort women and vulnerable groups. Police and courts still mostly proceed under cheating, obscenity, defamation, voyeurism or cyber stalking provisions, which were drafted for simple digital content, not AI synthesis.[8] Therefore enforcement practice sometimes struggles on questions like attribution of content, responsibility of platforms, and proving authenticity or manipulation of digital evidence.

C. Conceptualising "AI Created Harms" and "Identity Fraud"

AI created harms may be understood as harms where artificial intelligence tools are integral in generating, modifying or amplifying harmful content or conduct. This includes harms where identity, likeness or voice of real person is used without consent, as well as fully synthetic characters designed for deception.[9] Identity fraud in this context involves unauthorised use of identifying data, biometrics, images or voice prints to impersonate, deceive, or cause damage. The same deepfake engine can support harmless parody, entertainment or education when consent exists and context is clear, and still power severe exploitation when consent is absent. Hence, boundary between harmless and harmful use lies in factors like informed consent, notice, labelling, expectation of privacy and realistic foreseeability of reputational or financial damage.[10] Indian constitutional privacy jurisprudence and dignity oriented readings of Article 21 also push regulators to treat non consensual intimate deepfakes as serious violation of autonomy and personhood.[11]

D. New Penal Architecture and Regulatory Context

India has recently replaced Indian Penal Code 1860 with Bharatiya Nyaya Sanhita 2023, alongside new procedural and evidentiary codes. The BNS restructures offences related to cheating, forgery, criminal intimidation, cyber terrorism and sexual offences, with specific references to electronic records and digital

---

[3] Prachi Agnihotri, 'Legal Implications of Deepfake Technology in Criminal Law' (2025) 8(1) International Journal of Law Management and Humanities 1645.
[4] *Ibid.*
[5] Jaromir Cerny, 'The Admissibility of Deepfake Evidence within Context of Indian Law' (2025) 7(6) International Journal for Multidisciplinary Research (IJFMR) 1.
[6] Shinu Vig, 'Regulating Deepfakes: An Indian Perspective' (2024) 17(3) Journal of Strategic Security 70.
[7] Jaromir Cerny (n 3).
[8] Deepfake Prevention and Criminalisation Bill, 2023 (as introduced in Rajya Sabha).
[9] Prachi Agnihotri, 'Legal Implications of Deepfake Technology in Criminal Law' (2025) 8(1) International Journal of Law Management and Humanities 1645.
[10] Jaromir Cerny, 'The Admissibility of Deepfake Evidence within Context of Indian Law' (2025) 7(6) International Journal for Multidisciplinary Research (IJFMR) 1.
[11] K S Puttaswamy v Union of India (2017) 10 SCC 1.

communication.[12] The Bharatiya Sakshya Adhiniyam 2023 refines rules on admissibility and authenticity of electronic evidence, which matters lot when parties contest whether media is deepfake or genuine.[13] Side by side, Information Technology Act 2000 and Intermediary Guidelines of 2021 impose duties on platforms to remove unlawful content and cooperate with law enforcement.[14] The Digital Personal Data Protection Act 2023 further builds data protection rights, consent norms and obligations of data fiduciaries that intersect with AI training, profiling and misuse of biometric data.[15] Scholars therefore argue that India now faces moment to recalibrate criminalisation thresholds and penalties in light of this combined architecture and explicit deepfake proposals.

RESEARCH PROBLEM

A. Inadequacy and Fragmentation of Current Indian Legal Framework

Existing tools include provisions on cheating, identity impersonation, obscenity, defamation, voyeurism, criminal intimidation and cyber offences under IT Act and BNS. However, these provisions were drafted for simpler fraud, static images and conventional pornography, so they rarely capture layered nature of AI enabled composite harms.[16] Doctrinal writing consistently notes lack of statutory definition of deepfakes, synthetic media or AI face swapping, which complicates consistent classification of offences and defences.[17] Victims therefore depend on creative interpretation of general provisions, leading to uncertainty, under enforcement and uneven protection against similar deepfake incidents across jurisdictions.

B. Challenges in Attribution, Evidence, and Enforcement

Investigators face technical difficulty in detecting deepfakes, preserving metadata, and linking uploads to real world actors who often operate through VPNs and foreign servers.[18] Voice cloning scams or fake video calls may involve cross border routing and modular criminal chains, so attribution of liability amongst coders, sellers, uploaders and clients becomes very complex.[19] Courts must also decide when digital evidence is sufficiently authenticated, and how to treat defence claims that genuine videos are deepfakes or vice versa. Scholars have warned that if judiciary becomes over sceptical towards audio visual evidence, it can weaken prosecution of genuine crimes and reduce deterrence generally.[20]

C. Unclear Alignment between Penal Law and Data Protection Norms

The DPDP Act 2023 gives individuals rights over processing of personal data, including images, voice data and identifiers used to train or run generative models.[21] Criminal investigations meanwhile rely on compelled disclosure, search and seizure powers, which may sit uneasily with consent based data protection structure in deepfake cases. No clear statutory standard currently defines when platforms may process or share biometric and identity data for deepfake detection, without violating data minimisation

---

[12] Bharatiya Nyaya Sanhita 2023.
[13] Jaromir Cerny (n 3).
[14] Information Technology Act 2000 and Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules 2021.
[15] Digital Personal Data Protection Act 2023.
[16] Jaromir Cerny (n 3).
[17] Jaromir Cerny, 'The Admissibility of Deepfake Evidence within Context of Indian Law' (2025) 7(6) International Journal for Multidisciplinary Research (IJFMR) 1.
[18] Shinu Vig, 'Regulating Deepfakes: An Indian Perspective' (2024) 17(3) Journal of Strategic Security 70.
[19] Songyang Sai and Zifan Wang, 'Criminal Regulatory Approaches to Deepfake-Related Offenses: Focusing on Crime of Fraud' (2026) 3(1) International Journal of Asian Social Science Research 20.
[20] Jaromir Cerny (n 3).
[21] Digital Personal Data Protection Act 2023.

principles.[22] This tension between privacy and prosecution makes regulatory design for AI created harms especially delicate in Indian context.

D. Absence of Coherent Victim Centred Approach

Existing laws mainly focus on punishment and takedown but do not provide holistic remedies like rapid erasure, algorithmic downranking, digital hygiene support and long term counselling. Non consensual intimate deepfakes can live in shadow archives and resurfacing platforms for years, so impact on dignity, mental health and reputation is continuing, not one time.[23] Scholarly work on deepfakes and gender shows that women and marginalised persons face disproportionate targeting with limited access to effective compensation or rehabilitation mechanisms.[24] Thus, present framework still under values victim voice and fails to embed their interests at centre of criminal process and platform governance.

SCOPE AND LIMITATIONS OF THE STUDY

A. Substantive and Territorial Focus

The study focuses on Indian legal system and especially on interplay of BNS, BSA, IT Act and DPDP Act with proposed deepfake specific regulations.[25] Foreign and comparative material from EU, US and China is used only to derive possible models for reform, not as direct normative benchmark for Indian courts.[26]

B. Thematic Focus

The analysis concentrates on deepfakes, identity fraud, AI based impersonation and closely related harms like voice cloning scams, fake political videos and intimate image abuse.[27] Broader themes of AI governance, competition law, algorithmic transparency or copyright training datasets enter discussion only when they directly affect penal liability or victim protection.

C. Doctrinal and Normative Nature of Analysis

The research uses doctrinal method by examining statutes, case law, government reports and academic literature, and then builds normative arguments about desirable reform directions.[28] It does not conduct empirical surveys, interviews or quantitative analysis of case numbers, so factual patterns are drawn mainly from published studies and official data where available.

D. Temporal Limitations

The analysis reflects position after enactment of BNS, BSA and DPDP Act and in light of latest accessible deepfake proposals and scholarly debates. Subsequent judicial decisions, amendments or new AI specific regulations may therefore require updating of doctrinal arguments and policy recommendations in coming years.[29]

---

[22] *Ibid.*
[23] Prachi Agnihotri, 'Legal Implications of Deepfake Technology in Criminal Law' (2025) 8(1) International Journal of Law Management and Humanities 1645.
[24] Jaromir Cerny (n 3).
[25] *Ibid.*
[26] Shinu Vig, 'Regulating Deepfakes: An Indian Perspective' (2024) 17(3) Journal of Strategic Security 70.
[27] Songyang Sai and Zifan Wang, 'Criminal Regulatory Approaches to Deepfake-Related Offenses: Focusing on Crime of Fraud' (2026) 3(1) International Journal of Asian Social Science Research 20.
[28] Jaromir Cerny, 'The Admissibility of Deepfake Evidence within Context of Indian Law' (2025) 7(6) International Journal for Multidisciplinary Research (IJFMR) 1.
[29] Deepfake Prevention and Criminalisation Bill, 2023 (as introduced in Rajya Sabha).

RESEARCH OBJECTIVES

The researcher has formulated following research objectives:

1. To map existing and new penal provisions relevant to deepfakes, identity fraud, and AI created harms in India.
2. To critically examine adequacy of these provisions to address evolving technological abuses.
3. To analyse how Indian criminal law, IT law, and data protection norms can be harmonised for effective regulation.
4. To propose concrete legal and policy reforms for victim centred and technologically informed regulatory framework.

RESEARCH QUESTIONS

The researcher has formulated following research questions:

1. How do deepfakes, identity fraud, and AI created harms challenge traditional notions of offence, mens rea, and evidence in Indian criminal law?
2. To what extent do new penal provisions under BNS, along with IT Act and DPDP Act, respond to these challenges?
3. What are key doctrinal and practical gaps in current Indian framework on deepfake regulation and identity related offences?
4. What model of reform, including possible deepfake specific legislation or amendments, would best protect victims and ensure accountable AI use?

RESEARCH HYPOTHESES

The researcher has formulated following research hypotheses:

1. The new penal provisions under BNS, when read with IT Act and DPDP Act, are still insufficient to comprehensively regulate deepfakes and AI generated identity fraud.
2. Existing offences on cheating, impersonation, obscenity, and defamation do not fully capture layered harms caused by deepfakes and synthetic media.
3. A coherent and explicit statutory framework, including clear duties on platforms and AI developers, is necessary for effective prevention and redress.
4. Victim centred reforms, including speedy takedown, evidentiary support, and reparations, can be designed without unduly chilling legitimate expression and innovation.

RESEARCH METHODOLOGY

Doctrinal method guides whole study, so focus stays on legal texts and reasoned interpretation.[30] Primary sources include Bharatiya Nyaya Sanhita 2023, Bharatiya Sakshya Adhiniyam 2023, Information Technology Act 2000 and Digital Personal Data Protection Act 2023, read with Deepfake Prevention and Criminalisation Bill 2023.[31] Judicial decisions and reported Indian cases on identity misuse, online harassment and digital fakery are examined to trace emerging doctrinal patterns. Secondary sources consist of Indian and foreign scholarship on deepfakes, synthetic media, identity fraud and criminal law,

---

[30] Shinu Vig, 'Regulating Deepfakes: An Indian Perspective' (2024) 17(3) Journal of Strategic Security 70.
[31] The Deepfake Prevention and Criminalisation Bill 2023 (as introduced in Rajya Sabha, Bill No LXX of 2023).

including detailed Indian analysis of deepfake regulation and deepfake evidence.[32] Comparative material from European Union discussions on AI Act and deepfake transparency duties is used only to draw normative guidance, not to transplant whole frameworks.[33] Methodology therefore stays doctrinal, but it remains outward looking and comparative in spirit. Interpretation follows standard tools of statutory construction, constitutional reasoning and criminal law theory, particularly on harm, mens rea and protected interests.[34] The research also uses conceptual analysis to clarify notions like "deepfake", "digital content forgery", "synthetic media" and "AI created harms" by reading definitions in Indian Bill alongside foreign definitions.[35]

CONCEPTUAL FRAMEWORK OF DEEPFAKES, IDENTITY FRAUD, AND AI CREATED HARMS

1. Defining deepfakes, synthetic media, and AI generated content

Deepfakes are usually described as audio, visual or audiovisual media created or altered with advanced AI, which convincingly simulate reality.[36] They combine deep learning, neural networks and sophisticated image or sound processing, so ordinary viewers cannot easily detect manipulation in such content. Indian doctrinal work explains deepfakes as "manipulated or synthetic media" produced using artificial intelligence, which alters videos, audio or images to make someone appear to say or do things never done.[37] The stress lies on both technique and deceptive realism, not only on simple editing or filtering like old software tools.

The Deepfake Prevention and Criminalisation Bill 2023 gives statutory flavour to this idea, by defining deepfake as digitally manipulated or fabricated content created using AI or similar technologies, which deceptively depict persons or events that did not really exist.[38] It also introduces "digital content forgery" as deliberate use of such technologies to create altered audio, visual or textual material with intent to deceive, which is important for penal reasoning. Synthetic media in wider sense covers any media generated or heavily modified through AI, including entirely artificial faces, voices or bodies that do not correspond to real individuals.[39] AI generated content therefore includes deepfakes but also many benign or neutral outputs like filters, translation or upscaling tools, so legal focus must narrow to harmful uses.

European debates on AI Act show similar definitional path, describing deepfakes as manipulated or synthetic audio, image or video content that appears authentic and features persons doing or saying things they did not do, created using AI techniques.[40] Such foreign definitions support Indian understanding and help in sharpening terms without importing whole foreign frameworks blindly. Conceptual clarity matters

---

[32] Arjun Mehta, 'Deepfakes and Criminal Justice: Procedural Safeguards in India' (2022) 3 Journal of Law and Emerging Technology 55.

[33] Mateusz Łabuz, 'Regulating Deep Fakes in AI Act' (2023) 2(1) Artificial Intelligence and Governance.

[34] Bharatiya Nyaya Sanhita 2023; Bharatiya Sakshya Adhiniyam 2023; Information Technology Act 2000; Digital Personal Data Protection Act 2023.

[35] 'Criminal Regulatory Approaches to Deepfake Related Offenses: Focusing on Crime of Fraud' (2026) 3(1) International Journal of Asian Social Science Research 20.

[36] Shinu Vig, 'Regulating Deepfakes: An Indian Perspective' (2024) 17(3) Journal of Strategic Security 70.

[37] The Deepfake Prevention and Criminalisation Bill 2023 (as introduced in Rajya Sabha, Bill No LXX of 2023).

[38] Arjun Mehta, 'Deepfakes and Criminal Justice: Procedural Safeguards in India' (2022) 3 Journal of Law and Emerging Technology 55.

[39] Mateusz Łabuz, 'Regulating Deep Fakes in AI Act' (2023) 2(1) Artificial Intelligence and Governance.

[40] Bharatiya Nyaya Sanhita 2023; Bharatiya Sakshya Adhiniyam 2023; Information Technology Act 2000; Digital Personal Data Protection Act 2023.

because criminal law turns on precise description of conduct and harm, so vague notion of deepfakes can either over criminalise satire or under protect victims of severe abuse.[41]

2. **Forms of identity fraud and impersonation in digital spaces**

Identity fraud in AI context involves misuse of identifying data, likeness or credentials to deceive others for gain or to inflict harm. In deepfake environment, this includes face swapping, voice cloning, synthetic profile creation and real time video impersonation in financial or social interactions.[42] Scholars mapping deepfake enabled fraud speak of modular criminal chains, where one actor steals or buys personal data, another actor generates synthetic content, and third conducts scam using that content.[43] This modular structure weaken link between victim and direct perpetrator, which complicate attribution under traditional Indian criminal provisions on cheating and impersonation.

Face swapping fraud uses deepfake tools to overlay fraudster's face with that of manager, relative or public figure, often during live video calls used to instruct fund transfers or seek sensitive information.[44] Because victims feel they see known person in real time, usual warning signals of scam weaken, and law must treat such misuse of identity as aggravated cheating or personation. Another form involves AI voice cloning, where short audio samples allow creation of highly realistic synthetic speech which imitates pitch, accent and emotional tone of real person. Such voices can be used to instruct employees, reassure relatives, or issue fake emergency appeals, so harm goes far beyond ordinary anonymous phishing practices.[45]

Identity fraud also plays out through synthetic personas, where no real person exists but AI constructs consistent face, biography and behavioural pattern that gains trust over time. These synthetic agents can then promote misinformation, recruit for extremist causes, or pump and dump financial schemes by pretending to be genuine influencers or experts. From legal angle, important aspect is that identity fraud using deepfakes often simultaneously infringes personal information, privacy, property and sometimes bodily autonomy, for example when deepfake pornography uses stolen images.[46] Indian framework on cheating, identity theft and misuse of personal data therefore needs coordinated reading when dealing with such multi layered harm.

3. **Typology of AI related harms, from reputational and emotional harms to financial and democratic harms**

AI related harms associated with deepfakes can be grouped into several overlapping categories. First cluster contains reputational and emotional harms, often arising from non consensual intimate deepfakes, false confessions, or fabricated abusive statements attributed to victims.[47] Studies from Indian context show that women, journalists and activists face targeted deepfake campaigns aimed at shaming, silencing or discrediting them in public sphere. Such campaigns combine character assassination, doxing and coordinated trolling, so emotional trauma and social stigma persist even after content takedown.

Second cluster involves financial harms, where deepfakes act as instruments of fraud, extortion or market manipulation. Deepfake enabled fraud exploits information asymmetry, as victims trust highly realistic

[41] 'Criminal Regulatory Approaches to Deepfake Related Offenses: Focusing on Crime of Fraud' (2026) 3(1) International Journal of Asian Social Science Research 20.
[42] ibid.
[43] ibid 21–24.
[44] ibid 23.
[45] Prachi Agnihotri, 'Legal Implications of Deepfake Technology in Criminal Law' (2025) 8(1) International Journal of Law Management and Humanities 1645.
[46] Vig (n 1) 80–83.
[47] Vig (n 1) 82–88.

audio visual cues, which turns synthetic content into tool of property appropriation on scale.[48] These harms can appear in business email compromise, fake CEO calls, romance scams and investment schemes, so boundaries between cybercrime and traditional financial offences blur. Such typology matters because Indian penal provisions on cheating, breach of trust and forgery must be interpreted to cover deception built on AI content, not purely verbal misrepresentations.

Third category covers harms to democratic processes and public order, which arise when deepfakes mimic political leaders, election officials or community figures to spread inflammatory statements or false announcements. Foreign and Indian commentators point out that deepfakes can inflame communal tensions, incite violence, or depress voter participation by eroding trust in authentic political messages.[49] Such democratic harms rarely fit neatly into single offence, since they may involve combination of hate speech, incitement, defamation and unlawful assembly related risks. Therefore regulatory responses need broader view of protected legal interests, focusing on social trust, electoral integrity and peaceful public discourse alongside individual victims.

Finally, there are systemic harms to evidentiary reliability and institutional trust. When courts, police and public realise that audio and video can be easily forged, even genuine recordings may lose persuasive value, which threaten fair trial and effective enforcement.[50] Indian scholarship on deepfake evidence warns that over time, presence of deepfakes might support blanket denials by accused and suspicion against authentic digital exhibits. This "liar's dividend" could indirectly shield offenders and weaken deterrence, unless procedural safeguards and forensic capacities keep pace.

4. Constitutional and human rights dimensions, including privacy, dignity, and free speech

Constitutional analysis in India begins with recognition of privacy as fundamental right under Article 21 in K S Puttaswamy v Union of India. Deepfakes that use intimate images, private conversations or personal data without consent clearly invade informational and decisional privacy of individuals in digital age. Doctrinal writing on Indian deepfake regulation underlines that such content also attacks dignity, honour and autonomy, because they present false but convincing images of body and behaviour.[51] Unlike ordinary defamation, visual and audio realism of deepfakes make denial difficult and humiliation more intense, especially for women and minority groups.

From free speech angle, Article 19(1)(a) protects expression but is subject to reasonable restrictions in interests of public order, decency, morality, defamation and security of State. Regulatory design must therefore distinguish between satire, parody and artistic experimentation, and harmful impersonation or deceptive propaganda that crosses these constitutional limits.[52] European debates on AI Act and Digital Services Act show one approach, where transparency labelling, content marking and platform duties aim to reduce harm without banning technology outright. Indian scholars argue for similar graded measures, so freedom of expression survives but non consensual deepfakes, intimate imagery and fraud oriented content face stricter penal and civil responses.

Human rights concerns also extend to fair trial rights under Article 21, because introduction of deepfake evidence can affect presumption of innocence and reliability of proof. If manipulated recordings are wrongly admitted, wrongful convictions may occur, but if courts become overly sceptical, genuine victims

[48] 'Criminal Regulatory Approaches to Deepfake Related Offenses' (n 6) 23–28.
[49] Vig (n 1) 89–93; Łabuz (n 4) 24–29.
[50] Agnihotri (n 10) 1647–1653; Mehta (n 3) 60–63.
[51] Vig (n 1) 74–79; International Journal for Multidisciplinary Research, 'Deepfake Evidence and Judicial Accountability in India' IJFMR250660298 (2025).
[52] Łabuz (n 4) 27–30.

of online violence may fail to secure justice.[53] Data protection rights under DPDP Act 2023 intersect here, as biometric data like facial images and voiceprints qualify as sensitive personal information requiring strict consent and purpose limitation. Unlawful scraping of such data for training and deploying deepfake models may amount to both statutory breach and constitutional privacy violation, supporting combined civil and penal remedies.

International human rights instruments, including International Covenant on Civil and Political Rights, recognise rights to privacy, reputation and free expression, all relevant to deepfake harms. Indian courts often read these instruments as persuasive aids while interpreting Part III, so any new penal provisions on AI created harms must align with such broader human rights commitments.

MAPPING THE INDIAN LEGAL FRAMEWORK UNDER NEW PENAL PROVISIONS

1. Overview of Bharatiya Nyaya Sanhita, Bharatiya Sakshya Adhiniyam, and their relevance to AI crimes

Bharatiya Nyaya Sanhita 2023 replaces Indian Penal Code and consolidates general penal policy for digital as well as physical harms.[54] Text of BNS keeps traditional structure of general explanations, specific offences and punishment, but extends several concepts to electronic records and communications.[55] Definitions of "document" and related terms, carried forward in substance, now comfortably include electronic records, images, audio and video stored on digital devices.[56] This wide drafting allows courts to treat manipulated deepfake videos and synthetic audio as documents or records capable of supporting cheating, forgery or intimidation charges. Several BNS offences naturally intersect with AI crimes, even without using technological vocabulary. Provisions on cheating, cheating by personation, criminal intimidation, extortion, publishing sexually explicit material, and stalking can all apply where wrongful act is carried out through deepfake media rather than direct speech.[57] Scholars examining criminal law and deepfakes point out that BNS keeps focus on intention to deceive, harm or outrage modesty, not on specific tool used.[58] Therefore deepfake creator who fabricates obscene clip to blackmail victim may be charged similar to one who circulates secretly filmed genuine clip, though evidentiary issues differ.[59]

Bharatiya Sakshya Adhiniyam 2023 replaces Indian Evidence Act and modernises framework for proof of electronic records in criminal trials.[60] Analysis of deepfake evidence under BSA highlights that while electronic records remain admissible, questions about integrity, authenticity, and chain of custody become more pressing once hyper realistic manipulation is common.[61] The BSA provisions on electronic records, certificates, and presumptions are crucial when prosecution or defence relies on video or audio that might

---

[53] Agnihotri (n 10) 1650–1655; Mehta (n 3) 61–63.
[54] The Bharatiya Nyaya Sanhita 2023, No 45 of 2023 (Government of India, 25 December 2023).
[55] *Ibid.*
[56] *Ibid.*
[57] Aditya Pratap Singh, 'Legal Implications of Deepfake Technology in Criminal Law' (2025) 8(1) International Journal of Law Management and Humanities 1645.
[58] Anisha Bhatta, 'Deciphering Legal Regimes Governing Deep-Fakes' (2024) Indian Journal of Integrated Research in Law, vol V issue II, 1712.
[59] *Ibid.*
[60] Harmanjeet Singh and Ritu Panta, 'Deepfake Evidence and Indian Criminal Justice System: Challenges of Authenticity, Consent and Admissibility in Law' (2025) 7(6) International Journal for Multidisciplinary Research (IJFMR) 1.
[61] *Ibid.*

be challenged as deepfake.[62] Recent scholarship stresses that courts must adapt evidentiary standards and forensic practices so that genuine recordings are not casually dismissed as fake, yet manipulated content is rigorously screened.[63] Together, BNS and BSA set broad penal and evidentiary canvas on which specific AI harms, including deepfakes and identity fraud, will be litigated in coming years.[64] However, they still operate largely as technology neutral instruments, so doctrinal work must interpret their language in light of new digital risks and forensic capacities.[65]

Existing offences under IT Act and related rules relevant to deepfakes and identity fraud Information Technology Act 2000 remains central cyber law statute for India and it directly targets misuse of computer resources and electronic communication.[66] Although enacted long before deepfakes, its provisions on unauthorised access, data interference, identity theft, and publication of obscene material already cover many AI driven abuses when read purposively.[67] Section 66C criminalises identity theft based on fraudulent use of electronic signatures, passwords or unique identification features of another person. In context of AI crimes, this provision can catch unauthorised use of login credentials obtained through phishing or social engineering that then enable deepfake dissemination from victim accounts.[68]

Section 66D punishes cheating by personation using computer resources, which directly resonates with deepfake based impersonations on video calls or social media.[69] Where fraudster uses synthetic voice or face of company executive to instruct fund transfers, offence under section 66D can arise even though visual or audio identity is artificially generated. Sections 67, 67A and 67B address publication or transmission of obscene material, sexually explicit content, and child sexual abuse material in electronic form.[70] Indian deepfake literature rightly treats non consensual intimate deepfakes, including synthetic child abuse content, as falling within these provisions, since harm flows from sexualised representation regardless of fabrication.[71]

Intermediary liability framework under IT Act, elaborated by Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules 2021, also matters strongly.[72] These rules require intermediaries to remove unlawful content upon actual knowledge, maintain due diligence, and deploy reasonable efforts to prevent circulation of material that violates existing laws, which includes many deepfakes.[73] Scholars examining Indian deepfake regimes criticise current intermediary approach as reactive and notice based, not preventive in design.[74] They argue platforms should bear clearer duties to

[62] Harmanjeet Singh and Ritu Panta, 'Deepfake Evidence and Indian Criminal Justice System: Challenges of Authenticity, Consent and Admissibility in Law' (2025) 7(6) International Journal for Multidisciplinary Research (IJFMR) 1.
[63] *Ibid.*
[64] Aditya Pratap Singh, 'Legal Implications of Deepfake Technology in Criminal Law' (2025) 8(1) International Journal of Law Management and Humanities 1645.
[65] Anisha Bhatta, 'Deciphering Legal Regimes Governing Deep-Fakes' (2024) Indian Journal of Integrated Research in Law, vol V issue II, 1712.
[66] Shinu Vig, 'Regulating Deepfakes: An Indian Perspective' (2024) 17(3) Journal of Strategic Security 70.
[67] *Ibid.*
[68] Aditya Pratap Singh (n 2) 1650–1655.
[69] *Ibid.*
[70] Shinu Vig, 'Regulating Deepfakes: An Indian Perspective' (2024) 17(3) Journal of Strategic Security 70.
[71] Zhiwen Hao, 'Criminal Law Regulation of AI Face-Swapping Crimes: Responsible Subjects, Theoretical Integration and Comprehensive Governance Framework' (2023) Minzu University of China, School of Law working paper.
[72] 'Criminal Regulatory Approaches to Deepfake-Related Offenses: Focusing on Crime of Fraud' (2026) 3(1) International Journal of Asian Social Science Research 20.
[73] *Ibid.*
[74] The Deepfake Prevention and Criminalisation Bill 2023, Bill No LXX of 2023, as introduced in Rajya Sabha on 7 February 2025.

detect patterns of synthetic impersonation, to label manipulated content, and to cooperate with victims and law enforcement beyond bare minimum takedown.

Courts in India have already used blocking powers under section 69A IT Act for various harmful content online. In deepfake context, these powers allow government to block access to specific URLs or platforms distributing synthetic abuse, yet concerns about transparency and necessity of such measures remain.[75]

Overall, IT Act framework provides starting point for addressing computer facilitated identity fraud and obscene deepfakes, but it is stitched from general provisions rather than built for AI manipulation. This patchwork nature leads to overlaps with BNS offences and uneven enforcement response across states and platforms, as several commentators observe in their doctrinal analysis.[76]

2. **Role of DPDP Act and data protection norms in addressing AI created harms**

Digital Personal Data Protection Act 2023 introduces comprehensive regime for lawful processing of personal data in India, including data used for AI training and deployment.[77] Deepfakes and synthetic identity creations often rely on unauthorised scraping and processing of facial images, videos and voice samples, which fall squarely within scope of personal data under Act.[78] Scholars analysing constitutional dimension of deepfakes emphasise that manipulation of someone's digital likeness without consent violates privacy, autonomy, dignity and informational self determination under Article 21.[79] DPDP Act now translates those abstract rights into concrete obligations of data fiduciaries, such as consent, purpose limitation, data minimisation, and security safeguards for personal data processing.[80]

When platform or developer collects and processes biometric data to train generative models, it must secure valid consent or rely on limited legitimate uses recognised by statute. Using such models later to create targeted deepfakes, particularly intimate or defamatory ones, may constitute both breach of statutory obligations and infringement of constitutional privacy values.[81] DPDP Act also mandates notice to data principals about purposes and categories of personal data processing. If organisation silently scrapes public profile pictures or videos for deepfake training without meaningful notice, that practice may breach transparency and fairness requirements even before any specific offence occurs.[82]

However, DPDP Act focuses on regulatory penalties and compensation rather than imprisonment, so it does not directly displace criminal provisions under BNS or IT Act.[83] Instead, data protection norms supply background duties and standards that can help courts judge reasonableness of conduct and severity of culpability in AI related offences. One challenge lies in aligning investigative needs with privacy protections. Collecting training logs, model weights, and datasets for proving deepfake origin may require

---

[75] Aditya Pratap Singh (n 2) 1650–1655.
[76] Anisha Bhatta, 'Deciphering Legal Regimes Governing Deep-Fakes' (2024) Indian Journal of Integrated Research in Law, vol V issue II, 1712.
[77] Digital Personal Data Protection Act 2023 (No 22 of 2023).
[78] *Ibid.*
[79] Juhi Chandel and Manisha Kundu, 'AI-Generated Deepfakes and Legal Vacuum in India: A Constitutional Analysis of Privacy, Consent, and Digital Harm under Article 21' (2025) 10(11) International Journal for Research Trends and Innovation (IJRTI) a850.
[80] Digital Personal Data Protection Act 2023 (No 22 of 2023).
[81] Juhi Chandel and Manisha Kundu, 'AI-Generated Deepfakes and Legal Vacuum in India: A Constitutional Analysis of Privacy, Consent, and Digital Harm under Article 21' (2025) 10(11) International Journal for Research Trends and Innovation (IJRTI) a850.
[82] Digital Personal Data Protection Act 2023 (No 22 of 2023).
[83] *Ibid.*

wide access to personal data, so law must balance due process and data minimisation principles carefully.[84]

Commentators suggest that DPDP Act can support victim centric remedies such as data erasure, restriction of processing and accountability for onward sharing, which are very relevant once deepfake harms occur.[85] Yet these remedies need coordination with criminal process, notice to affected persons, and technical capacity on part of regulators and courts to implement meaningful deletion or de indexing. Indian debate on deepfakes therefore increasingly treats DPDP Act as complementary pillar to penal law. It supplies language of consent, lawful purpose and digital harm that helps articulate why certain AI uses are not just offensive or immoral, but legally wrongful in structured way.[86]

Proposed deepfake specific legislative initiatives and policy papers in India Deepfake Prevention and Criminalisation Bill 2023, introduced in Rajya Sabha, represents first serious attempt to craft bespoke statute for synthetic media harms in India.[87] The Bill defines "deepfake" as digitally manipulated or fabricated content, including images, videos or audio created with advanced technologies, with intent to deceptively depict subjects or issues.[88] It also defines "digital content forgery" as use of technologies such as artificial intelligence and machine learning to create or alter audio, visual or textual content with purpose of deceiving viewers.[89] These definitions are wider than conventional notions of forgery or obscenity, because they focus on manipulation process and deceptive realism rather than only on subject matter.

Bill's core policy choice is to criminalise creation, dissemination and use of deepfake content without consent or without digital watermark.[90] By tying legality to presence of consent or visible watermark, drafters attempt to separate malicious impersonation and covert abuse from legitimate uses like satire, education or cinema where informed consent or labelling exist. The proposal envisages role for Task Force to recommend regulatory guidelines, penalties and technical safeguards, including possible visual protection features and blockchain based authenticity tools.[91] Scholars see value in such multi stakeholder mechanism, though they caution that vague delegation should not replace clear legislative standards and procedural safeguards.[92]

Indian academic writing on deepfakes, across several journals, consistently urges move from scattered reliance on IT Act and BNS towards integrated framework that recognises unique nature of AI created harms.[93] These works highlight gaps in victim support, insufficiency of existing offences for political deepfakes, and need for obligations on platforms and AI developers to detect and label manipulated

---

84 Juhi Chandel and Manisha Kundu, 'AI-Generated Deepfakes and Legal Vacuum in India: A Constitutional Analysis of Privacy, Consent, and Digital Harm under Article 21' (2025) 10(11) International Journal for Research Trends and Innovation (IJRTI) a850.
85 *Ibid.*
86 *Ibid.*
87 The Deepfake Prevention and Criminalisation Bill 2023, Bill No LXX of 2023, as introduced in Rajya Sabha on 7 February 2025.
88 *Ibid.*
89 The Deepfake Prevention and Criminalisation Bill 2023, Bill No LXX of 2023, as introduced in Rajya Sabha on 7 February 2025.
90 *Ibid.*
91 The Deepfake Prevention and Criminalisation Bill 2023, Bill No LXX of 2023, as introduced in Rajya Sabha on 7 February 2025.
92 Shinu Vig, 'Regulating Deepfakes: An Indian Perspective' (2024) 17(3) Journal of Strategic Security 70.
93 Anisha Bhatta, 'Deciphering Legal Regimes Governing Deep-Fakes' (2024) Indian Journal of Integrated Research in Law, vol V issue II, 1712.

content.[94] Comparative studies, including work on AI face swapping crimes in China and on EU regulatory approaches, feed into Indian policy debate by showing possible models for responsibility allocation.[95] They argue for combined approach where criminal law, platform duties, technical standards and public awareness operate together, instead of expecting penal provisions alone to solve systemic deepfake risks.[96]

Indian scholars also emphasise that any deepfake specific statute must respect constitutional commitments to privacy and free speech, as elaborated in Justice K S Puttaswamy v Union of India and related jurisprudence.[97] That means precise drafting of offences, safeguards against over broad censorship, and strong procedural protections for accused in cases where authenticity of contested media becomes central issue. Policy conversation in India is therefore shifting toward comprehensive regulation of deepfakes, identity fraud and AI created harms, anchored in new penal provisions but drawing guidance from data protection and human rights perspectives.[98] Deepfake Prevention and Criminalisation Bill 2023, along with rich academic commentary, now provides concrete starting point for Parliament and regulators to refine such framework in coming years.[99]

DOCTRINAL AND PRACTICAL CHALLENGES IN REGULATION AND ENFORCEMENT

1. Mens rea, attribution, and multiple actors in AI driven offences

Deepfake crimes often arise from long technological chains, not from one lonely wrongdoer.[100] Model creators, platform providers, app developers, prompt engineers and end users all participate at different stages of harmful content lifecycle.[101] Traditional offences in Bharatiya Nyaya Sanhita like cheating, impersonation, forgery and defamation still expect clear mental element in one identified accused.[102] When one person codes model, another hosts service, third uploads prompt and hundreds amplify content, locating mens rea becomes doctrinally messy.

Section 316 and 319 BNS treat cheating and impersonation as individualised conduct directed at specific victim or group.[103] Deepfake financial fraud often uses cloned executive voices and synthetic videos generated abroad, so intention to cheat is spread across multiple human and algorithmic choices.[104] Defamation under Sections 356(1) and 356(2) BNS presumes identifiable author of imputation.[105] Deepfake clip may be automatically re-encoded, re-captioned and re-posted by bots, so court struggles to decide whose mind carried animus injuriandi.

[94] Aditya Pratap Singh (n 2) 1650–1655.
[95] Zhiwen Hao, 'Criminal Law Regulation of AI Face-Swapping Crimes: Responsible Subjects, Theoretical Integration and Comprehensive Governance Framework' (2023) Minzu University of China, School of Law working paper.
[96] *Ibid.*
[97] Juhi Chandel and Manisha Kundu, 'AI-Generated Deepfakes and Legal Vacuum in India: A Constitutional Analysis of Privacy, Consent, and Digital Harm under Article 21' (2025) 10(11) International Journal for Research Trends and Innovation (IJRTI) a850.
[98] Shinu Vig, 'Regulating Deepfakes: An Indian Perspective' (2024) 17(3) Journal of Strategic Security 70.
[99] The Deepfake Prevention and Criminalisation Bill 2023, Bill No LXX of 2023, as introduced in Rajya Sabha on 7 February 2025.
[100] Aditya Pratap Singh, 'Legal Implications of Deepfake Technology in Criminal Law' (2025) 8(1) International Journal of Law Management and Humanities 1645.
[101] *Ibid.*
[102] *Ibid.*
[103] Harmanjeet Singh and Ritu Panta, 'Deepfake Evidence and Indian Criminal Justice System: Challenges of Authenticity, Consent and Admissibility in Law' (2025) 7(6) International Journal for Multidisciplinary Research 1.
[104] *Ibid.*
[105] *Ibid.*

Forgery provisions such as Section 336 BNS contemplate conscious making of false document or electronic record.[106] In deepfake pipeline, some intermediaries may only supply datasets or plug-ins, yet their tools materially shape deceptive output without clear direct intention toward particular victim. Scholars therefore argue that doctrinal tests of abetment, common intention or criminal conspiracy do not sit comfortably with modular AI ecosystems.[107]

Attribution models that treat each participant as joint perpetrator risk over-criminalisation, while narrow focus on uploader alone leaves serious impunity gaps. Indian constitutional jurisprudence on criminal statutes requires clarity and foreseeability of penal liability.[108] If data annotators or generic model providers cannot reasonably foresee that their tools will be used for specific pornographic or electoral deepfake, attaching mens rea becomes normatively doubtful.

The difficulty of attribution is compounded by anonymity and rapid dissemination across platforms and devices.[109] Even where police identify device used to upload deepfake, it may be impossible to prove who actually engineered core synthetic manipulation at earlier stage. Some authors suggest graded liability models, where knowing deployment of deepfake tools for clearly unlawful purposes attracts primary liability, while reckless failure to implement safeguards triggers secondary liability.[110] Such proposals however still depend on robust factual proof of knowledge or recklessness across many technical actors, which present criminal procedure is not designed to collect.

2. Evidentiary challenges: detection, authentication, and chain of custody

Deepfake evidence directly destabilises assumptions of authenticity, integrity and reliability that underpin Bharatiya Sakshya Adhiniyam 2023.[111] Indian courts already struggle with ordinary electronic records, and synthetic media raises still sharper doubts about what photographs or videos really depict.[112] The Deepfake Evidence and Indian Criminal Justice System study notes that deepfakes are intended to look more convincing than genuine footage.[113] Conventional visual inspection or simple metadata checks usually fail, so judges and lawyers may misread fabricated clip as reliable corroboration.

BSA incorporates framework analogous to former Sections 65A and 65B Evidence Act for admissibility of electronic records.[114] Certificates regarding manner of production, device and integrity presume that underlying file is not algorithmically fabricated from start. Where complainant produces explicit deepfake video as proof of non consensual pornography, court needs far more than standard authenticity certificate.[115]

Forensic experts must identify traces of GAN based synthesis, artifacts in face blending, lip sync anomalies and model fingerprints, which many state labs currently lack capacity to do. BNSS procedural rules anticipate reliance on digital forensics but also warn about risks of manipulation during

[106] *Ibid.*
[107] Aditya Pratap Singh, 'Legal Implications of Deepfake Technology in Criminal Law' (2025) 8(1) International Journal of Law Management and Humanities 1645.
[108] Shreya Singhal v Union of India (2015) 5 SCC 1.
[109] Information Technology Act 2000.
[110] Aditya Pratap Singh, 'Legal Implications of Deepfake Technology in Criminal Law' (2025) 8(1) International Journal of Law Management and Humanities 1645.
[111] Harmanjeet Singh and Ritu Panta, 'Deepfake Evidence and Indian Criminal Justice System: Challenges of Authenticity, Consent and Admissibility in Law' (2025) 7(6) International Journal for Multidisciplinary Research 1.
[112] *Ibid.*
[113] *Ibid.*
[114] State of Maharashtra v Praful B Desai (2003) 4 SCC 601.
[115] *Ibid.*

investigation.[116] Because deepfake material can be edited between seizure and trial, defence may legitimately demand independent expert review, which slows proceedings and complicates chain of custody.

Scholars emphasise that every step from initial online capture to forensic copy must be carefully logged, hashed and preserved.[117] Any gap in hash values, storage records or access logs allows defence to suggest that prosecution itself may have altered or even created synthetic image. Indian jurisprudence on electronic evidence already insists on rigorous authentication.

In State of Maharashtra v Praful B Desai Supreme Court accepted video conferencing evidence but stressed need for procedural safeguards around recording and transmission.[118] In Anvar P V v P K Basheer Court treated Section 65B certificate as mandatory gateway for secondary electronic evidence.[119] Deepfake disputes will likely demand still higher threshold, including expert testimony on AI manipulation techniques and demonstration of forensic tools used.

The Shreya Singhal decision recognised dangers of overbroad regulation of online speech but did not confront synthetic media.[120] Courts now face double bind, where distrust of digital evidence may protect accused from wrongful conviction, yet also prevent victims from proving authentic sexual violence or extortion recordings. Scholars propose statutory amendments to BSA introducing express category of AI generated evidence and requirement of authenticity verification certificate from accredited forensic laboratories.[121] Such certification could become linchpin for admissibility, while leaving room for defence cross examination on technical methodology and error margins.

3. Jurisdictional and cross border enforcement issues

Deepfake and online identity fraud routinely operate across territories, while Indian penal law still anchored in territorial notions of occurrence.[122] Synthetic videos targeting Indian voters or investors may be generated on servers abroad, routed through foreign platforms and funded via overseas wallets. The criminal context analysis notes that deepfake financial fraud often replicates voices of senior officials to induce international fund transfers.[123]

Where victim bank in India acts on instruction, but manipulated call originates outside India, investigation faces complex conflict of laws and mutual legal assistance hurdles.[124] Political disinformation deepfakes also circulate widely on messaging platforms whose parent companies sit in other jurisdictions.[125] Obtaining subscriber details, training data logs or internal moderation records then depends on cross border cooperation regimes and platform policies, rather than on direct sovereign control.

Jurisdiction clauses in IT Act and general criminal law enable some extraterritorial reach, particularly where computer resources located in India or effect felt in India.[126] However, practical enforcement still depends on willingness of foreign states and corporations to respond quickly to Indian requests, which

[116] State of Maharashtra v Praful B Desai (2003) 4 SCC 601.
[117] Harmanjeet Singh and Ritu Panta, 'Deepfake Evidence and Indian Criminal Justice System: Challenges of Authenticity, Consent and Admissibility in Law' (2025) 7(6) International Journal for Multidisciplinary Research 1.
[118] State of Maharashtra v Praful B Desai (2003) 4 SCC 601.
[119] Anvar PV v PK Basheer (2014) 10 SCC 473.
[120] Shreya Singhal v Union of India (2015) 5 SCC 1.
[121] *Ibid.*
[122] Information Technology Act 2000.
[123] Anvar PV v PK Basheer (2014) 10 SCC 473.
[124] *Ibid.*
[125] Information Technology Act 2000.
[126] *Ibid.*

often comes too late for electoral or reputational harm. Scholars describe time lag between detection of morphed content, filing of legal action and grant of court order as major obstacle.[127] By time blocking order reaches overseas platform or mirror sites, deepfake may already be downloaded, reshared and embedded into countless private devices.

The Regulating Deepfakes study highlights challenge of identifying original uploader and tracing path of dissemination across platforms.[128] Even when first upload occurred from Indian IP address, use of VPNs, temporary accounts and encrypted messaging creates evidentiary gaps that weaken prosecution narrative. Cross border evidence gathering also interacts uneasily with privacy and data protection debates. Requests for access to platform logs, IP addresses and biometric datasets must be balanced with privacy guarantees under DPDP Act and under foreign data protection regimes.[129] Without clear bilateral and multilateral protocols tailored to AI generated harms, Indian authorities risk both under-enforcement and diplomatic friction.[130] Victims meanwhile experience continuing availability of harmful content abroad even after partial domestic takedown, which undermines faith in criminal justice.

4. Platform and intermediary liability, safe harbour, and obligations of AI developers

Information Technology Act and its Intermediary Guidelines framework represent central regulatory pillar for platforms that host or distribute deepfake content.[131] Sections 66C, 66D and 66E criminalise identity theft, cheating by personation and violation of privacy through electronic means, offering hooks against deepfake creators and sharers.[132] However IT Act deals only indirectly with synthetic media and retains safe harbour in Section 79 for intermediaries acting as mere conduits.[133]

Regulating Deepfakes article notes that intermediaries are expected to remove illegal content once notified or ordered, yet their deeper algorithmic role remains under-defined.[134] Section 66E and 66D have been invoked for deepfake pornography and impersonation, while 66F addresses cyber terrorism in extreme disinformation scenarios.[135] Still, penalties and offence definitions were not drafted with AI generated, hyper-realistic content in mind, so they ignore unique scale and persistence of harm.

The same study discusses Section 79 safe harbour and Delhi High Court decision in MySpace Inc v Super Cassettes Industries Ltd, where platform was required to take down infringing content upon notice even without prior court order.[136] This reasoning suggests that platforms hosting deepfakes may bear duty of rapid removal once they receive credible complaint or knowledge.

Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules 2021 increase responsibilities for significant social media intermediaries.[137] They must appoint grievance officers, enable traceability in certain contexts and act expeditiously on user complaints or government directions about unlawful content. In November 2023, Government instructed social media intermediaries to remove morphed videos or deepfakes within twenty four hours of complaint under IT Rules 2021.[138] Directive

[127] Mateusz Łabuz, 'Regulating Deep Fakes in Artificial Intelligence Act' (2023) 2(1) ACIG.
[128] *Ibid.*
[129] Digital Personal Data Protection Act 2023.
[130] Aditya Pratap Singh, 'Legal Implications of Deepfake Technology in Criminal Law' (2025) 8(1) International Journal of Law Management and Humanities 1645.
[131] Information Technology Act 2000.
[132] State of Maharashtra v Praful B Desai (2003) 4 SCC 601.
[133] Shreya Singhal v Union of India (2015) 5 SCC 1.
[134] *Ibid.*
[135] *Ibid.*
[136] MySpace Inc v Super Cassettes Industries Ltd (2017) 236 DLT 478 (Del).
[137] Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules 2021.
[138] Shinu Vig, 'Regulating Deepfakes: An Indian Perspective' (2024) 17(3) Journal of Strategic Security 70.

followed high profile non consensual deepfakes of actors and reflects move toward strict notice-and-takedown standard in India.[139]

Scholars emphasise that identification of uploader and speed of viral spread still limit effectiveness of takedown based models.[140] Even with strict deadlines, deepfakes can inflict lasting reputational, psychological and financial damage before platforms complete internal verification and removal. Legal Implications article shows how MeitY now experiments with watermarking, AI based fact checking and authenticity mechanisms within new iterations of IT Rules.[141] CERT-In advisory in 2024 urges platforms to deploy AI detection tools, collaborate with cybersecurity firms and report deepfake incidents to authorities.

DPDP Act 2023 introduces consent based processing model for personal data, including biometric identifiers such as facial images and voice samples.[142] Where platform or AI developer processes such data without lawful consent to generate harmful deepfake, it may face regulatory penalties from Data Protection Board alongside criminal exposure. Yet DPDP Act does not expressly govern secondary or synthetic reuse of lawful data in AI training and inference, creating grey zone around responsibility of AI model providers.[143] Absent clear rules, platforms may blame upstream developers, while developers claim that they never targeted specific individual victim and only provided general purpose tools.

Constitutional jurisprudence under Article 21 on privacy and dignity, and under Article 19(1)(a) on free speech, further complicates platform liability debates.[144] Overbroad duties to monitor and filter content can chill legitimate satire, political commentary and artistic uses of synthetic media, whereas under-regulation erodes dignity and informational self determination of individuals. Comparative scholarship notes that EU Artificial Intelligence Act moves toward risk based obligations for high risk and general purpose AI systems, including transparency for deepfake content.[145] Indian literature suggests similar layered duties on AI developers and platforms, such as provenance labelling, watermarking, auditable logs and mandatory reporting of malicious use.

Courts and regulators therefore confront difficult calibration problem. If safe harbour is narrowed too drastically, smaller platforms and open source developers may exit market or resort to heavy handed censorship, but if it remains too broad, victims of deepfake harms will continue to face remediless online abuse.

COMPARATIVE PERSPECTIVES AND REFORM PROPOSALS FOR INDIA

1. Lessons from foreign criminal and regulatory approaches to deepfakes and AI harms

Foreign jurisdictions increasingly treat deepfakes as composite technological and social harms rather than isolated cyber offences. Regulatory focus shifts from intent of single offender to ecosystem responsibility. United States law addresses deepfakes through fragmented federal and state interventions, not through comprehensive criminal code reform. California Civil Code targets election related and non consensual sexual deepfakes, emphasising injunctions over imprisonment.[146] Virginia criminalises malicious

---

[139] Shinu Vig, 'Regulating Deepfakes: An Indian Perspective' (2024) 17(3) Journal of Strategic Security 70.
[140] Mateusz Łabuz, 'Regulating Deep Fakes in Artificial Intelligence Act' (2023) 2(1) ACIG.
[141] Aditya Pratap Singh, 'Legal Implications of Deepfake Technology in Criminal Law' (2025) 8(1) International Journal of Law Management and Humanities 1645.
[142] Digital Personal Data Protection Act 2023.
[143] *Ibid.*
[144] Justice KS Puttaswamy (Retd) v Union of India (2017) 10 SCC 1.
[145] Mateusz Łabuz, 'Regulating Deep Fakes in Artificial Intelligence Act' (2023) 2(1) ACIG.
[146] California Civil Code §1708.86 (United States).

dissemination of synthetic pornography, reflecting victim protection priority rather than abstract moral condemnation. Statutes still require proof of knowledge, which remains difficult where tools automate content generation. Federal proposals like DEEPFAKES Accountability Act stress disclosure and provenance duties, not broad criminalisation. Legislators recognise chilling effect risk on political satire and artistic expression. United Kingdom adopts platform centric approach under Online Safety Act 2023, imposing duties of care on intermediaries hosting manipulated media. Law focuses on systemic risk management rather than attributing mens rea to upstream developers.[147]

UK criminal law still relies on traditional offences like harassment and fraud. Deepfake harm addressed indirectly through enhanced platform compliance obligations. European Union advances more integrated framework through Digital Services Act and Artificial Intelligence Act. AI Act classifies deepfake systems as limited risk, imposing transparency and labelling obligations.[148] EU model avoids blanket criminal liability for model developers. Instead regulatory sanctions and compliance audits operate as primary enforcement tools. Digital Services Act imposes notice and action duties with strict timelines. Large platforms face risk assessments for systemic dissemination of manipulated media.[149] Comparative scholarship suggests criminal law remains blunt instrument for synthetic media harms. Regulatory law provides flexibility and faster response mechanisms. These models demonstrate preference for ex ante governance rather than ex post punishment. Indian law still leans heavily on post harm criminal prosecution. Foreign approaches also emphasise civil remedies, administrative fines and takedown orders. Victim relief often prioritised over symbolic penal severity.

2. Evaluating suitability of EU style transparency and labelling duties in Indian context

EU style labelling duties require clear disclosure when content is AI generated or materially manipulated. Objective is preventing deception rather than suppressing speech.[150] Indian constitutional framework under Article 19 demands narrow tailoring of speech restrictions. Mandatory labelling could satisfy proportionality if content neutral and purpose driven. Transparency duties align with Indian Supreme Court emphasis on informed autonomy and dignity. Privacy jurisprudence under Justice K S Puttaswamy supports informational self determination.[151] However enforcement capacity poses challenge. Many Indian platforms lack technical ability to deploy reliable watermarking or provenance tools.

Open source AI models complicate compliance. Developers often lack control over downstream deployments or user modifications. EU regime differentiates between deployers and developers. Similar distinction needed in Indian regulatory design. Mandatory labelling may conflict with satire and parody traditions. Over inclusive disclosure rules risk chilling creative political commentary. Indian electoral context heightens urgency. Deepfake political speech threatens democratic deliberation more acutely due to linguistic diversity and rapid messaging circulation. Labelling duties may function effectively during elections through temporary enhanced obligations. Election Commission advisories already move in this direction. EU experience shows labelling insufficient once content goes viral. Users often ignore or misunderstand disclaimers. Thus transparency must combine with algorithmic demotion and rapid takedown mechanisms. Indian IT Rules partially enable such responses. Adopting EU style duties requires

[147] Online Safety Act 2023 (United Kingdom).
[148] Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act).
[149] Regulation (EU) 2022/2065 on Single Market for Digital Services (Digital Services Act).
[150] Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act).
[151] Justice K S Puttaswamy (Retd) v Union of India (2017) 10 SCC 1.

phased implementation and regulatory sandboxing. Immediate blanket mandates may burden smaller intermediaries unfairly.

3. Proposals for specific statutory amendments or dedicated deepfake and identity fraud statute

Indian Penal Code successors like Bharatiya Nyaya Sanhita address deception and impersonation but not synthetic fabrication explicitly. Statutory clarity remains inadequate. Amendments could introduce definition of synthetic media and deepfake manipulation. Clear statutory language improves foreseeability and constitutional validity. Specific offence of malicious creation or knowing dissemination of harmful deepfake may be inserted. Mens rea threshold should require knowledge or reckless disregard. Separate aggravated offences may apply where deepfakes target sexual privacy, elections, or financial systems. Existing Sections 66C and 66D IT Act partially cover identity fraud.[152] Procedural law under Bharatiya Sakshya Adhiniyam requires amendment recognising AI generated evidence. Special authentication certificates by accredited laboratories necessary.

A dedicated Deepfake and Identity Fraud Act could consolidate criminal, civil and regulatory provisions. Fragmented regulation currently confuses enforcement agencies. Such statute should incorporate graded liability. Primary creators bear higher culpability than passive intermediaries. Safe harbour under Section 79 IT Act should remain but with conditional compliance duties. Failure to act on credible notice should trigger liability. Data Protection Digital Personal Data Act should clarify obligations for biometric misuse in AI training. Secondary synthetic reuse requires explicit regulation.[153] Regulatory oversight may vest in MeitY with technical advisory board. Criminal prosecution should remain last resort. Special investigative powers for digital forensics units required. Capacity building essential for meaningful enforcement.

4. Designing victim centred model, including takedown, rectification, compensation, and restorative remedies

Victims of deepfakes suffer reputational psychological and economic harm. Criminal conviction alone rarely restores dignity or autonomy. Immediate takedown constitutes most urgent remedy. IT Rules 2021 mandate prompt removal upon complaint.[154] Yet notice based systems often delay relief. Automated detection and trusted flagger mechanisms necessary. Rectification includes correction notices and platform amplification of clarifications. Such remedies counter lingering misinformation effects. Compensation mechanisms remain underdeveloped. Victims rarely receive monetary redress despite severe harm. Statutory compensation funds similar to victim compensation schemes may be extended. Funding could derive from regulatory fines on non compliant platforms. Restorative remedies include apology orders and content provenance disclosures. These measures acknowledge harm publicly. Civil remedies under tort law remain slow and expensive. Streamlined digital tribunals could offer faster adjudication. Anonymity protections for complainants essential in sexual deepfake cases. Fear of further exposure deters reporting. Cross border takedown cooperation crucial. Without international coordination victims face perpetual online victimisation. Victim centric approach aligns with Article 21 dignity jurisprudence. Law must prioritise lived harm not only doctrinal purity.

CONCLUSION, SUGGESTIONS AND WAY FORWARD

1. Consolidated findings from doctrinal and comparative analysis

Indian criminal law struggles with attribution and mens rea in modular AI ecosystems. Comparative

[152] Information Technology Act 2000, ss 66C, 66D.
[153] Digital Personal Data Protection Act 2023 (India).
[154] Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules 2021.

regimes adopt regulatory and civil solutions. Foreign models prioritise platform duties transparency and risk management. Criminal law used selectively for egregious conduct. Evidentiary frameworks require urgent modernisation to address synthetic manipulation. Existing electronic evidence rules insufficient. Victim interests often sidelined in punitive frameworks. Comparative systems emphasise remediation and rapid relief.

2. Assessment of whether current and proposed laws meet constitutional and policy goals

Current Indian laws partially address deepfake harms through IT Act and BNS. Gaps undermine effectiveness and foreseeability. Constitutional proportionality requires precise offences and limited speech intrusion. Over broad criminalisation risks Shreya Singhal type invalidation.[155] Regulatory approaches better satisfy least restrictive means principle. Transparency and takedown duties align with free speech balance. Policy goals of trust dignity and electoral integrity remain unmet under existing framework.

3. Final recommendations for legislators, regulators, platforms, and legal practitioners

Legislators should enact targeted amendments or dedicated statute addressing deepfakes. Clarity and graded liability essential. Regulators must issue technical standards for watermarking detection and forensic certification. Capacity building should be prioritised. Platforms should invest in detection tools and rapid grievance handling. Proactive compliance reduces reputational and legal risk. Legal practitioners need technical literacy to litigate deepfake cases effectively. Interdisciplinary training necessary.

4. Future research directions in AI criminal law and deepfake governance

Further research needed on algorithmic attribution models and evidentiary thresholds. Comparative empirical studies remain limited. Impact of deepfakes on marginalized communities warrants focused inquiry. Gendered harms often under reported. International cooperation frameworks for synthetic media enforcement require development. Bilateral protocols remain inadequate. Long term governance must integrate ethics technical design and law. Static statutes may fail against evolving AI capabilities.

REFERENCES / BIBLIOGRAPHY

1. Cases

a. Anvar PV v PK Basheer (2014) 10 SCC 473.
b. Justice KS Puttaswamy (Retd) v Union of India (2017) 10 SCC 1.
c. MySpace Inc v Super Cassettes Industries Ltd (2017) 236 DLT 478 (Del).
d. Shreya Singhal v Union of India (2015) 5 SCC 1.
e. State of Maharashtra v Praful B Desai (2003) 4 SCC 601.

2. Legislation and Bills (India)

a. Bharatiya Nyaya Sanhita 2023 (No 45 of 2023).
b. Bharatiya Sakshya Adhiniyam 2023 (No 46 of 2023).
c. Digital Personal Data Protection Act 2023 (No 22 of 2023).
d. Information Technology Act 2000.
e. Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules 2021.
f. The Deepfake Prevention and Criminalisation Bill 2023, Bill No LXX of 2023, as introduced in Rajya Sabha on 7 February 2025.

[155] Shreya Singhal v Union of India (2015) 5 SCC 1.

3. Foreign and International Instruments
a. California Civil Code §1708.86 (United States).
b. Online Safety Act 2023 (United Kingdom).
c. Regulation (EU) 2022/2065 of European Parliament and of Council of 19 October 2022 on Single Market for Digital Services (Digital Services Act).
d. Regulation (EU) 2024/1689 of European Parliament and of Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act).

4. Secondary Sources
a. Agnihotri P, 'Legal Implications of Deepfake Technology in Criminal Law' (2025) 8(1) International Journal of Law Management and Humanities 1645.
b. Bhatta A, 'Deciphering Legal Regimes Governing Deep-Fakes' (2024) Indian Journal of Integrated Research in Law, vol V issue II, 1712.
c. Cerny J, 'The Admissibility of Deepfake Evidence within Context of Indian Law' (2025) 7(6) International Journal for Multidisciplinary Research (IJFMR) 1.
d. Chandel J and Kundu M, 'AI-Generated Deepfakes and Legal Vacuum in India: A Constitutional Analysis of Privacy, Consent, and Digital Harm under Article 21' (2025) 10(11) International Journal for Research Trends and Innovation (IJRTI) a850.
e. 'Criminal Regulatory Approaches to Deepfake Related Offenses: Focusing on Crime of Fraud' (2026) 3(1) International Journal of Asian Social Science Research 20.
f. Hao Z, 'Criminal Law Regulation of AI Face-Swapping Crimes: Responsible Subjects, Theoretical Integration and Comprehensive Governance Framework' (2023) Minzu University of China, School of Law working paper.
g. International Journal for Multidisciplinary Research, 'Deepfake Evidence and Judicial Accountability in India' IJFMR250660298 (2025).
h. Łabuz M, 'Regulating Deep Fakes in Artificial Intelligence Act' (2023) 2(1) Artificial Intelligence and Governance.
i. Łabuz M, 'Regulating Deep Fakes in Artificial Intelligence Act' (2023) 2(1) ACIG.
j. Mehta A, 'Deepfakes and Criminal Justice: Procedural Safeguards in India' (2022) 3 Journal of Law and Emerging Technology 55.
k. Sai S and Wang Z, 'Criminal Regulatory Approaches to Deepfake-Related Offenses: Focusing on Crime of Fraud' (2026) 3(1) International Journal of Asian Social Science Research 20.
l. Singh A P, 'Legal Implications of Deepfake Technology in Criminal Law' (2025) 8(1) International Journal of Law Management and Humanities 1645.
m. Singh H and Panta R, 'Deepfake Evidence and Indian Criminal Justice System: Challenges of Authenticity, Consent and Admissibility in Law' (2025) 7(6) International Journal for Multidisciplinary Research 1.
n. Vig S, 'Regulating Deepfakes: An Indian Perspective' (2024) 17(3) Journal of Strategic Security 70.