INTERNATIONAL JOURNAL FOR LEGAL RESEARCH AND ANALYSIS

Open Access, Refereed Journal Multi Disciplinary
Peer Reviewed

www.ijlra.com

ISSN: 2582-6433

DISCLAIMER

No part of this publication may be reproduced, stored, transmitted, or distributed in any form or by any means, whether electronic, mechanical, photocopying, recording, or otherwise, without prior written permission of the Managing Editor of the *International Journal for Legal Research & Analysis (IJLRA)*.

The views, opinions, interpretations, and conclusions expressed in the articles published in this journal are solely those of the respective authors. They do not necessarily reflect the views of the Editorial Board, Editors, Reviewers, Advisors, or the Publisher of IJLRA.

Although every reasonable effort has been made to ensure the accuracy, authenticity, and proper citation of the content published in this journal, neither the Editorial Board nor IJLRA shall be held liable or responsible, in any manner whatsoever, for any loss, damage, or consequence arising from the use, reliance upon, or interpretation of the information contained in this publication.

The content published herein is intended solely for academic and informational purposes and shall not be construed as legal advice or professional opinion.

ISSN: 2582-6433

ABOUT US

The *International Journal for Legal Research & Analysis (IJLRA)* (ISSN: 2582-6433) is a peer-reviewed, academic, online journal published on a monthly basis. The journal aims to provide a comprehensive and interactive platform for the publication of original and high-quality legal research.

IJLRA publishes Short Articles, Long Articles, Research Papers, Case Comments, Book Reviews, Essays, and interdisciplinary studies in the field of law and allied disciplines. The journal seeks to promote critical analysis and informed discourse on contemporary legal, social, and policy issues.

The primary objective of IJLRA is to enhance academic engagement and scholarly dialogue among law students, researchers, academicians, legal professionals, and members of the Bar and Bench. The journal endeavours to establish itself as a credible and widely cited academic publication through the publication of original, well-researched, and analytically sound contributions.

IJLRA welcomes submissions from all branches of law, provided the work is original, unpublished, and submitted in accordance with the prescribed submission guidelines. All manuscripts are subject to a rigorous peer-review process to ensure academic quality, originality, and relevance.

Through its publications, the *International Journal for Legal Research & Analysis* aspires to contribute meaningfully to legal scholarship and the development of law as an instrument of justice and social progress.

ISSN: 2582-6433

PUBLICATION ETHICS, COPYRIGHT & AUTHOR RESPONSIBILITY STATEMENT

The *International Journal for Legal Research and Analysis (IJLRA)* is committed to upholding the highest standards of publication ethics and academic integrity. All manuscripts submitted to the journal must be original, unpublished, and free from plagiarism, data fabrication, falsification, or any form of unethical research or publication practice. Authors are solely responsible for the accuracy, originality, legality, and ethical compliance of their work and must ensure that all sources are properly cited and that necessary permissions for any third-party copyrighted material have been duly obtained prior to submission. Copyright in all published articles vests with IJLRA, unless otherwise expressly stated, and authors grant the journal the irrevocable right to publish, reproduce, distribute, and archive their work in print and electronic formats. The views and opinions expressed in the articles are those of the authors alone and do not reflect the views of the Editors, Editorial Board, Reviewers, or Publisher. IJLRA shall not be liable for any loss, damage, claim, or legal consequence arising from the use, reliance upon, or interpretation of the content published. By submitting a manuscript, the author(s) agree to fully indemnify and hold harmless the journal, its Editor-in-Chief, Editors, Editorial Board, Reviewers, Advisors, Publisher, and Management against any claims, liabilities, or legal proceedings arising out of plagiarism, copyright infringement, defamation, breach of confidentiality, or violation of third-party rights. The journal reserves the absolute right to reject, withdraw, retract, or remove any manuscript or published article in case of ethical or legal violations, without incurring any liability.

ISSN: 2582-6433

WHEN ALGORITHMS OFFEND: EMERGING AI-DRIVEN CYBER THREATS, LEGAL CHALLENGES AND THE REGULATORY RESPONSE

AUTHORED BY - VANDANA CHAUDHARY & SHITI KANTH DUBEY
Faculty of Law, Agra College
Dr. Bhimrao Ambedkar University, Agra, India

Abstract

The rapid deployment of Artificial Intelligence has fundamentally changed the ecology of cybercrime, transforming its nature, scale, sophistication and geographical outreach. Conduct once requiring skilled human offenders is increasingly automated: generative models produce persuasive phishing at industrial scale, synthetic audio and video enable real-time impersonation of executives and public figures, and machine-learning-assisted malware adapts to defensive controls in flight. The National Crime Records Bureau records an alarming 31.2 per cent rise in cybercrime in India in 2023, reaching 86,420 registered cases, while the United States Internet Crime Complaint Center logged losses exceeding USD 16.6 billion in 2024 and USD 20.9 billion in 2025, the first year in which "AI-related" was recognised as a distinct crime description. Against this backdrop, this paper examines whether India's legal framework, comprising the Information Technology Act, 2000, the Bharatiya Nyaya Sanhita, 2023, the Bharatiya Sakshya Adhiniyam, 2023, the Digital Personal Data Protection Act, 2023, and the associated intermediary and CERT-In regime, is doctrinally and institutionally equipped to address AI-enabled criminal conduct. Adopting a doctrinal and comparative legal methodology, the paper analyses the constitutional guarantees under Articles 14, 19 and 21 as interpreted in Justice K.S. Puttaswamy v. Union of India, evaluates the evidentiary framework built around Section 63 of the Bharatiya Sakshya Adhiniyam, and compares Indian responses with the European Union Artificial Intelligence Act, the Budapest Convention on Cybercrime, the newly adopted United Nations Convention against Cybercrime (Hanoi Convention, 2024), the United States Computer Fraud and Abuse Act, and the United Kingdom Computer Misuse Act. It argues that although several AI-enabled offences are covered by technology-neutral provisions, distinctive problems of attribution, mens rea, synthetic evidence and cross-border enforcement disclose serious doctrinal and institutional lacunae. The paper concludes with a

ISSN: 2582-6433

calibrated reform agenda spanning legislative, procedural, forensic, judicial-training and international-cooperation dimensions.

Keywords: Artificial Intelligence; AI-enabled Cybercrime; Deepfakes; Digital Evidence; Cyber Law; AI Governance.

I. Introduction

Artificial Intelligence has moved decisively from research laboratories into the fabric of contemporary economic, social and criminal life. Systems capable of generating text, images, audio and video that are effectively indistinguishable from human production are now widely accessible, cheaply operable, and improving at a pace no prior information technology has matched. This democratisation of computational capability is simultaneously a driver of productivity and a multiplier of criminal opportunity. The National Crime Records Bureau (NCRB) records that cybercrime in India rose by 31.2 per cent in 2023, reaching 86,420 registered cases, of which nearly 68.9 per cent were motivated by fraud.[1] The Federal Bureau of Investigation's Internet Crime Complaint Center (IC3) reported losses of USD 16.6 billion in 2024, a 33 per cent year-on-year rise, and USD 20.9 billion in 2025, the first year in which "AI-related" was formally designated as a distinct crime descriptor.[2] India's own Indian Cyber Crime Coordination Centre ("I4C") recorded financial-fraud complaints climbing from 2.62 lakh in 2021 to 24.02 lakh by mid-2026.[3] The scale, speed and personalisation of these harms are increasingly attributable to the deployment of machine learning by hostile actors.

The central question that animates this paper is whether India's existing legal framework - an assemblage of the Information Technology Act, 2000 (IT Act),[4] the Bharatiya Nyaya Sanhita, 2023 (BNS),[5] the Bharatiya Nagarik Suraksha Sanhita, 2023 (BNSS), the Bharatiya Sakshya Adhiniyam, 2023 (BSA),[6] the Digital Personal Data Protection Act, 2023 (DPDP Act),[7] and

[1]National Crime Records Bureau. (2025). Crime in India 2023. Ministry of Home Affairs, Government of India. Retrieved from https://ncrb.gov.in
[2]Federal Bureau of Investigation. (2025). 2024 Internet Crime Report. Internet Crime Complaint Center (IC3). https://www.ic3.gov/AnnualReport/Reports/2024_IC3Report.pdf
[3]Ministry of Home Affairs, Government of India. (2026). Indian Cyber Crime Coordination Centre (I4C): Performance and Interventions. Press Information Bureau. Retrieved from https://pib.gov.in
[4]The Information Technology Act, 2000, No. 21 of 2000, Acts of Parliament, 2000 (India). India Code. https://www.indiacode.nic.in
[5]The Bharatiya Nyaya Sanhita, 2023, No. 45 of 2023, Acts of Parliament, 2023 (India). India Code.
[6]The Bharatiya Sakshya Adhiniyam, 2023, No. 47 of 2023, Acts of Parliament, 2023 (India). India Code.
[7]The Digital Personal Data Protection Act, 2023, No. 22 of 2023, Acts of Parliament, 2023 (India). Digital Personal Data Protection Rules, 2025, notified vide G.S.R. 843(E), Ministry of Electronics and Information

ISSN: 2582-6433

the derivative rules and directions issued under them, is doctrinally and institutionally capable of responding to AI-enabled criminal conduct. That inquiry cannot be answered in the abstract. It requires a careful appraisal of substantive offences, evidentiary standards, investigative capacity, judicial jurisprudence, constitutional guarantees, comparative frameworks and international cooperation. The paper argues that although many AI-enabled harms are covered by technology-neutral provisions, distinctive problems of attribution, mens rea, synthetic evidence and cross-border enforcement expose serious lacunae that require calibrated reform rather than either legislative complacency or hasty over-criminalisation.

II. Research Methodology

The paper adopts a doctrinal and comparative legal methodology supplemented by policy analysis. Primary sources consulted include the Constitution of India, Indian statutes, subordinate rules and government notifications, judicial pronouncements of the Supreme Court and High Courts, and international instruments (the Budapest Convention on Cybercrime, the United Nations Convention against Cybercrime, and the European Union Artificial Intelligence Act). Secondary sources comprise reports of the National Crime Records Bureau, the Ministry of Home Affairs, the Ministry of Electronics and Information Technology, the Indian Computer Emergency Response Team, the Reserve Bank of India, the Federal Bureau of Investigation, Europol, ENISA, the United Nations Office on Drugs and Crime, and the Organisation for Economic Co-operation and Development, along with peer-reviewed academic literature. Empirical claims are drawn only from cited institutional data, and expert projections are identified as such.

III. Historical Evolution of Cybercrime and the Emergence of AI-Enabled Threats

The evolution of cybercrime can be traced across five broad phases. The first phase, from the late 1980s to the early 1990s, comprised individually authored viruses and unauthorised access performed largely by hobbyists. The second phase saw the professionalisation of hacking and the emergence of financially motivated fraud in the wake of e-commerce. The third phase, from the mid-2000s, introduced organised cybercriminal enterprises, characterised by crime-as-a-service, botnets rented by the hour, and ransomware syndicates. The fourth phase, roughly from 2015, was defined by automated attacks, large-scale credential stuffing, industrialised phishing

Technology, 13 November 2025.

ISSN: 2582-6433

kits, and cryptocurrency-enabled anonymous payments. The current fifth phase is characterised by AI-enabled cybercrime: generative models produce phishing communications at industrial scale, voice cloning enables real-time impersonation, and adversarial machine learning permits the systematic evasion of security controls. Each transition has altered offender profile, victim reach, attack speed, anonymity, and evidentiary difficulty; none has been more consequential than the current shift, in which even individuals without technical expertise can access tools once confined to state actors or specialised criminal enterprises.

IV. Conceptualising AI-Enabled Cybercrime: A Doctrinal Taxonomy

Doctrinal precision demands that four distinct categories be kept separate. First, traditional cybercrime encompasses conduct such as unauthorised access, phishing or ransomware executed without AI. Second, AI-assisted cybercrime denotes conventional offences that are enhanced by AI, for example, phishing emails drafted with the aid of generative models. Third, AI-enabled cybercrime describes offences whose commission substantially depends on AI, such as real-time deepfake impersonation of a chief executive to authorise a fraudulent wire transfer. Fourth, autonomous AI-generated cyber threats denote conduct in which AI systems iterate, adapt or execute malicious activity with reduced human intervention, including polymorphic malware that alters its own code to evade detection. This taxonomy matters because criminal responsibility, mens rea, attribution, sentencing and regulatory allocation vary sharply across the four categories. Conflating them, the common failing of both regulators and commentators, produces either over-criminalisation or under-regulation.

V. Emerging Forms of AI-Enabled Cybercrime

Six broad categories of AI-enabled cyber-harm have crystallised. First, AI-generated phishing and business email compromise (BEC) deploy grammatically flawless, culturally tailored and highly personalised messages. IC3 recorded USD 2.77 billion in BEC losses in 2024 across 21,442 complaints, and voice cloning has become a standard follow-up tactic to make fraudulent instructions appear authentic.[8] Second, deepfake and synthetic identity fraud has scaled dramatically. Analyses estimate USD 2.19 billion in global deepfake-driven fraud by 2025, with USD 1.65 billion in 2025 alone, over half attributable to impersonation of officials and celebrities.[9] The engineering firm Arup lost USD 25.6 million in a single deepfake-video-

[8]2024 Internet Crime Report, supra note 2.
[9]Surfshark Research. (2025). AI-driven Deepfake Fraud Analysis 2019–2025; Vectra AI. (2026). The State of AI-Enabled Fraud, March 2026; World Economic Forum. (2026). Global Cybersecurity Outlook 2026, WEF, Geneva.

ISSN: 2582-6433

call fraud in early 2024, underscoring the operational maturity of the threat. Third, AI-assisted malware and automated vulnerability discovery increases both the volume and the adaptability of attacks; ransomware operations increasingly incorporate AI-driven target selection.

Fourth, credential and identity attacks are supercharged by generative AI: identity-fraud attempts using deepfakes reportedly surged by approximately 3,000 per cent in 2023 alone, and businesses lost roughly USD 500,000 per major deepfake-related incident on average in 2024.[10] Fifth, AI-generated misinformation and disinformation intersects directly with election integrity, public order and reputational harm; the November 2023 deepfake targeting a leading Indian actor illustrates the reputational, dignitary and evidentiary dimensions of the threat. Sixth, novel offences against AI systems themselves, model theft, data poisoning, adversarial evasion, and prompt injection, raise doctrinal questions for which criminal law is currently unprepared.

Digital arrest scams, a form of AI-enabled extortion in which criminals impersonate law-enforcement officers over videoconference and stage "virtual custody" - rose from 39,925 cases in 2022 to 1,23,672 in 2024, with reported losses climbing from ₹91 crore to ₹1,935 crore.[11] Approximately 46 per cent of such operations originate from criminal compounds in Southeast Asia, particularly Cambodia, Myanmar and Laos, where the United Nations Office on Drugs and Crime has documented forced labour by trafficked persons compelled to run online scams.[12] These numbers reveal not merely a domestic law-enforcement problem but a transnational organised-crime challenge with distinctive AI-driven attack vectors.

A related development is the maturation of "cybercrime-as-a-service," in which specialised criminal enterprises rent phishing kits, credential-stuffing infrastructure, AI-generated deepfake production, and ransomware toolchains to lower-skilled operators on subscription terms. The commodification of AI-enabled offensive capability has substantially reduced the technical threshold for entry into cybercrime, expanded the offender population, and blurred the traditional distinction between organised and individual criminality. In the financial-

[10]Surfshark & Vectra AI Deepfake Analysis, supra note 9.
[11]Ministry of Home Affairs, Government of India. (2026). Data on Digital Arrest Complaints and Financial Losses, Indian Cyber Crime Coordination Centre (I4C), Reply to Rajya Sabha Unstarred Question.
[12]United Nations Office on Drugs and Crime. (2024). Transnational Organized Crime and the Convergence of Cyber-enabled Fraud, Human Trafficking for Forced Criminality and Money Laundering in South-East Asia. UNODC Regional Office for Southeast Asia and the Pacific.

ISSN: 2582-6433

services domain, the Reserve Bank of India has repeatedly cautioned regulated entities about impersonation, mule accounts, and synthetic-identity fraud, underscoring the sectoral pressure that AI-enabled crime has begun to exert on the payments system.

VI. Statistical Dimension and Empirical Landscape

Any responsible analysis must distinguish general cybercrime statistics from AI-specific statistics. In India, the NCRB's Crime in India 2023 reports 86,420 registered cybercrime cases (a 31.2 per cent rise from 65,893 in 2022), with fraud comprising 68.9 per cent (59,526 cases), sexual exploitation 4.9 per cent, and extortion 3.8 per cent.[13] IT Act offences rose by 36 per cent, driven principally by cheating by personation. The National Cyber Crime Reporting Portal received approximately 7.4 lakh complaints in the first four months of 2024; between 2021 and mid-2025 financial-fraud complaints rose from 2.62 lakh to 24.02 lakh.[14] By mid-2026, the Citizen Financial Cyber Fraud Reporting and Management System had helped save more than ₹11,158 crore across 32.80 lakh complaints; over 15.75 lakh SIMs and 5.77 lakh IMEIs had been blocked; and fraudulent transactions worth more than ₹25,698 crore had been declined through the Suspect Registry.

Internationally, IC3 recorded 859,532 complaints and USD 16.6 billion in reported losses in 2024, up 33 per cent from 2023, with phishing and spoofing (193,407 complaints), extortion and personal-data breaches leading the categories.[15] The 2025 figure of USD 20.9 billion marks the first year in which IC3 formally recognised "AI-related" complaints, logging almost 22,000 such matters and approximately USD 900 million in AI-attributed losses, widely regarded by analysts as a floor rather than a ceiling given systemic under-reporting. Projections such as the Deloitte Center for Financial Services' estimate that generative-AI-enabled fraud in the United States alone may reach USD 40 billion by 2027 must be presented as projections and not confirmed statistics.[16] The European Union Agency for Cybersecurity has similarly identified AI-driven threats as the most rapidly escalating category in its 2024 Threat Landscape assessment.[17]

[13]Crime in India 2023, supra note 1.
[14]I4C Performance and Interventions, supra note 3.
[15]2024 Internet Crime Report, supra note 2.
[16]Deloitte Center for Financial Services. (2024). Generative AI is Expected to Magnify the Risk of Deepfakes and Other Fraud in Banking. Deloitte Insights.
[17]European Union Agency for Cybersecurity. (2024). ENISA Threat Landscape 2024. Publications Office of the European Union.

ISSN: 2582-6433

VII. Indian Constitutional Framework

The constitutional foundations for regulating AI-enabled cybercrime rest on Articles 14, 19 and 21 of the Constitution of India.[18] Article 14 speaks to non-arbitrariness and, in the AI context, requires that algorithmic decision-making by the State satisfy the standards of reasoned action and non-discrimination. Article 19(1)(a) protects freedom of speech and expression, extended, since Shreya Singhal v. Union of India, to online communication, while Article 19(2) permits reasonable restrictions on the grounds enumerated therein.[19] Any regulatory response to deepfakes or synthetic content must therefore satisfy the twin requirements of narrow tailoring and proportionality. Article 21 protects life and personal liberty and, following the nine-judge Bench decision in Justice K.S. Puttaswamy v. Union of India, encompasses informational privacy, dignity and identity as fundamental rights.[20] The Puttaswamy jurisprudence is doctrinally significant for the AI-cybercrime interface because deepfake impersonation, non-consensual synthetic imagery, and unauthorised training on personal data all engage privacy interests that the State is now constitutionally bound to protect. Restrictions must be lawful, necessary in a democratic society, proportionate, and procedurally safeguarded.

VIII. Indian Statutory and Regulatory Framework

The IT Act, 2000 remains the principal cyber-specific legislation. Section 43 imposes civil liability for unauthorised access and damage; Section 66 criminalises those acts when done dishonestly or fraudulently; Section 66C punishes identity theft with imprisonment up to three years and a fine of up to ₹1 lakh; Section 66D criminalises cheating by personation using computer resources with the same maximum term; Section 66E punishes violation of bodily privacy through electronic capture, publication or transmission of intimate images; Section 66F addresses cyber terrorism, punishable with imprisonment up to life; Sections 67, 67A and 67B govern obscene, sexually explicit and child-sexual-abuse material; Section 69A empowers blocking of unlawful content; and Section 70B statutorily anchors the Indian Computer Emergency Response Team.[21] These provisions, though technology-neutral, were drafted with human offenders in mind and do not squarely address AI-generated content, autonomous adaptation, or model-level attacks.

[18]The Constitution of India, arts. 14, 19, 21.
[19]Shreya Singhal v. Union of India, (2015) 5 SCC 1 (Supreme Court of India, 24 March 2015).
[20]Justice K.S. Puttaswamy (Retd.) v. Union of India, (2017) 10 SCC 1, Writ Petition (Civil) No. 494 of 2012 (Supreme Court of India, 24 August 2017).
[21]Information Technology Act, 2000, supra note 4.

ISSN: 2582-6433

The BNS, 2023, which came into force on 1 July 2024, restates and consolidates a range of applicable offences: Section 111 addresses organised crime; Section 318 defines cheating (replacing Sections 415 and 420 of the erstwhile Indian Penal Code); Section 319 criminalises cheating by personation; Section 335 defines making of a false document or electronic record; Section 336 punishes forgery, including forgery of electronic records; Sections 337 and 338 address forgery of valuable security; Section 351 covers criminal intimidation; Section 353 addresses publication of statements conducing to public mischief; and Section 356 codifies defamation.[22] AI-generated defamatory content, synthetic threats, and deepfake-driven cheating fall squarely within these provisions, but their application to AI-generated conduct raises questions of authorship, agency and intent that the drafters did not directly address.

The BSA, 2023 replaced the Indian Evidence Act, 1872 with effect from 1 July 2024. Section 61 preserves the admissibility of electronic and digital records; Section 62 sets out the mode of proof; and Section 63 - the direct successor to the erstwhile Section 65B of the Evidence Act - governs the admissibility of computer output. Section 63(4) mandates a certificate identifying the electronic record, describing the manner of its production, and, importantly, signed both by a person in charge of the device and by an "expert," and further requires inclusion of the hash value in the Schedule.[23] This dual-certification and hash-value regime is a doctrinally important step forward for authentication of digital evidence, though it also intensifies the practical burden on investigators.

The DPDP Act, 2023 (Act No. 22 of 2023) established India's first standalone personal-data-protection statute. It received Presidential assent on 11 August 2023 but remained substantially inoperative until the Digital Personal Data Protection Rules, 2025 were notified on 13 November 2025 through G.S.R. 843(E), which also triggered a phased commencement of the Act.[24] The Act applies to processing within India and, extraterritorially, to entities offering goods or services to individuals in India. It operates primarily on consent, contains obligations concerning data-breach notification, storage limitation and cross-border transfer, and creates the Data Protection Board of India.

Two subordinate regimes are of particular importance. First, the CERT-In Directions of 28

[22]Bharatiya Nyaya Sanhita, 2023, supra note 5.
[23]Bharatiya Sakshya Adhiniyam, 2023, supra note 6.
[24]Digital Personal Data Protection Act, 2023, supra note 7.

ISSN: 2582-6433

April 2022, issued under Section 70B(6) of the IT Act, mandate reporting of specified cyber-incidents within six hours of noticing and require retention of information-and-communication-technology logs for 180 days within Indian jurisdiction.[25] Second, the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 impose due-diligence obligations on intermediaries, with enhanced obligations on "significant social media intermediaries." Draft amendments notified on 22 October 2025 introduce a formal definition of "synthetically generated information" under Rule 2(1)(wa) and require labelling, metadata traceability, and reasonable technical measures to detect and moderate deepfakes.[26]

IX. Judicial Approach in India

Indian courts have developed a modest but doctrinally consequential jurisprudence on cyber and privacy matters. Justice K.S. Puttaswamy v. Union of India, a nine-judge decision, recognised the right to privacy as a fundamental right protected under Articles 14, 19 and 21 and supplied the constitutional touchstone for evaluating both surveillance and data-protection regimes.[27] Shreya Singhal v. Union of India struck down Section 66A of the IT Act as unconstitutionally vague and read down Section 79 to require a court order or authorised government direction before intermediaries lose their safe-harbour protection.[28] These decisions collectively frame the balance between free expression and reasonable restriction that must inform any regulatory response to AI-generated speech.

On electronic evidence, three decisions are foundational. Anvar P.V. v. P.K. Basheer held that *"Section 65B of the Indian Evidence Act was a complete code and that a certificate under Section 65B(4) was mandatory for admission of secondary electronic evidence."*[29] Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal reaffirmed Anvar, clarified that *"where the original device is produced no certificate is required, and provided procedural guidance for cases where the certificate is unavailable through no fault of the party seeking to adduce the*

[25]Ministry of Electronics and Information Technology, Government of India. (2022). Directions under sub-section (6) of Section 70B of the Information Technology Act, 2000, relating to information security practices, procedure, prevention, response and reporting of cyber incidents for Safe & Trusted Internet, Notification No. 20(3)/2022-CERT-In, 28 April 2022.
[26]Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, G.S.R. 139(E), Ministry of Electronics and Information Technology, 25 February 2021; Draft Amendments notified for public consultation on 22 October 2025, introducing the definition of 'synthetically generated information' under Rule 2(1)(wa).
[27]Puttaswamy, supra note 20.
[28]Shreya Singhal, supra note 19.
[29]Anvar P.V. v. P.K. Basheer, (2014) 10 SCC 473 (Supreme Court of India, 18 September 2014).

ISSN: 2582-6433

record."[30] Section 63 of the BSA now carries forward this jurisprudence with the added requirement of expert certification and hash values. State of Tamil Nadu v. Suhas Katti remains historically important as the first Indian conviction under the IT Act, illustrating early judicial receptivity to digital evidence in an online-defamation-through-impersonation context.[31] The Delhi High Court in Chaitanya Rohilla v. Union of India constituted a committee to examine deepfake regulation, indicating growing judicial engagement with generative-AI harms.[32]

X. International Legal Framework and Comparative Perspectives

At the international level, the Council of Europe's Convention on Cybercrime, 2001 ("Budapest Convention") remains the principal harmonising instrument on substantive offences, procedural powers, and international cooperation, and its Second Additional Protocol of 2022 addresses enhanced cross-border access to electronic evidence including direct cooperation with service providers.[33] India is not a State Party. The United Nations General Assembly adopted the United Nations Convention against Cybercrime by Resolution 79/243 of 24 December 2024; the Convention was opened for signature on 25 October 2025 at Hanoi - hence the label "Hanoi Convention" - and will enter into force ninety days after deposit of the fortieth instrument of ratification.[34] It criminalises specified ICT offences, provides for cross-border cooperation and electronic-evidence sharing, and creates a twenty-four-hour cooperation network.

In the European Union, Regulation (EU) 2024/1689 - the Artificial Intelligence Act, entered into force on 1 August 2024 and adopts a tiered risk-based approach, prohibiting certain AI practices and imposing extensive obligations on providers and deployers of high-risk systems.[35] It interlocks with the General Data Protection Regulation, the NIS 2 Directive on

[30]Arjun Panditrao Khotkar v. Kailash Kushanrao Gorantyal, (2020) 7 SCC 1 (Supreme Court of India, 14 July 2020).
[31]State of Tamil Nadu v. Suhas Katti, C.C. No. 4680 of 2004 (Chief Metropolitan Magistrate, Egmore, Chennai, 5 November 2004) — first conviction in India under the Information Technology Act, 2000.
[32]Chaitanya Rohilla v. Union of India & Ors., W.P. (C) No. 15596 of 2023 (Delhi High Court, pending) — petition seeking regulation of deepfakes.
[33]Council of Europe. (2001). Convention on Cybercrime, ETS No. 185, opened for signature 23 November 2001, entered into force 1 July 2004; Council of Europe. (2022). Second Additional Protocol to the Convention on Cybercrime on Enhanced Co-operation and Disclosure of Electronic Evidence, CETS No. 224, opened for signature 12 May 2022.
[34]United Nations. (2024). United Nations Convention against Cybercrime: Strengthening International Cooperation for Combating Certain Crimes Committed by Means of Information and Communications Technology Systems and for the Sharing of Evidence in Electronic Form of Serious Crimes, G.A. Res. 79/243 (24 December 2024); opened for signature at Hanoi, 25 October 2025.
[35]Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), OJ L, 12 July 2024. Entered into force on

ISSN: 2582-6433

cybersecurity across essential services, and the Digital Services Act.[36] The United States relies principally on the Computer Fraud and Abuse Act, 18 U.S.C. § 1030, together with sectoral statutes, agency guidance, and state-level privacy law, while executive-branch initiatives address AI safety. The United Kingdom Computer Misuse Act, 1990 criminalises unauthorised access, unauthorised acts with intent to impair, and the making, supplying or obtaining of hacking tools; the Online Safety Act, 2023 imposes duties on services to address illegal and harmful content.[37] Comparative analysis reveals that no jurisdiction has produced a fully coherent AI-cybercrime framework; each addresses fragments of the challenge, and India should therefore learn selectively rather than imitatively.

XI. Law Enforcement and Practical Challenges

The distinctive challenges AI poses to enforcement fall into six clusters. First, attribution: identifying the natural person or entity behind AI-generated conduct is exceptionally difficult when generative outputs efface stylistic fingerprints and criminal infrastructure is distributed across jurisdictions. Second, mens rea: proving criminal intention where AI systems interpose autonomous decisions between the operator and the harm resists conventional doctrinal categorisation. Third, evidentiary uncertainty: distinguishing genuine from synthetic content requires forensic capacities that most Indian district-level investigative units do not possess; the certificate-and-hash regime under Section 63 of the BSA, while doctrinally sound, presumes technical expertise that is unevenly distributed. Fourth, cross-border investigation: cybercriminal operations are routinely spread across five jurisdictions, offender, victim, cloud infrastructure, financial rail and beneficiary, and India's non-participation in the Budapest Convention constrains the efficacy of mutual legal assistance. Fifth, capacity deficits: shortage of trained investigators, prosecutors and judicial officers, low conviction rates, and inadequate digital-forensic laboratories collectively impede outcomes; independent analyses suggest that in 2023 only around 22 per cent of registered cybercrimes were charged and fewer than 3 per cent resulted in conviction.[38] Sixth, victim under-reporting: shame, procedural complexity, and the perceived futility of complaint remain systemic obstacles.

1 August 2024.

[36]Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (General Data Protection Regulation), OJ L 119, 4 May 2016; Directive (EU) 2022/2555 (NIS 2 Directive), OJ L 333, 27 December 2022.

[37]Computer Fraud and Abuse Act, 18 U.S.C. § 1030 (United States); Computer Misuse Act 1990, c. 18 (United Kingdom); Online Safety Act 2023, c. 50 (United Kingdom).

[38]Crime in India 2023, supra note 1.

ISSN: 2582-6433

Beyond investigation, private-sector cooperation, real-time takedown, and platform disclosure of first-originator information remain contested. The Sahyog portal launched by the Ministry of Electronics and Information Technology has, by March 2025, facilitated the blocking of 1,11,185 items of suspicious content, including 83,867 WhatsApp accounts linked to cybercrime.[39] Yet the scale of the threat vastly outpaces institutional capacity, and this asymmetry is arguably the single greatest determinant of enforcement outcomes. The domestic Citizen Financial Cyber Fraud Reporting and Management System, though evidently useful in freezing fraudulent transactions in real time, remains asymmetric with the pace of scam industrialisation, and the geographic concentration of scam infrastructure in Southeast Asia leaves Indian investigators reliant on diplomatic and mutual-legal-assistance pathways whose operational tempo is measured in months rather than the minutes necessary to intercept illicit fund flows.

XII. Liability Framework and the Attribution Problem

The liability question is multi-layered. Direct criminal liability of the human perpetrator is uncontroversial where mens rea can be established. Developers of AI systems used for criminal purposes are not, without more, criminally liable under Indian law; general principles of abetment, conspiracy and negligence apply, but with significant evidentiary hurdles. Deployers and platform operators may attract civil and, in specified circumstances, criminal liability under the IT Act and the Intermediary Rules, if they fail to observe due diligence or comply with valid takedown directions. Corporate liability under Section 85 of the IT Act extends to officers where offences are committed by companies with their consent, connivance, or attributable neglect. The doctrinal question whether AI systems themselves can be treated as instruments, agents or "legal persons" for the purposes of liability remains open and, as a matter of Indian law, should be answered with caution: legal personhood is not the appropriate response to attribution difficulty, since it risks displacing accountability from the human actors who deploy, monetise, and instrumentalise these systems.

XIII. Artificial Intelligence as a Defensive Cybersecurity Tool

Artificial Intelligence is simultaneously the most powerful defensive technology available to modern security practice, automated threat detection, anomaly identification, malware classification, and forensic reconstruction all benefit from machine learning. Yet defensive AI

[39] I4C Performance and Interventions, supra note 3.

ISSN: 2582-6433

is not consequence-free: false positives may result in unjustified restriction of lawful activity; predictive policing risks entrenching bias; and biometric surveillance implicates the privacy interests recognised in Puttaswamy.[40] The regulatory task is to enable defensive deployment while installing meaningful safeguards, including algorithmic transparency, human oversight of adverse decisions, and independent auditing.

XIV. Regulatory Gaps and Reform Framework

Six concrete regulatory gaps warrant attention. First, no Indian statute expressly addresses AI-generated synthetic content, deepfake impersonation, or non-consensual intimate deepfakes, though the draft 2025 amendments to the Intermediary Rules mark an initial step. Second, evidentiary standards under Section 63 of the BSA require clarification of how AI-generated material, including deepfakes should be authenticated and how expert witnesses on AI ought to be qualified. Third, no dedicated cybercrime coordination legislation places I4C on a statutory footing, though it has been designated an attached office of the Ministry of Home Affairs with effect from 1 July 2024. Fourth, cross-border evidence gathering remains constrained by India's non-participation in the Budapest Convention and its cautious approach to the Hanoi Convention. Fifth, the DPDP Act contemplates automated processing without providing a specific enforcement pathway against AI-facilitated data crimes. Sixth, corporate accountability for AI safety, covering testing, red-teaming, incident reporting, and safe deployment, remains substantially absent from Indian law and must not be left to voluntary codes alone.

XV. Recommendations

A calibrated reform agenda ought to include the following measures. First, targeted legislative amendment: the IT Act, or a dedicated Digital India Act, should introduce express offences for the creation, distribution and monetisation of non-consensual synthetic intimate imagery, deepfake-based cheating and impersonation, and coordinated inauthentic behaviour that produces measurable harm; definitions must be narrowly tailored to satisfy the constitutional standards articulated in Shreya Singhal. Second, evidentiary reform: a statutory or judicial protocol should govern authentication of AI-generated evidence, empanelment of AI-forensic experts, and admissibility of provenance-metadata such as content-credential attestations of the C2PA family. Third, capacity building: a national AI-forensic infrastructure, adequately staffed

[40]Puttaswamy, supra note 20.

ISSN: 2582-6433

cybercrime cells in every district, and specialised prosecutors and judges are essential.

Fourth, mandatory incident reporting: the CERT-In six-hour framework should be complemented by AI-specific reporting obligations for developers and deployers of foundation models above defined capability thresholds, together with a national AI-safety incident registry. Fifth, public-private cooperation and threat-intelligence sharing on a defined legal footing, including safe-harbour protection for good-faith disclosure. Sixth, international engagement: India's decision whether to accede to the Budapest Convention or ratify the Hanoi Convention should be taken on the merits of each; either path materially enhances mutual legal assistance and cross-border evidence access. Seventh, judicial and prosecutorial training on AI, machine-learning forensics, and synthetic evidence must be institutionalised through the National Judicial Academy and the State Judicial Academies. Eighth, a national AI-cybercrime strategy should articulate a whole-of-government approach across the Ministry of Electronics and Information Technology, the Ministry of Home Affairs, the Reserve Bank of India, the Telecom Regulatory Authority of India, the Department of Telecommunications, and State police forces, with measurable objectives, budgetary lines, and periodic public reporting. Ninth, victim-support mechanisms—including psychological counselling, legal aid, and reintegration support for those trafficked into Southeast Asian scam compounds—must be scaled substantially. Tenth, corporate accountability and cyber-insurance obligations for large deployers of consumer-facing AI must be considered as part of a broader liability architecture.

XVI. Conclusion

The transformation of cybercrime by Artificial Intelligence is not merely quantitative. It is a qualitative shift in the architecture, economics, and reach of criminal conduct. India's existing legal framework, technology-neutral in aspiration, covers a substantial part of the substantive terrain but reveals significant institutional and doctrinal deficits when tested against attribution difficulties, synthetic evidence, cross-border operations, and autonomous adaptation. Constitutional guarantees, particularly the privacy jurisprudence of Puttaswamy, provide a principled foundation on which both offensive-facing regulation and defensive AI deployment must be built. Comparative frameworks, the European Union Artificial Intelligence Act, the Budapest Convention and its Second Additional Protocol, the Hanoi Convention, the Computer Fraud and Abuse Act, and the United Kingdom Computer Misuse and Online Safety Acts - offer partial models but no complete template. The reforms proposed here are neither unusually ambitious nor exceptionally novel; they are the ordinary responsibilities of a mature legal

ISSN: 2582-6433

system responding to a transformative technology. The next phase of Indian cyber-law reform must be evidence-based, constitutionally grounded, institutionally deliverable, and internationally cooperative. Only then will the legal system keep faith with the citizens whose safety in cyberspace increasingly depends upon it.