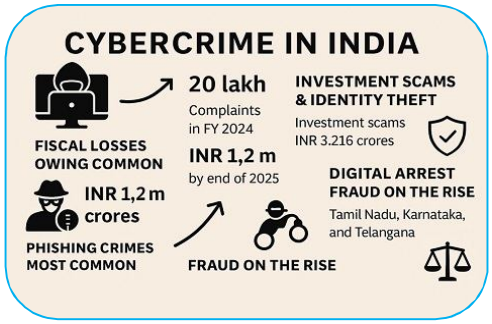
INDIAN STREAMS RESEARCH JOURNAL
ISSN NO : 2230-7850
IMPACT FACTOR : 5.1651 (UIF)
VOLUME - 15 | ISSUE - 10 | NOVEMBER – 2025

CYBERCRIME IN INDIA: LAW, POLICY, & ENFORCEMENT IN AI ERA

Prof. Narendra Kumar Dhaka[1] and Navaneet Kumar[2]
[1]Professor and Guide at Nirwan University, Jaipur, Discipline – Law.
[2]Research Scholar at Nirwan University, Jaipur, Discipline – Law.

ABSTRACT:



Advances in artificial intelligence (AI) particularly deepfakes and voice cloning, have catalyzed a new wave of cybercrime in India. The scale ranges from targeted impersonation and sextortion to mass financial fraud that exploits the Unified Payments Interface (UPI) rails and social-media platforms. This paper examines the Indian legal framework, principally the Information Technology Act, 2000, the Digital Personal Data Protection Act, 2023, criminal provisions in the Indian Penal Code (IPC)/Bharatiya Nyaya Sanhita (BNS), intermediary due diligence obligations, and the emerging jurisprudence of the Supreme Court and High Courts, alongside institutional responses such as CERT-In directions, the Indian Cyber Crime Coordination Centre (I4C), the National Cybercrime Reporting Portal (NCRP), and the 1930 helpline.

Drawing from recent case studies and official advisories, it identifies normative and operational gaps and proposes actionable reforms on provenance, platform governance, evidence, cross-border cooperation, and victim-centric remedies.

KEYWORDS: *Cybercrime, Artificial Intelligence (AI), Deepfakes and Voice Cloning, Digital Arrest Scams, Digital Personal Data Protection Act, 2023 (DPDP Act), Intermediary Liability and Due Diligence, CERT-In and I4C (Indian Cyber Crime Coordination Centre), National Cybercrime Reporting Portal (NCRP), International Cooperation (MLAT, CERT-to-CERT).*

I. INTRODUCTION

Indian cyberspace is experiencing a dual transformation: rapid digital adoption and a corresponding escalation in cybercrime sophistication. In Delhi, an elderly victim recently transferred money after hearing a cloned voice of a family member pleading for help.[1]

Across states, 'digital arrest' scams, courier-parcel frauds, and task-based investment rackets have proliferated on encrypted messaging apps and social media, while AI-manipulated images and videos of public figures flood timelines. Law enforcement capacity has grown e.g., dedicated cyber police stations, I4C's ecosystem, and NCRP workflows but detection and arrest rates in several jurisdictions still lag behind incident volumes.

This paper pursues three objectives:

i. First, it maps the legal architecture that applies to AI-enabled cyber harms, substantive offences, intermediary liability, privacy and data protection, and sectoral standards.
ii. It synthesizes recent practice, including state statistics and central blocking/advisory measures, to assess efficacy against impersonation, sextortion, phishing, account takeovers, data breaches, and cross-border laundering.
iii. It proposes a legislative and policy roadmap that is risk-proportionate, speech-respecting, and enforceable in India's federal, multilingual, and mobile-first context.

II. THE AI-ENABLED THREAT LANDSCAPE

a. Deepfakes and Voice Cloning

Deepfakes: synthetic audio-visual content generated or manipulated using AI are increasingly used for impersonation, sexualized abuse, and financial deception. High-profile Indian actors have reported deepfake clips that combined swapped visuals with dubbed audio to advertise fraudulent schemes; government leaders have publicly acknowledged the systemic risk.[2]

Voice cloning more directly targets trust. In the Delhi incident, a WhatsApp message combined with a fabricated audio plea led to an immediate transfer via digital payments, an illustration of 'human-in-the-loop' social engineering amplified by AI.[3]

b. Mass-Market Fraud over UPI and Messaging Apps

Task-based and investment frauds simulate legitimate workflows (liking videos, completing 'tasks') and then escalate payment demands. The Ministry of Home Affairs (MHA) and I4C have recommended and effected blocking of websites used for such rackets and emphasized reporting through NCRP.[4]

UPI remains a prime vector for instant value transfers. Police reports across cities describe fake bills, QR-code phishing, KYC and loan frauds, and fake customer-care numbers that trigger remote-access control or credential capture.

c. Data Breaches, Account Takeovers, and Platform Exploits

Credential stuffing against consumer platforms and targeted exploits against critical service providers continue to expose personal data at scale. In 2023, multiple Indian and global incidents highlighted the consequences of weak password hygiene, third-party vulnerabilities, and delayed patching regimes.[5]

d. 'Digital Arrest' and Coercive Extortion

Scammers increasingly stage law-enforcement backdrops on video calls, display forged IDs, and coerce continuous online presence what Indian reportage has termed 'digital arrest'. Victims are told to transfer funds 'for audit' or to avoid immediate arrest, often through UPI or mule accounts, causing substantial losses.[6]

III. EMPIRICAL GLIMPSES: CASES, TRENDS, AND CAPACITY

State-level data reveal rising complaints but uneven detection and arrest ratios, reflecting the scale of cross-jurisdictional money flows and the use of shell entities and mule accounts. Karnataka's statistics, for instance, show a sharp rise in reported cases between 2021 and 2023 with a concurrent decline in detection.[7]

At the same time, targeted enforcement, such as dismantling payment-gateway compromise rings and freezing proceeds via NCRP workflows has produced measurable recoveries.[8]

Central advisories and blocks under the Information Technology Act, 2000 (IT Act) have been used against task-based and investment-fraud infrastructure, while public campaigns by NPCI ('Gyaan Se, Dhyaan Se') seek to improve user hygiene.[9]

IV. THE INDIAN LEGAL ARCHITECTURE

a. Substantive Offences: IT Act and IPC/BNS

AI-enabled harms rarely create entirely novel offence categories; rather, they intensify classic wrongs like cheating, forgery, obscenity, criminal intimidation and extortion, defamation, and privacy intrusions. The IT Act's Sections 66C (identity theft), 66D (cheating by personation using computer resources), 66E (violation of privacy), and 67–67B (obscenity/sexually explicit content including child sexual abuse material) are frequently invoked, supplemented by IPC/ BNS provisions such as cheating, forgery, criminal defamation, and insulting the modesty of a woman.[10]

Prosecutions increasingly involve 'hybrid' fact patterns: a forged or morphed image (forgery/defamation), an AI-synthesised clip circulated without consent (IT Act s. 66E; 67A or POCSO where applicable), and a money trail through UPI into layered mule accounts (cheating and criminal conspiracy).

b. Intermediary Liability and Due Diligence

Intermediaries are granted conditional safe harbor under Section 79 of the IT Act, provided that they adhere to the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, as amended. These rules require reasonable efforts to prevent unlawful content, expedited takedowns of intimate imagery, appointment of grievance officers, and responsiveness to lawful orders. Government advisories have reiterated obligations concerning impersonation, morphed images, and deepfakes.[11]

The Grievance Appellate Committee (GAC) layer provides an administrative avenue for individual grievances against platform decisions. However, effective mitigation of deepfake virality requires provenance and detection tooling, hash-sharing for non-consensual imagery, and circuit-breakers for rapidly spreading content.

c. Privacy and Data Protection

The Digital Personal Data Protection Act, 2023 (DPDP Act) establishes core obligations on data fiduciaries and rights for data principals. Although not a cybercrime statute, it provides a compliance backbone data minimization, purpose limitation, security safeguards and breach notifications, that reduces breach surface area. Sectoral standards under RBI, CERT-In's 2022 Directions (logging, time synchronization, breach reporting), and MeitY advisories complement this layer.

d. Evidence Law and Forensics

AI-synthesised content challenges authenticity and chain-of-custody assessments. The Supreme Court's jurisprudence and Bharatiya Sakshya Adhiniyam, 2023 on electronic evidence requiring certification under section 63, while allowing limited curatives, must now be operationalized with forensic capabilities such us hashing, metadata capture, watermark detection at the police-station level. Platform logs, KYC trails, and payment timestamps are critical to establishing identity and intent.

e. Judicial Interventions

Courts have begun to address the balance between regulating deepfakes and preserving online freedoms. The Delhi High Court has cautioned that sweeping judicial directions to police the entire internet may be impractical, and that calibrated executive rulemaking is better suited for fast-evolving technologies.[12]

Separately, the Supreme Court's recognition of privacy as a fundamental right and its directions in matters concerning violent content takedowns inform a victim-centric approach to platform governance and proactive hashing of reported content.

V. COMPARATIVE AND INTERNATIONAL PERSPECTIVES

International responses range from watermarking mandates and provenance disclosures to platform liability regimes and election-specific prohibitions. The European Union's Digital Services Act (DSA) imposes risk assessments and crisis response duties on very large online platforms, while its AI Act framework emphasises transparency around deepfakes.[13]

The United States has seen state-level criminalisation of election-related and nonconsensual deepfakes, alongside federal proposals for labelling and provenance. China's deep synthesis rules require labelling and traceability. These models provide reference points for India's own watermarking and provenance proposals mooted under ongoing policy discussions and putative reforms.[14]

VI. GAPS AND CHALLENGES

a. Provenance and Authenticity

Without standardised provenance, platforms struggle to reliably detect manipulated media at scale, and users lack contextual cues. Watermarking and cryptographic signatures must be designed to be robust to compression and format-shifting while preserving privacy. A national hash database for non-consensual imagery linked to GAC and law-enforcement portals can reduce recirculation.

b. Cross-Border Money Flows and Mule Networks

Fraud proceeds often exit quickly through international fintech rails, card networks, or cash-outs. Strengthening bank and fintech-level transaction monitoring, tightening offboarding of mule accounts, and expanding real-time cooperation through NCRP/I4C can increase recoveries and deterrence.

c. Capacity and Tools for Police and Prosecutors

Frontline investigators need AI-enabled triage: on-device tools for media authenticity checks, rapid freezing through the 1930 escalation workflow, and standardised eDiscovery from platforms and payment intermediaries. State cyber labs require staffing for multimedia forensics and malware analysis; prosecutors need training on explaining AI-generated evidence and intent to courts.

d. Victim Support and Due Process

A victim-centric approach demands fast takedowns, confidential reporting, psychosocial support for sextortion survivors, and restitution pathways. At the same time, provenance and detection measures must avoid over-removal of lawful speech or creating surveillance risks. Transparent appeals via GAC and judicial review remain essential.

VII. Policy Recommendations for India

a. Mandate Provenance for High-Risk Generative Outputs: Adopt risk-tiered provenance requirements (watermarking/content credentials) for model providers and distribution platforms when outputs are photorealistic or voice-like. Require interoperable disclosure signals surfaced in UI and via APIs for newsrooms, civil society, and researchers.

b. Platform Circuit-Breakers and Hash-Sharing: Operationalise rapid circuit-breakers for virality spikes of suspected deepfakes involving minors, intimate imagery, or public-order risks. Establish a secure, audited hash-sharing consortium for non-consensual intimate imagery (NCII) with opt-in victim consent and cross-platform erasure.

c. Evidence Modernisation: Expand national standards for digital evidence acquisition (hashing, metadata capture, time-stamping) and integrate 65B-compliant workflows into police station SOPs. Encourage courts to recognise certified provenance signals and watermark detection reports as prima facie authenticity aids, subject to cross-examination.

d. Payments Safeguards and Account Hygiene: Require real-time risk scoring for first-time, high-value UPI transfers and display of enhanced counterparty identity signals. Incentivise PSPs to auto-throttle suspicious bursts and to publish quarterly transparency reports on freezes, recoveries, and response times through the 1930/NCRP channel.

e. Intermediary Due Diligence Enhancements: Clarify obligations for deepfake labelling, anti-impersonation UX, and responsive takedown SLAs for intimate imagery. Codify periodic model-card disclosures from large generative services that distribute within India.

f. **Capacity Building and Public Awareness:** Fund state-level forensic units for synthetic media, expand CCITR-style training hubs, and scale NPCI and MHA-led public campaigns in regional languages to counter 'digital arrest', QR-code and KYC scams.
g. **International Cooperation:** Use MLAT channels and CERT-to-CERT MoUs to target cross-border mule networks and ransomware affiliates. Support global standards on content credentials to reduce provenance fragmentation.

VIII. CONCLUSION

In conclusion, India's approach to regulating deepfakes and generative AI must be grounded in a principled balance between technological advancement, constitutional freedoms, and public safety. Drawing from global models such as the European Union's risk based Digital Services Act and AI Act, the United States' targeted criminalisation of harmful deepfakes, and China's traceability driven deep synthesis framework, India can formulate a hybrid system that aligns with its democratic, socio legal, and infrastructural realities. This necessitates a statutory foundation mandating provenance and watermarking for high risk generative outputs, coupled with adaptive intermediary liability norms, circuit breaker mechanisms for viral synthetic content, and harmonised standards for digital evidence admissibility under Section 63 of the Bhartiya Nayay Sanhita 2023. Parallel strengthening of financial surveillance and cooperative frameworks through NCRP and I4C is critical to curbing the monetisation of synthetic frauds and cross border mule networks. Further, India's policy design must centre victims, ensuring swift redressal, privacy respecting takedowns, psychosocial support, and transparency in content moderation decisions while avoiding overbroad censorship or chilling effects on innovation.
Ultimately, embedding AI provenance, due diligence, and accountability within a rights preserving digital governance architecture will enable India not merely to mitigate the harms of synthetic media, but to set normative leadership in shaping a secure and trustworthy global AI ecosystem which might be capable of fighting with the evil of AI and protect those who are not much qualified to take care of self from the AI evil.

REFERENCES

[1] NDTV News Desk, *AI Voice Cloning: What It Is and How to Avoid Getting Scammed By It*, NDTV (Feb. 12, 2024). Mukesh Singh Sengar, *Elderly Delhi Man Duped By AI-Cloned Voice, Loses Rs 50,000 in Extortion*, NDTV (Dec. 12, 2023).
[2] Sunil Bharti Mittal, *Sunil Mittal Exposes AI Scam, Says "My Voice Was Perfectly Articulated ... Cloning Attempt"*, *Econ. Times* (Oct. 21, 2024).
[3] NDTV News Desk, *supra note – i,* Dheeraj Mishra, *Voice Cloning Scam: How Cyber-Criminals Are Using Kids' Voices to Dupe Parents*, *Indian Express* (Feb. 12, 2024).
[4] Ministry of Home Affairs, Government of India, *Press Release: I4C Identifies 100+ Websites Involved in Part-Time/Investment Fraud; MeitY Blocks Them; Complaints via 1930 Helpline & NCRP*, Press Information Bureau (Dec. 6, 2023). Indian Cybercrime Coordination Centre (I4C), Ministry of Home Affairs, *National Cybercrime Reporting Portal (NCRP) – Features & Helpline 1930*, https://www.i4c.mha.gov.in/ncrp.aspx (last updated May 13, 2025).
[5] Arctic Wolf, *2023 Data Breaches in Review*, ARCTIC WOLF (Jan. 25, 2024). F5 Labs, *2023 Identity Threat Report: Executive Summary*, F5 LABS (2023). AhnLab Security Emergency-Response Center (ASEC), *Cases & Countermeasures of Credential Stuffing Attacks Using Leaked Accounts*, ASEC (2023).
[6] NDTV News Desk, *73-Year-Old Noida Lawyer Loses Over Rs 3 Crore in "Digital Arrest" Scam*, NDTV (Aug. 27, 2024). Zoya Mateen, *The "Digital Arrest" Scam Terrifying India's Middle Class*, BBC News (Apr. 5, 2024). *Woman Duped of Over Rs 11 Lakh*, *Times of India* (Nov. 29, 2023). The Logical Indian Team, *Digital Arrest Scam: How a Noida Family Was Held Hostage for 5 Days and Lost ₹1.1 Crore*, *Logical Indian* (Apr. 3, 2024). Times News Network, *Police Arrest CA for Siphoning ₹25 Crore from Payment Gateway Firm*, *Times of India* (Nov. 29, 2023). NDTV News Desk, *3 More Arrested in Connection with ₹25 Crore Cyber-Fraud Case in Maharashtra*, NDTV (Oct. 12, 2023). ET Bureau, *Thane Payment Gateway Fraud:*

₹16,180 Crore Siphoned Off After Hacking, *Econ. Times* (Oct. 6, 2023). Hindustan Times Correspondent, *Accused Who Hacked Into Payment Gateway and Stole ₹25 Crore Booked*, *Hindustan Times* (Oct. 8, 2023).

[7] *Not Sparing Even Pettiest of Crimes: Police on Rise in Bengaluru's Cases*, *Deccan Herald* (Jan. 3, 2024). Johnson T. A., *What the NCRB 2023 Report Reveals About Growing Cybercrimes in Karnataka*, *Indian Express* (Oct. 4, 2025).

[8] Nishikant Karlikar, *Police Arrest CA, One Other for Siphoning 25 Crore from Payment Gateway Firm*, *Times of India* (Dec. 3, 2023). *3 More Arrested in Connection with ₹25 Crore Cyber-Fraud Case in Maharashtra*, NDTV (Oct. 27, 2023). PTI, *Thane Payment Gateway Fraud: ₹16,180 Crore Siphoned Off After Hacking*, *Econ. Times* (Oct. 8, 2023). Anamika Gharat, *Accused Who Hacked Into Payment Gateway and Stole ₹25 Crore Booked*, *Hindustan Times* (Oct. 8, 2023).

[9] Vijay Kumar Yadav, *Mumbai Loses Rs 1,200 Crore to Cyber Crime This Year, a Rise of 350%*, *Indian Express* (Dec. 23, 2024). Mohamed Thaver, *Woman Loses Rs 7.8 Crore in Cyber Fraud, Mumbai Police Stops Ongoing "Digital Arrest"*, *Indian Express* (Aug. 14, 2025). *UPI Scams and Tips to Stay Safe — "Gyaan Se, Dhyaan Se" Campaign*, *Times of India* (Dec. 7, 2023). *NPCI Launches UPI Safety Awareness Campaign with Pankaj Tripathi*, *Outlook Money* (Nov. 6, 2024).

[10] Information Technology Act, No. 21 of 2000, section 66C–66E, 67–67B (India). Indian Penal Code, No. 45 of 1860, section - 420, 465, 468, 469, 499, 509 (India). Protection of Children from Sexual Offences Act, No. 32 of 2012 (India).

[11] Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, Gazette of India, Extraordinary, Part II, Sec. 3(i), G.S.R. 139(E) (India). Information Technology Act, No. 21 of 2000, section 79 (India). Ministry of Electronics and Information Technology (MeitY), Government of India, *Advisory to Social Media Intermediaries on Deepfakes and Impersonation*, Press Information Bureau (Jan. 3, 2024). Ministry of Electronics and Information Technology (MeitY), Government of India, *Advisory to All Intermediaries and Platforms Regarding Identification and Removal of Deepfakes*, Press Information Bureau (May 2, 2025). Internet Freedom Foundation, *Dealing with Deepfakes: India's Evolving Legal Framework*, INTERNET FREEDOM (last visited Oct. 18, 2025).

[12] *Delhi High Court Directs the Government to Nominate Members to Committees Relating to Deepfake Issue* (Nov. 28, 2024). *Delhi HC Grants Panel 3 Months for Comprehensive Report*, *Econ. Times* (Mar. 24, 2025). *Deepfakes and AI-Generated Content: A Landmark Ruling by the Delhi High Court in Response to a Harassment Campaign Targeting a Public Activist*, DDG (Jul. 21, 2025).

[13] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and Amending Directive 2000/31/EC (Digital Services Act), 2022 O.J. (L 277) 1. M. Łabuz, *Regulating Deep-Fakes in the Artificial Intelligence Act*, *ACI GJournal* (2023). *The Digital Services Act (DSA)*, EU Digital Services Act (last visited Oct. 19, 2025).

[14] National Conference of State Legislatures, *Deceptive Audio or Visual Media ("Deepfakes") 2024 Legislation*, NCSL (Nov. 22, 2024). *Each and Every Person from False Appearances by Keeping Exploitation Subject to Accountability Act of 2023* (DEEPFAKES Accountability Act), H.R. 5586, 118th Cong. (2023–24).