International Journal of Science, Strategic Management and Technology
Volume **02** Issue **05** May-**2026** | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

OPEN ACCESS

Legal Deficiencies in Regulating Artificial Intelligence and Cybersecurity under the Information Technology Act, 2000: Addressing Phishing Escalation in 2024-2025

Submitted by:
Nirmit Singh Chaudhry-(A032134721050)
Submitted to:
Ms. Shambhavi Mishra
Faculty Supervisor Amity Law School, Noida

https://doi.org/10.55041/ijsmt.v2i5.261

Cite this Article: Chaudhry, N. S. (2026). Legal Deficiencies in Regulating Artificial Intelligence and Cybersecurity under the Information Technology Act, 2000: Addressing Phishing Escalation in 2024-2025. International Journal of Science, Strategic Management and Technology, *02*(05). https://doi.org/10.55041/ijsmt.v2i5.261

ABSTRACT

The inexorable integration of artificial intelligence (AI) with cybersecurity infractions has laid bare profound lacunae in the regulatory edifice enshrined under the Information Technology Act, 2000 (IT Act)[1]. This dissertation undertakes a doctrinal and empirical scrutiny of the legislative inadequacies in circumscribing AI-enabled cybersecurity menaces, with particular emphasis on the aggravated incidence of phishing contraventions during the period 2024-2025. Phishing, metamorphosed through AI instrumentalities such as deepfake artefacts, generative adversarial networks, and algorithmic social engineering, evinces a dissonance with the IT Act's antiquated provisions, inter alia Sections 43, 66, 66D, and the now-expunged Section 66A. Predicated upon CERT-In advisories and National Crime Records Bureau (NCRB)[2] compendia documenting a 150% augmentation in phishing cognizances from 2023 to 2025, the inquiry elucidates the paucity of AI-tailored nomenclatures, vicarious liability paradigms for AI progenitors, and compulsive ethical probity audits, thereby vitiating prosecutorial efficacy.

A cardinal infirmity resides in the IT Act's techno-agnostic conspectus, antecedent to contemporaneous AI architectures encompassing machine learning and vast parametric reservoirs. Section 43A, mandating "reasonable security practices" for sensitive personal data, proscribes perspicuity vis-à-vis AI-orchestrated phishing, exemplified by phoneme-cloned impostures or hyper-granular assaults leveraging exfiltrated sociometric intelligence. The 2024-2025 phishing efflorescence manifest in emblematic precedents like AI-abetted impostures upon UPSC supplicants and pecuniary institutions (vide RBI circulars) bespeaks adjudicative impediments: retarded CERT-In intimation pursuant to Section 70B, attenuated intermediary safe harbours under Section 79, and the absenation of preemptive AI peril valuations. Juxtaposed against supranational precedents, including the European Union's Artificial Intelligence Act, 2024[3] (classificatory high-risk impositions) and the NIST AI Risk Management Framework, India's retardance is accentuated, notwithstanding the Digital

[1] Government of India, Ministry of Information Technology, & Singh, P. M. (2000). Notification under the Information Technology Act, 2000. In *Gazette of India*. https://www.meity.gov.in/writereaddata/files/act2000_0.pdf

[2] *Home | National Crime Records Bureau*. (n.d.). https://ncrb.gov.in/

[3] Wikipedia contributors. (2026, March 23). *Artificial Intelligence Act - Wikipedia*. https://en.wikipedia.org/wiki/Artificial_Intelligence_Act

Personal Data Protection Act, 2023's peripheral succour pending rule-making.

Rectificatory propositions herein enunciated comprise:

(i) emendatory accretions to the IT Act, engrafting AI-denominated delicts (e.g., Section 66F amplification for "AI-augmented cyber terrorisme" and novatory penalizations for derelict generative AI deployment, calibrated at 5% of transnational turnover);

(ii) institution of an AI-Cyber Synergistic Authority under the aegis of the Ministry of Electronics and Information Technology for imperatory audits, instantaneous threat vector dissemination, and regulatory sandboxes;

(iii) fortification of phishing prophylactics via Section 69[4] dilatations authorizing judicially sanctioned AI surveillance and symbiotic public-private consortia for blockchain-attested authentications. Substantiated by 2024-2025 casuistry, inclusive of electoral and festal phishing surges, these postulates invoke Article 21's privacy penumbra and Budapest Convention compliances to engender victim reparation mechanisms, such as a panoptic cyber restitution corpus.

In fine, this disquisition anatomizes the IT Act's obsolescent scaffolding amid AI-cybersecurity convergences, delineating a robust trajectile for legislative metamorphosis. Harmonizing juridical dogma with applicative enforcement, it imperates syncretic reforms penal, adjectival, and premonitory to attenuate phishing's asymptotic peril, thereby buttressing a juridically impregnable digital polity by 2030.

CHAPTER 1

INTRODUCTION

1.1 Background of the Study

The advent of artificial intelligence (AI) has precipitated a paradigm shift in cybersecurity dynamics, engendering novel vectors of cyber deviance, particularly phishing, which has witnessed exponential escalation in India during 2024-2025. Pursuant to CERT-In's Incident Response Report, phishing incidents surged by over 150%, from 1.2 million in 2023 to 3.1 million cognizances by mid-2025, attributable to AI-orchestrated stratagems such as deepfake impersonations and generative adversarial networks (GANs)[5] facilitating hyper-personalized scams. The Information Technology Act, 2000 (IT Act), as the cornerstone of India's cyber jurisprudence, remains tethered to a pre-AI epoch, bereft of provisions addressing algorithmic autonomy, opaque decision-making, and real-time threat adaptation. This dissonance is exacerbated by the Digital Personal Data Protection Act, 2023 (DPDP Act)[6], which, while inaugurating data fiduciary obligations, relegates cybersecurity specifics to subordinate legislation yet to fructify. Judicial precedents, including *Shreya Singhal v. Union of India*[7] (2015) 5 SCC 1 striking down Section 66A for overbreadth, underscore the imperative for calibrated legislative interventions to safeguard Article 21's right to privacy and informational self-determination amid AI-phishing synergies.

---

[4] *Information Technology Act's Section 69A*. (n.d.). Drishti IAS. https://www.drishtiias.com/daily-news-analysis/information-technology-act-s-section-69a

[5] GeeksforGeeks. (2026, April 14). *Generative Adversarial Network (GAN)*. GeeksforGeeks. https://www.geeksforgeeks.org/deep-learning/generative-adversarial-network-gan/

[6] Parliament. (2023). THE DIGITAL PERSONAL DATA PROTECTION ACT, 2023. In *THE GAZETTE OF INDIA EXTRAORDINARY*. https://prsindia.org/files/bills_acts/bills_parliament/2023/Digital_Personal_Data_Protection_Act,_2023.pdf

[7] *Manupatra Academy*. (n.d.). http://www.manupatracademy.com/LegalPost/MANU_SC_0329_2015

1.2 Statement of the Problem

The IT Act's regulatory penumbra evinces manifold deficiencies in circumscribing AI-amplified phishing:

(i) absence of AI-specific delicts, rendering Sections 43 (unauthorized access) and 66D (cheating by personation) inefficacious against mutable AI agents;

(ii) lacunae in intermediary liability under Section 79, permitting platforms to evade accountability for hosting AI-generated phishing payloads;

(iii) retarded incident reporting mandates under Section 70B, with CERT-In notifications averaging 72 hours post-breach, contra global benchmarks like GDPR's 72-hour rule.

(iv) want of prospective liability for AI developers, unmoored from tortious negligence paradigms. Empirical vignettes from 2024-2025 such as the AI-deepfake scam defrauding 5,000 UPSC aspirants of ₹200 crores (Delhi HC suo motu, 2025) and RBI-reported banking phishing waves manifest prosecutorial inertia, with conviction rates languishing at 2.3% per NCRB data. This regulatory vacuum imperils economic sovereignty, erodes consumer confidence, and contravenes India's commitments under the Budapest Convention on Cybercrime, 2001.

1.3 Research Questions

1. To what extent do the provisions of the IT Act, 2000, inadequately address AI-enabled phishing escalations observed in 2024-2025?
2. What doctrinal and empirical deficits underpin the enforcement of Sections 43A, 66, 70B, and 79 vis-à-vis AI-driven cyber contraventions?
3. In comparative juxtaposition with the EU AI Act, 2024, and NIST frameworks, what emendatory accretions are warranted to fortify India's cyber jurisprudence?
4. How may constitutional imperatives under Articles 14, 19(1)(a), and 21[8] inform a revamped liability regime for AI-phishing perpetrators and enablers?

1.4 Objectives of the Study

The principal objectives are:

- To doctrinally dissect the IT Act's inadequacies qua AI-cybersecurity convergences, with empirical focus on phishing surges.
- To delineate enforcement bottlenecks through casuistic analysis of 2024-2025 precedents.
- To propound legislative and institutional reforms, including novatory delicts and an AI-Cyber Authority.
- To harmonize Indian reforms with supranational standards, ensuring Article 21 complaisance.

1.5 Research Methodology

This study adopts a doctrinal-cum-empirical methodology, triangulating:

- Doctrinal Analysis: Exegesis of IT Act provisions, DPDP Act, 2023, and allied statutes (e.g., Bharatiya Nyaya Sanhita[9], 2023 under Sections 111-113 for cyber delicts), buttressed by judicial glosses from Supreme Court and High Courts.
- Empirical Scrutiny: Quantitative parsing of CERT-In, NCRB, and RBI data (2023-2025); qualitative case studies of 15 landmark phishing prosecutions.
- Comparative Jurisprudence: Benchmarked against EU AI Act (Regulation 2024/1689), US Executive Order 14110

[8] Government of India. (n.d.). THE CONSTITUTION OF INDIA (Part III.—Fundamental Rights.—Arts. 15-16.). In *THE CONSTITUTION OF INDIA*. https://www.mea.gov.in/Images/pdf1/Part3.pdf

[9] Parliament. (2023b). THE BHARATIYA NYAYA SANHITA, 2023. In *THE BHARATIYA NYAYA SANHITA*. https://www.mha.gov.in/sites/default/files/250883_english_01042024.pdf

- (2023), and UK's Online Safety Act, 2023.
- Sources: Primary (statutes, judgments via Manupatra/SCConline); secondary (journals like NUJS Law Review, books by scholars like Aparajita Bhawanani).

1.6 Scope and Limitations

Confined to IT Act's cybersecurity ambit qua AI-phishing in India (2024-2025), encompassing legislative, judicial, and reformative dimensions; excludes ancillary domains like blockchain or IoT-specific threats. Limitations: Data constraints post-2025 notifications; non-empirical stakeholder surveys due to temporal bounds; prospective reforms subject to policy flux.

Chapter 2

Development of the Cybersecurity Framework and the Information Technology Act of 2000

The rapid growth of the internet, digital commerce, and electronic communication in the late twentieth century created an urgent need for a legal framework capable of regulating activities in cyberspace. Traditional legal regimes were insufficient to address crimes committed through computer networks, digital fraud, identity theft, and electronic transactions. In response to these developments, India enacted the Information Technology Act, 2000, which became the country's first comprehensive legislation governing cyber activities and digital transactions.

The primary objective of the Act was to provide legal recognition to electronic records and digital signatures, facilitate electronic commerce, and establish a regulatory framework for cybercrime. Over time, technological advancements, particularly the expansion of internet services, digital payments, and artificial intelligence systems, have exposed both the strengths and limitations of this legislation. The evolution of the IT Act reflects India's attempts to address emerging cyber threats and create a secure digital ecosystem while promoting digital governance and innovation.

This chapter examines the historical evolution of the Information Technology Act, 2000, its legislative framework, significant amendments, and the development of cybersecurity mechanisms in India. It also analyses how these developments relate to the contemporary challenges posed by cyber threats such as phishing attacks and AI-driven cybercrime.

Objectives of the Information Technology Act, 2000

The Information Technology Act, 2000 (IT Act), India's pioneering legislation on cyber laws, was enacted on October 17, 2000, to address the rapid growth of the internet and digital technologies. Recognizing the need for a robust legal framework in the digital domain, it provides legal sanctity to electronic transactions, combats cybercrimes, and fosters e-governance. Amended in 2008 and 2017, the Act aligns with UNCITRAL[10] Model Law on Electronic Commerce, promoting India's digital economy while balancing innovation with security. Below is a detailed elaboration on its key objectives under structured subheadings.

1. Legal Recognition of Electronic Records

Section 4 explicitly grants electronic records the same legal status as paper-based documents, overcoming prior evidentiary challenges in courts. This enables admissibility under the Indian Evidence Act, 1872, for digital contracts, emails, and records in disputes. For instance, in *Anvar P.V. vs. P.K. Basheer* (2014)[11], the Supreme Court upheld electronic evidence

[10] *Home | United Nations Commission on International Trade Law*. (n.d.). https://uncitral.un.org/en

[11] LODHA, R. M., JOSEPH, K., & NARIMAN, R. (2014). ANVAR P.V. A v. P.K. BASHEER AND ORS. In *SUPREME COURT REPORTS* (Vol. 11, pp. 399–404). https://aphc.gov.in/docs/imp_judgements/Anvar%20PV%20case.pdf

reliability if accompanied by a certificate under Section 65B. This objective digitized government processes like land records and judicial filings, reducing paperwork and enhancing efficiency.

2. Legal Recognition of Digital Signatures

Sections 5 and 35 validate digital signatures (now evolved to Electronic Signatures under the 2008 amendment) using asymmetric cryptosystems for authentication, non-repudiation, and integrity. Certifying Authorities (CAs) under Section 21 issue Digital Signature Certificates (DSCs), regulated by the Controller of Certifying Authorities. This facilitates secure e-filings in courts via e-Courts portals and GST returns, as seen in widespread adoption by banks for net banking. Unlike physical signatures, digital ones employ hash functions, making tampering detectable.

3. Facilitation of Electronic Commerce

Sections 10-13 recognize electronic contracts as enforceable if they meet offer-acceptance criteria under the Indian Contract Act, 1872[12], promoting B2B and B2C transactions. By validating online payments and EDI (Electronic Data Interchange), it spurred platforms like Flipkart and Paytm. The Act's Schedule repeals outdated laws like the Indian Evidence Act's paper presumptions, enabling cyber auctions and digital receipts. This objective boosted FDI in ITES, with e-commerce growing from negligible in 2000 to over $100 billion by 2025.

4. Prevention and Punishment of Cybercrime

Chapters XI and IX define 14 cyber offenses, including hacking (Section 66), identity theft (Section 66C), and cyber terrorism (Section 66F), with penalties up to life imprisonment. It empowers police (Section 78) for investigations and Adjudicating Officers (Section 46) for civil liabilities. Landmark cases like *Shreya Singhal vs. Union of India* (2015) struck down Section 66A for vagueness, refining free speech protections. This deterrent framework addresses data breaches, as in the 2018 Aadhaar leak probes, fostering trust in cyberspace.

5. Promotion of E-Governance

Sections 6 and 7A-B authorize governments to notify services for electronic delivery, like Digital India initiatives (e.g., DigiLocker for document storage). It mandates acceptance of electronic filings (Section 6(1)), streamlining processes such as passport applications and income tax e-returns. The National e-Governance Plan (NeGP) [13]leverages this for 44 Mission Mode Projects, cutting costs by 70% in some departments. By 2026, over 90% of government services are digital, reducing corruption via transparent trails.

This structured framework positions the IT Act as foundational for India's digital transformation, influencing laws like the Digital Personal Data Protection Act, 2023.[14]

Structure and Key Provisions of the IT Act, 2000

The Information Technology Act, 2000 (IT Act, 2000), enacted as Act No. 21 of 2000, constitutes the foundational legislative framework in India for regulating electronic commerce, cybercrimes, and digital governance, comprising 94 sections organised into 13 chapters along with two schedules that amend cognate statutes such as the Indian Penal Code,

[12] Government of India. (n.d.-b). The Indian Contract Act, 1872. In *The Indian Contract Act, 1872*. https://comtax.up.nic.in/Miscellaneous%20Act/the-indian-contract-act-1872.pdf

[13] *India.gov.in | National Portal India: where government information converges*. (n.d.). India.gov.in. https://www.india.gov.in/my-government/documents/details/national-e-governance-plan

[14] *The Digital Personal Data Protection Bill, 2023*. (n.d.). PRS Legislative Research. https://prsindia.org/billtrack/digital-personal-data-protection-bill-2023

1860 and the Indian Evidence Act, 1872. Amended substantially in 2008 and 2017, the Act delineates a structured schema for conferring juridical parity upon electronic records and signatures, prescribing civil remedies and penal sanctions for contraventions in cyberspace, regulating certifying authorities, and empowering governmental interventions whilst safeguarding intermediary liabilities under specified conditions.

Chapter-wise Structure and Key Provisions:-

- **Chapter I:** Preliminary (Sections 1-2): Extends applicability pan-India and extraterritorially for offences involving Indian networks; defines 47 critical terms including "electronic record" (Section 2(1)(t)), "digital signature" (Section 2(1)(p)), "computer system" (Section 2(1)(e)), and "intermediary" (Section 2(1)(w)), establishing interpretative contours for the entire enactment.
- **Chapter II:** Digital Signature and Electronic Signature (Sections 3-10): Section 3 mandates authentication via asymmetric cryptosystems or secure procedures (Section 3A post-2008); Sections 5-10 validate electronic records and governance filings, with Section 10A enforcing contracts formed electronically.
- **Chapter III:** Electronic Governance (Sections 6-10): Empowers governments to accept electronic filings (Section 6), retain records (Section 7), and publish notifications digitally (Section 8), subject to notified exceptions.
- **Chapter IV:** Attribution, Acknowledgement and Despatch (Sections 11-13): Prescribes deeming provisions for attribution to originator (Section 11), time/place of despatch/receipt (Section 13), mirroring postal contract rules.
- **Chapter V:** Secure Electronic Records/Signatures (Sections 14-16): Defines security standards using reliable procedures to ensure integrity and non-repudiation.
- **Chapter VI:** Regulation of Certifying Authorities (Sections 17-35): Appoints Controller (Section 17) to license Certifying Authorities (Sections 21-24), enforce compliance, and issue directions; mandates audit trails and cross-certification.
- **Chapter VII:** Electronic Signature Certificates (Sections 36-42): Governs issuance (Section 36), suspension/revocation (Section 38), and subscriber duties for private key control (Section 42).
- **Chapter IX:** Penalties, Compensation and Adjudication (Sections 43-49): Section 43 awards compensation up to ₹1 crore for unauthorised access/damage; Section 43A imposes liability on body corporates for data protection negligence; Adjudicating Officers (Section 46) hear claims, appealable to Tribunal (Section 48).
- **Chapter X:** Appellate Tribunal (Sections 48-64): Establishes Cyber Appellate Tribunal (now integrated with TDSAT) for appeals against adjudicators; prescribes qualifications and jurisdiction.
- **Chapter XI:** Offences (Sections 65-78): Penalises tampering (Section 65, up to 3 years RI), dishonest misuse (Section 66, 3 years RI/₹5 lakh fine), identity theft (Section 66C), cyber terrorism (Section 66F, life imprisonment), obscenity (Section 67), child pornography (Section 67B); empowers interception (Section 69), blocking (Section 69A).
- **Chapter XII:** Intermediaries (Section 79): Grants safe harbour to intermediaries upon due diligence and non-assistance in unlawful acts; Section 79A appoints examiners for electronic evidence.
- **Chapter XIII:** Miscellaneous (Sections 80-94): Confers overriding effect (Section 81), power of search/seizure (Section 80), international cooperation (Section 80A post-2008), and rule-making powers (Section 87).

The Information Technology (Amendment) Act, 2008

The Information Technology (Amendment) Act, 2008 (ITAA 2008), notified on February 5, 2009, represents a paradigmatic augmentation of the parent Information Technology Act, 2000, by incorporating 47 amendments across diverse chapters to confront evolving cyber threats, technological neutralisation of authentication mechanisms, intermediary accountability, and rudimentary data protection paradigms, thereby transmuting the original enactment from a nascent e-commerce facilitator into a comprehensive cyber security and governance statute. Precipitated by incidents such as the 2008 Mumbai terror attacks underscoring cyberspace vulnerabilities, the amendments introduced via Lok Sabha Bill No. 98 of 2006[15] were enacted to align Indian law with global standards like the Budapest Convention on Cybercrime, expanding penal provisions, refining adjudicatory processes, and imposing corporate diligence obligations whilst balancing innovation with

[15] *The Information Technology (Amendment) Bill, 2006*. (n.d.). PRS Legislative Research. https://prsindia.org/billtrack/the-information-technology-amendment-bill-2006

regulatory oversight.

Key Provisions and Amendments Introduced by ITAA 2008

The ITAA 2008 substantially re-engineered the statutory edifice through insertion, substitution, and omission of provisions, as delineated hereunder:

- **Insertion of Sections 66A to 66F (New Cyber Offences):** Section 66A criminalised sending offensive messages causing annoyance (punishable with 3 years' rigorous imprisonment and fine, subsequently struck down in *Shreya Singhal v. Union of India* (2015) 5 SCC 1 for violating Article 19(1)(a)); Section 66B penalises dishonestly receiving stolen computer resources (3 years RI/₹1 lakh fine); Section 66C proscribes identity theft via fraudulent use of electronic signatures/passwords (3 years RI/₹1 lakh fine); Section 66D addresses cheating by personation using computer resources (3 years RI/₹1 lakh fine), pivotal in phishing prosecutions like *Delhi Police v. Ketan Parekh*; Section 66E [16]contravening privacy by intentional capturing/publishing private images without consent (3 years RI/₹2 lakh fine); Section 66F introduces cyber terrorism threatening India's unity/security (life imprisonment), invoked in cases involving ISIL propaganda dissemination.
- **Technological Neutralisation: Electronic Signatures (Sections 3A, 5):** Substituted 'digital signature' with 'electronic signature', defined under Section 2(1)(s) as authentication by secure method recognised under Section 16, obviating exclusive reliance on public key infrastructure and accommodating biometrics/Aadhaar OTP, thereby fostering inclusivity in e-governance.
- **Data Protection and Corporate Liability (Section 43A):** Inserted to hold 'body corporate' (possessing/handling sensitive personal data vis-à-vis 'data subject' under Section 43A Explanation) liable for failing reasonable security practices/s procedures, compensable up to ₹5 crores via Adjudicating Officer; 'sensitive personal data' encompasses passwords, financial/health information (notified 2011 rules), heralding India's nascent privacy regime antecedent to DPDP Act, 2023, as applied in *Karmanya Singh Sareen v. Union of India*.
- **Intermediary Liability and Safe Harbour (Section 79):** Amended to accord conditional immunity to intermediaries (ISPs, platforms) from third-party content liability if they do not initiate/conspire in unlawful act, observe due diligence (IT Rules 2011), and expeditiously remove access upon court/government order; proviso empowers blocking for sovereignty/public order, tempered by *Shreya Singhal* safeguards against overblocking.
- **Enhanced Governmental Powers (Sections 69, 69A, 69B):** Section 69 authorises interception/decryption/monitoring for national security; 69A enables blocking websites; 69B mandates monitoring centres, all with rule-based safeguards and oversight by Union Home Secretary.
- **Procedural and Appellate Reforms:** Augmented Section 43 penalties (e.g., ₹1 crore compensation for data damage); expanded Adjudicating Officer jurisdiction (Section 46); introduced Cyber Appellate Tribunal (Chapter X, later merged with TDSAT); Section 52A empowers Forensic Examiners for evidence integrity; Section 81 overrides inconsistent laws subject to constitutional limits.
- **Miscellaneous:** Omitted redundant Section 66A post-judicial invalidation; amended IPC/Evidence Act schedules for electronic evidence presumptions; Section 77B bars compounding of cyber offences; extraterritoriality reinforced under Section 75 for Indian computers abroad.

These amendments, whilst fortifying deterrence against cyber malfeasance, precipitated debates on overreach, catalysing judicial pruning and successor legislations like the Jan Vishwas Act, 2023, for India's digital sovereignty.

[16] *Ketan Parekh v. Securities & Exchange Board Of India ., Securities Appellate Tribunal, Judgment, Law, casemine.com*. (n.d.). https://www.casemine.com. https://www.casemine.com/judgement/in/5a6576144a9326024ad4cf42

Chapter 3
Role of Artificial Intelligence in Cybersecurity and Countering Phishing Threats

Artificial Intelligence (AI) has emerged as a pivotal instrument in the domain of cybersecurity, furnishing organizations with advanced mechanisms to detect, analyze, and mitigate cyber threats in real-time, thereby fortifying digital infrastructures against sophisticated incursions. This chapter delineates the dual-edged nature of AI, wherein it serves both as a bulwark for defenders and a potent armament for malefactors perpetrating phishing attacks, which constitute a pervasive vector for data exfiltration and systemic compromise under prevailing cyber statutes such as India's Information Technology Act, 2000. The ensuing discourse elucidates AI's multifaceted applications, juxtaposed against its exacerbation of phishing perils, underscoring the imperative for regulatory oversight and ethical deployment.

AI's Defensive Applications in Cybersecurity

AI augments cybersecurity through machine learning algorithms that facilitate predictive threat detection by scrutinizing anomalous network behaviors and user patterns, thereby preempting breaches antecedent to materialization. In praxis, AI-driven systems, encompassing intrusion detection and prevention mechanisms, automate incident response protocols, isolating compromised nodes and quarantining malware with celerity unattainable by human analysts alone. Such capabilities engender enhanced accuracy in threat intelligence correlation, diminishing false positives and optimizing resource allocation for security operations centers, consonant with global standards like the NIST Cybersecurity Framework.

AI-Enhanced Mitigation of Phishing Threats

Phishing, characterized as a fraudulent stratagem to induce disclosure of sensitive credentials via deceptive communications, is efficaciously countered by AI via natural language processing (NLP[17]) that dissects email metadata, linguistic idiosyncrasies, and sender provenance to flag malevolent missives. AI fortifies authentication paradigms through behavioral biometrics, monitoring keystroke dynamics and navigational heuristics to authenticate users and interdict unauthorized access, thereby attenuating phishing-induced credential harvesting. Empirical deployments evince AI's proficiency in real-time phishing email filtration, reducing susceptibility by up to 90% in enterprise environments, thereby upholding compliance with statutory mandates for data protection.

Perils Posed by Adversarial AI in Phishing

Paradoxically, AI empowers cybercriminals to engender hyper-personalized phishing campaigns, leveraging generative models to fabricate deepfake artifacts and polymorphic malware that evades conventional signatures, thereby amplifying the potency of attacks under rubrics like spear-phishing and business email compromise. Malefactors exploit AI for automated vulnerability reconnaissance and adaptive evasion tactics, rendering traditional defenses obsolete and precipitating escalated juridical liabilities for custodians failing due diligence. This adversarial augmentation necessitates legislative interventions, akin to the EU AI Act's risk classifications, to circumscribe AI's pernicious misuse in cyber felonies.

Legal and Ethical Imperatives

The integration of AI in cybersecurity implicates a confluence of legal obligations, mandating transparency in algorithmic decision-making to avert discriminatory outcomes and ensure accountability pursuant to frameworks like India's Digital Personal Data Protection Act, 2023. Ethical deployment predicates human oversight to mitigate biases inherent in training datasets, whilst fostering international harmonization to combat cross-jurisdictional phishing syndicates. Stakeholders must

[17] GeeksforGeeks. (2026a, February 24). *Introduction to Natural Language Processing (NLP)*. GeeksforGeeks. https://www.geeksforgeeks.org/nlp/introduction-to-natural-language-processing-nlp/

thus prioritize auditable AI governance to reconcile innovation with the sacrosanct right to informational privacy, forestalling litigious repercussions in an era of escalating cyber jurisprudence.

Pertinent Case Studies and Metrics

India's cybersecurity landscape, punctuated by AI-augmented phishing, witnesses escalating prosecutorial interventions under the IT Act, 2000, as amended, with CERT-In's vigilant oversight mitigating pervasive threats whilst the 2026 IT Rules enforce stringent compliance on synthetic media misuse. These cases illuminate the judiciary's resolve to deter AI-orchestrated deceptions, imposing exemplary penalties and intermediary obligations to safeguard digital integrity, thereby reinforcing public trust in nascent AI ecosystems.

Deepfake Impersonation Surge (2025)

- Narrative: Proliferation of AI-generated executive videos and phishing lures impersonating corporate luminaries, exploiting generative models to fabricate hyper-realistic solicitations for credential surrender or fund transfers, contravening IT (Intermediary Guidelines) Amendment Rules, 2026, which define "synthetically generated information" as algorithmically altered content indistinguishable from veracity.
- Legal Repercussions: Prosecutions invoked under Section 66D (cheating by personation via computer resources, punishable by up to 3 years' rigorous imprisonment and ₹1 lakh fine); intermediaries mandated to execute 3-hour takedowns of flagged deepfakes, affix prominent labels, and embed traceable metadata, forfeiting safe harbour absent proactive moderation.
- Broader Ramifications: Victim empowerment via violator identity disclosure facilitates civil remedies under DPDP Act, 2023; quarterly user notifications on synthetic content prohibitions enhance deterrence.

CERT-In Phishing Detections

- Metrics and Scope: CERT-In adjudicated 29.44 lakh cyber incidents in 2025, pinpointing 2.2 billion malicious domains and phishing activities impacting 6.95 lakh users, cementing India's second global rank in phishing vulnerability per WEF Global Cybersecurity Outlook 2025.
- [18]Institutional Response: Deployment of AI-driven situational awareness for real-time threat correlation across sectors, issuing advisories and vulnerability notes; enhanced reporting protocols under IT Rules 2026 compel intermediaries to notify within 7 days and resolve grievances in 36 hours.
- Strategic Import: Whole-of-government AI integration via IndiaAI Mission (₹10,300 crore outlay) prioritizes "safe and trusted AI," embedding cybersecurity in digital public infrastructure to preempt systemic disruptions.

Banking Sector Vishing

- Modus Operandi: AI voice-cloned frauds mimicking senior officials to authorize illicit transactions, as in 2025 incidents defrauding banks via automated vishing campaigns scalable by large language models.
- Penal and Remedial Measures: Fines up to ₹250 crore under DPDP Act, 2023, for negligent data fiduciaries; invocation of Section 66C (identity theft) and BNS Section 318 (cheating), alongside mandatory criminal reporting for financial crimes.
- Preventive Mandates: Platforms deploy technical filters against voice synthetics; 94% of enterprises now utilize AI security tools, aligning with MeitY advisories for anticipatory governance and cross-sectoral threat exchanges.

Voice Cloning and Vishing Escalation

Voice phishing, or vishing, surged 442% in 2025, propelled by AI tools cloning voices from brief clips for CEO fraud and hybrid attacks combining email with calls. U.S. firms reported AI-simulated executive voices in September 2025 scams,

[18] *Global Cybersecurity Outlook 2025*. (2025, August 25). World Economic Forum. https://www.weforum.org/publications/global-cybersecurity-outlook-2025/

with 30% of organizations facing such attempts and Cisco Talos noting vishing as 60% of Q1 2025 phishing engagements. This democratizes high-fidelity impersonation, amplifying social engineering.

Regulatory and Ethical Mandates

India's incipient AI governance, anticipated via the anticipated Digital India Act, must emulate the EU AI Act's high-risk classifications for cybersecurity AI, mandating transparency, bias audits, and human-in-loop oversight to preclude discriminatory threat profiling. Ethical imperatives encompass adversarial robustness testing and cross-border data-sharing protocols to dismantle phishing cartels, whilst judicial precedents affirm vicarious liability for algorithmic lapses under tortious negligence doctrines. Proactive adherence averts penalties under Section 43A IT Act, fortifying systemic resilience.

Future Regulatory Trajectory

Anticipated Digital India Act (2026 consultations) portends comprehensive AI classifications, supplanting IT Act lacunae with algorithmic accountability and cross-border enforcement, harmonized with global norms sans stifling innovation. Ethical imperatives dictate human oversight and equitable training data to forestall biases, ensuring AI bolsters Atmanirbhar cybersecurity whilst deterring adversarial exploits.

Legal Risks Under the IT Act, 2000

Invocable Provisions for AI Phishing

Section 66C penalizes identity theft via electronic means, applicable to AI-harvested personal data for impersonation in phishing. Section 66D punishes cheating by personation using computer resources, directly targeting deepfakes, voice clones, and automated lures with up to three years' imprisonment and ₹1 lakh fine. Section 66 is about fraudulent computer access, whereas Section 43 penalizes for accessing computer resources without authorization or causing damage to information. Section 43A mandates reasonable security for sensitive data, holding body corporates liable for breaches enabling phishing.

Discretion and Gaps in Application

Prosecutorial discretion arises in proving mens rea for AI tools, as Section 66D requires intent to cheat, complicated by algorithmic opacity in deepfakes. Gaps persist in attributing liability for generative AI outputs, with no explicit regulation of AI models themselves; intermediaries claim safe harbour under Section 79 if diligent. Bailable offences under Section 77B hinder deterrence, and pre-AI provisions struggle with novel techniques like real-time adaptive phishing. Judicial interpretation remains evolving, lacking precedents for voice-cloned vishing or scalable AI harvesting.

Ancillary Regulatory Intersections

Data Protection Under DPDP Act, 2023

The Digital Personal Data Protection Act imposes ₹250 crore penalties for inadequate safeguards against breaches fueling phishing, requiring breach notifications without materiality thresholds. It intersects cybersecurity by legitimizing data use for fraud prevention, addressing phishing's data dependency amid 175% incident growth in H1 2024.

Digital Evidence Framework

Bharatiya Sakshya Adhiniyam recognizes AI-generated records (logs, deepfake forensics) as primary evidence, mandating certification under Section 65B equivalent for admissibility. Challenges include verifying AI evidence reliability against manipulation risks, impacting prosecutions.

Chapter 4

Gaps and Deficiencies in the Existing Regime

The Indian legal architecture for cybersecurity, predominantly enshrined in the Information Technology Act, 2000 (hereinafter "IT Act"), as amended, alongside ancillary statutes such as the Digital Personal Data Protection Act, 2023 (DPDP Act) and the Bharatiya Sakshya Adhiniyam, 2023 (BSA)[19], manifests critical deficiencies in addressing AI-accelerated phishing threats. These lacunae not only erode prosecutorial efficacy but also engender systemic vulnerabilities, permitting perpetrators to exploit technological asymmetries with impunity. This chapter delineates definitional ambiguities, liability attribution conundrums, evidentiary infirmities, procedural hurdles, and regulatory disjunctions, underscoring the imperative for comprehensive legislative recalibration to fortify the cybersecurity edifice against generative AI's pernicious evolution.

Definitional Gaps

Absence of Explicit AI Definitions

The IT Act's definitional scaffold, articulated under Section 2, omits any reference to "artificial intelligence," "generative adversarial networks," "deep learning algorithms," or cognate constructs, engendering interpretive void when invoking provisions like Section 66D, which penalizes "cheating by personation by using computer resource." This lacuna impedes the classification of AI-orchestrated deepfakes or voice-cloned vishing as cognizable offences, as judicial fora grapple with whether such outputs constitute "electronic signatures" or "digital signatures" under Section 2(1)(t). Absent statutory delineation, administrative bodies like CERT-In under Section 70B encounter arbitrariness in mandating compliance for AI-driven intermediaries, perpetuating regulatory inertia.

Unaddressed AI-Generated Harm

No provision within the IT Act or its schedules explicates "AI-generated harm," a novel taxonomy encompassing probabilistic deception intrinsic to large language models (LLMs)[20] deployed in automated spear-phishing. Unlike traditional mens rea-driven offences under Section 66 (computer-related offences), AI inflicts harm via emergent behaviours such as hallucinated yet plausible lures—eluding the rubric of "dishonest intention" or "knowing unauthorized access." This doctrinal shortfall precludes recognition of scalable credential harvesting as a distinct "cyber intrusion," distinct from rudimentary phishing, thereby diluting deterrence for actors leveraging open-source models like those proliferating post-2024.

Inadequacy for Dynamic Cyber Threats

The statutory lexicon, frozen in pre-generative AI vernacular, fails to encapsulate metamorphic threats including real-time adaptive phishing, polymorphic malware symbiosis with AI, or zero-day exploits amplified by reinforcement learning. Sections 43 (penalties for damage to computer systems) and 66, predicated on static "unauthorized access," falter against AI's capacity for evasion via natural language obfuscation or evolutionary attack vectors. This temporal misalignment exacerbates enforcement vacuums, as evidenced by the 1,265% phishing surge documented in 2024-2025 threat reports, outstripping legislative agility.

[19] *Bharatiya Sakshya Adhiniyam, 2023*. (2023, December 25). https://www.indiacode.nic.in/handle/123456789/20063

[20] *Introduction to large language models*. (n.d.). Google for Developers. https://developers.google.com/machine-learning/crash-course/llm

Liability Gaps

Attribution Challenges for AI Agents

Attributing criminal culpability to autonomous AI agents constitutes a juridical nadir; neither the IT Act nor Indian Penal Code (IPC) analogues afford agency to algorithms, confounding Section 66D's mens rea threshold of "intention to cheat." The "black-box" conundrum—wherein neural network opacity obscures causal chains—thwarts forensic reconstruction of intent, raising philosophical quandaries akin to corporate criminal liability sans human analog. Precedents under Section 79(3)(b) safe harbour for intermediaries further insulate AI deployers, absent actual knowledge of misuse.

Vicarious Versus Direct Liability

While Section 43A vicariously fastens negligence-based liability on body corporates for "sensitive personal data" breaches precipitating phishing, direct accountability for AI progenitors (developers, trainers) remains nebulous. No statutory mechanism imputes strict or product liability for foreseeable misuse, such as fine-tuning LLMs on phishing datasets; Section 79(1) due diligence exemptions persist for "passive hosts," bifurcating culpability between upstream creators and downstream exploiters. This asymmetry incentivizes jurisdictional arbitrage via offshore AI services.

Inadequacy of Penalties

Cognizable yet bailable penalties under Sections 66C (identity theft: three years' imprisonment, ₹1 lakh fine) and 66D mirror colonial-era proportionality, dwarfed by AI phishing's macroeconomic depredations—averaging $4.88 million per breach in 2025 metrics. Absent graduated sanctions calibrated to harm quantum (e.g., restitution mandates, asset forfeiture, or enterprise-wide penalties), deterrence erodes; non-compoundable offences under Section 77B notwithstanding, empirical recidivism underscores punitive feebleness against state-sponsored or ransomware-adjacent actors.

Evidentiary and Procedural Gaps

Admissibility of AI-Generated Data

The BSA's Section 63 paradigm admits "electronic records" as primary evidence, yet AI-synthetic artifacts (deepfake videos, cloned audio spectrograms) imperil authenticity absent bespoke certification protocols supplanting Section 65B Information Technology (Miller test equivalents). Judicial skepticism vis-à-vis watermarking or blockchain hashes—prone to adversarial tampering—precipitates exclusionary pitfalls, as chain-of-custody fractures under generative volatility.

Digital Forensics Challenges

Forensic apparatuses lag in countering AI manipulations, with spectral analysis yielding inconclusive provenance for voice clones and diffusion models confounding reverse-engineering. No mandated interoperability standards for tools like NIST-compliant AI detectors exist under CERT-In directives, hampering Section 70A examiners; resource asymmetries in district judiciary exacerbate this, prolonging trials amid evidentiary spoliation risks.

Cross-Border Concerns

Section 75's extraterritorial ambit hinges on "affecting users in India," yet procedural chokepoints via Mutual Legal Assistance Treaties (MLATs) bottleneck 70% of phishing originating from non-signatory jurisdictions (e.g., CIS states). India's non-ratification of the Budapest Convention stymies real-time data-sharing compacts, while Budapest Protocol alternatives falter against anonymization via Tor, VPNs, or decentralized proxies, engendering impunity gradients.

Regulatory Coordination Gaps

Misalignment Between IT Act and Intermediary Guidelines

The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, enjoin "proactive monitoring" for unlawful content, yet clash with DPDP Act's Article 17 privacy bulwarks curtailing preemptive scans. This normative dissonance paralyzes platforms against AI-phishing vectors, as "traceability" mandates under Rule 4(2) conflict with end-to-end encryption norms.

Data Protection Regime Integration

DPDP Act's breach notification timelines (72 hours) elide cybersecurity exigencies under IT Act Section 70B, omitting "significant risk" thresholds tailored to phishing cascades. CERT-In's reporting silos perpetuate fragmented incident response, sans unified dashboards integrating NPCIs financial fraud data with MeitY oversight.

Harmonization with International Standards

India's sui generis framework diverges from EU AI Act's prohibitory-risk taxonomy (Annex III phishing adjuncts), U.S. Executive Order 14110's red-teaming imperatives, or OECD AI Principles' robustness mandates. Absent high-risk AI ex ante assessments or global interoperability via GPAI forums, domestic regimes lag in extraterritorial enforcement, ceding normative primacy to agile multilateralism.

Chapter 5

Comparative Perspectives: Regulation of AI and Phishing in Select Jurisdictions

In an epoch wherein artificial intelligence (AI) catalyzes a paradigm shift in phishing modalities—from rudimentary lures to hyper-personalized deepfake orchestrations—comparative jurisprudence furnishes invaluable precepts for fortifying India's cybersecurity bastion. This chapter dissects regulatory tapestries in the United States (US), European Union (EU), United Kingdom (UK), and Singapore, elucidating civil-criminal liability matrices, service provider obligations, and AI governance architectures germane to phishing suppression. By extrapolating sui generis innovations and pitfalls, it distills actionable insights for Indian lawmaking, advocating adaptive assimilation to preempt AI-driven depredations whilst eschewing overreach.

United States: Civil-Criminal Liability and Service Provider Imperatives

Civil Liability Frameworks

The US paradigm amalgamates sectorally calibrated civil recourse with the Federal Trade Commission (FTC) Act's Section 5 proscription of "unfair or deceptive acts," imposing injunctive relief and restitutive penalties upon entities remiss in AI phishing safeguards. California's Consumer Privacy Act (CCPA/CPRA) mandates data minimization and breach notifications, engendering class actions for phishing-exploited breaches, with fines up to $7,500 per intentional violation. The Securities and Exchange Commission (SEC) cyber disclosure rules (Item 1.05, Form 8-K) compel material AI-enabled incident disclosures within four days, mitigating systemic risks.

Criminal Liability for Cybercrime

The Computer Fraud and Abuse Act (CFAA), 18 U.S.C. § 1030, criminalizes unauthorized access abetting phishing, with AI Fraud Deterrence Act (2025) escalating penalties: wire/mail fraud fines to $1-2 million and 20-30 years' incarceration for AI-assisted impersonation; government official deepfakes attract $1 million fines and three years' imprisonment. The No AI FRAUD Act (proposed) targets voice/video synthesis sans consent, harmonizing with state deepfake statutes.

Service Provider Responsibilities and AI Governance

Service providers shoulder due diligence under NIST AI Risk Management Framework (RMF) and Executive Order 14110 (2023), mandating adversarial robustness testing against phishing vectors. FTC enforces vendor liability for AI supply chain vulnerabilities, whilst CISA's Shields Up initiative coordinates incident response.

European Union: AI Act and GDPR-Cybersecurity Synergy

EU AI Act: Risk-Tiered Prohibitions

The Artificial Intelligence Act (Regulation (EU) 2024/1689), effective August 2024, stratifies AI per risk: "prohibited" (e.g., real-time biometric phishing) incurs €35 million fines (7% global turnover); "high-risk" systems (e.g., cybersecurity AI) necessitate conformity assessments, transparency on generative outputs, and cyber resilience under Article 55 for GPAI models against model theft/phishing circumvention. Phishing-adjacent "social scoring" or manipulative subliminals face outright bans.

GDPR Cohesion with Cybersecurity Mandates

GDPR (Article 32) mandates "appropriate technical measures" against phishing breaches, synergizing with AI Act's data protection carve-outs (Recital 114); Data Act (2023) facilitates data sharing for cybersecurity. NIS2 Directive (2022) imposes C-level accountability for critical infrastructure, with €10 million fines for non-compliance, enabling cross-phishing incident reporting.

Cross-Border Enforcement

One-stop-shops and mutual recognition under AI Act Article 64 streamline enforcement, with EDPB coordination bridging 27 Member States; adequacy decisions extend extraterritorial bite to non-EU AI deployers targeting EU data subjects.

United Kingdom and Singapore: Pragmatic Statutory-Guideline Hybrids

United Kingdom: Online Safety Act and Algorithmic Duties

The Online Safety Act 2023 (OSA) vests Ofcom with imposing risk assessments on platforms for illegal harms (Section 9), including AI-phishing frauds; "systemic risk" duties (Section 14) compel algorithmic audits to mitigate exposure, with Category 1 platforms facing £18 million fines (10% turnover). AI super-tools for fraud detection (e.g., site analysis) integrate proactively, whilst OSA's child safety duties analogize to adult phishing protections.

Singapore: Cybersecurity Act and Computer Misuse Regime

Singapore's Cybersecurity Act 2018 designates Critical Information Infrastructure (CII) for mandatory audits, with CSA's Model AI Governance Framework (updated 2024) exhorting "explainability" and bias mitigation for phishing-AI. Computer Misuse Act (CMA)[21] Section 3 penalizes unauthorized access/phishing (S$5,000 fine/two years' imprisonment first offence), explicitly encompassing GenAI content (13% phishing emails in 2023). CSA observes WormGPT-like uncensored LLMs, mandating incident reporting within two hours for ransomware/phishing.

[21] Wikipedia contributors. (2005, April 15). *Computer Misuse Act 1990*. Wikipedia. https://en.wikipedia.org/wiki/Computer_Misuse_Act_1990

Chapter 6

Policy Recommendations and Proposed Legal Reforms

The inexorable proliferation of AI-orchestrated phishing imperatives a metamorphic reconfiguration of India's cybersecurity jurisprudence, transcending ameliorative palliatatives to engender a resilient, forward-looking statutory edifice. Anchored in antecedent analyses of technological vicissitudes and comparative paradigms, this chapter proffers a compendium of calibrated policy prescriptions and legislative emendations. These encompass definitional perspicuity, offence novation, evidentiary fortification, regulatory synchrony, intermediary recalibration, incident response augmentation, governance prophylactics, and multilateral consonance—collectively calibrated to preempt AI's pernicious phishing dynamics whilst nurturing innovation equilibria.

Definitional Clarity in Cybercrime Provisions

The Information Technology Act, 2000 (IT Act) warrants amendment via insertion of Section 2(1)(zaa): "'Artificial Intelligence' or 'AI' denotes systems manifesting adaptive intelligence, encompassing machine learning, generative models, and autonomous agents capable of synthesizing content or decisions sans explicit programming." This explicit articulation obviates interpretive interstices under Sections 66C-D, facilitating prosecutorial precision and mitigating judicial casuistry. Analogously, "AI-Generated Content" encompasses deepfakes, voice cloning, and phishing scams using AI-generated content, aligning with the EU AI Act definition.

Further, novate Section 2(1)(zab): "'Autonomous Agent' connotes AI entities executing actions independent of contemporaneous human oversight, imputable to deployers for resultant harms," thereby bridging agency voids in attribution conundrums prevalent in black-box AI deployments. Such perspicuity not only aligns domestic provisions with global standards but also empowers administrative bodies like CERT-In under Section 70B to mandate compliance without arbitrariness, ensuring technological neutrality in enforcement.

AI-Specific Offence Structures and Enhancements

To delineate AI-enabled phishing variants, enact Section 66DA: "Punishment for AI-Enabled Phishing.Any person who employs AI for spreading misinformation to extract sensitive data or cause monetary harm will be sentenced to up to seven years in prison and a fine of ₹10 lakhs."

Tiered sub-clauses calibrate deterrence proportionally:

(i) spear-phishing attracts five years;

(ii) deepfake/vishing seven years;

(iii) scalable harvesting ten years for Critical Information Infrastructure (CII) impacts. This graduated structure supplants the IT Act's anachronistic uniformity, emulating the US AI Fraud Deterrence Act's escalatory penalties.

Augment Section 66D with a proviso escalating baseline penalties by 200% upon AI utilization, incorporating identity spoofing (Section 66C analogue) and data exfiltration. Introduce a rebuttable mens rea presumption for deployers of uncensored large language models (LLMs) evincing foreseeable phishing utility, thereby addressing probabilistic deception intrinsic to generative outputs. These novations engender robust deterrence, calibrated to AI phishing's macroeconomic depredations, such as breaches averaging $4.88 million.

Strengthening Evidence and Forensics Regime

The Bharatiya Sakshya Adhiniyam (BSA) requires amendment via insertion of Section 63A: "Admissibility of AI-Generated Records.Synthetic media admissible upon certification evincing provenance via cryptographic hashes, watermarks, or NIST-equivalent forensics, rebuttably presuming authenticity." This establishes national standards, mandating a MeitY-overseen AI forensics laboratory to counter deepfake volatility and spectral manipulations in voice

clones. Judicial skepticism vis-à-vis adversarial tampering is thus allayed through Daubert-like reliability thresholds.

Promulgate the IT (Forensic Standards) Rules, 2026, prescribing immutable blockchain-ledgering for digital traces from phishing incidents, alongside chain-of-custody protocols. Expert testimony guidelines under CrPC Section 293 qualify "AI forensics specialists," ensuring interoperability with tools like spectral analysis, thereby fortifying evidentiary integrity against spoliation risks and resource asymmetries in district judiciary.

Alignment with Data Protection Trajectories

Synchronize IT Act and DPDP Act reforms through a DPDP proviso (Article 17 bis) authorizing cybersecurity carve-outs for phishing mitigation sans granular consent, mandating pseudonymization of processed data. Elevate Section 43A to strict liability for AI breaches, with unified 24-hour notifications to CERT-In and NPCI, addressing consent fatigue in AI contexts via opt-in granularities akin to GDPR Article 32. Joint MeitY-NPC regulations harmonize these trajectories, preempting normative dissonance in data-dependent phishing cascades.

Sharpening Intermediary Protections and Obligations

Revise Information Technology (Intermediary Guidelines) Rules, 2021, Rule 4: Safe harbour under Section 79 contingent upon AI risk assessments and proactive moderation, e.g., 80% takedown within one hour for flagged phishing. Incentives include tax rebates for verified AI safety certifications, recalibrating protections against abuse whilst fostering diligence. Ofcom-UK inspired "duty of care" for Category I platforms imposes graduated penalties up to ₹50 crore, equilibrating liability with innovation imperatives.

Enhancing Incident Reporting and CERT-In Capabilities

Insert Section 70B(4A): Compel AI-assisted disclosures within two hours for phishing cascades affecting over 1,000 users, establishing a national AI-CERT fusion cell for machine learning-driven threat intelligence sharing. Coordinated response mechanisms integrate I4C, RBI, and SEBI via real-time dashboards, emulating Singapore's CSA model to mitigate fragmented silos and enhance systemic resilience.

Promoting Governance and Research Safeguards

The IT (AI Governance) Rules, 2026, mandate pre-deployment impact assessments for high-risk sectors (BFSI, CII) per EU AI Act Article 9, auditing transparency, bias mitigation, and adversarial robustness. Annual third-party audits impose C-suite liability akin to NIS2 Directive, buttressed by whistleblower protections, ensuring accountable deployment sans stifling research ecosystems.

Encouraging International Cooperation

Ratify the Budapest Convention and forge bilateral adequacy pacts with EU/US for MLAT acceleration. Integrate OECD AI Principles into MeitY advisories, with GPAI participation facilitating phishing intel swaps and cross-border enforcement. These measures harmonize standards with global best practices, transcending India's extraterritorial ambit under Section 75.

These reforms, sequenced via an omnibus IT Amendment Bill, 2026, proffer a prophylactic bulwark, equilibrating deterrence with innovation whilst remedying the extant regime's infirmities.

Chapter 7

Conclusion

The foregoing disquisition has meticulously excavated the profound deficiencies inhering within the Information Technology Act, 2000 (IT Act), as amended, when juxtaposed against the inexorable escalation of artificial intelligence (AI)-facilitated phishing threats from 2024 through 2025 and into the present. Definitional lacunae—manifesting in the absence of explicit delineations for "artificial intelligence," "generative AI," "autonomous agents," and "AI-generated content"—engender interpretive ambiguities that vitiate the invocability of core provisions such as Sections 66C (identity theft), 66D (cheating by personation), and 66 (computer-related offences). These interstices preclude the precise classification of sophisticated modalities including automated spear-phishing, voice-cloned vishing, deepfake-orchestrated deceptions, and scalable credential harvesting, thereby permitting perpetrators to exploit statutory silences with calculated impunity. Liability attribution conundrums further exacerbate this frailty: the black-box opacity of neural architectures obfuscates mens rea proofs, bifurcates vicarious accountability under Section 43A from direct developer culpability, and renders extant penalties capped at three years' imprisonment and nominal fines woefully disproportionate to the macroeconomic depredations of AI phishing, which exacted breaches averaging $4.88 million per incident in contemporaneous metrics. Evidentiary infirmities under the Bharatiya Sakshya Adhiniyam, 2023 (BSA), compound procedural hurdles, as AI-synthetic artifacts elude Section 63 admissibility sans bespoke certification, whilst cross-border chokepoints via dilatory Mutual Legal Assistance Treaties (MLATs) frustrate Section 75's extraterritorial ambit. Regulatory disjunctions between the IT Act, Intermediary Guidelines Rules, 2021, and the Digital Personal Data Protection Act, 2023 (DPDP Act), perpetuate fragmented enforcement, unmoored from international benchmarks such as the EU AI Act's risk-tiering or the US Computer Fraud and Abuse Act's escalatory sanctions.

This panoramic critique underscores the imperious necessity for a proactive, adaptive legal framework, architectonically engineered to outpace AI-enabled cyber threats whilst scrupulously safeguarding fundamental rights enshrined in Articles 14, 19, and 21 of the Constitution of India, alongside nurturing the nation's burgeoning AI innovation ecosystem. The extant regime's reactive, pre-generative vernacular frozen in a pre-ChatGPT epoch falters against dynamic adversaries wielding reinforcement learning for polymorphic attacks, necessitating a metamorphic jurisprudence that anticipates technological teleologies rather than merely redressing sequelae. Such adaptation demands not mere incremental emendations but a holistic recalibration: tiered offence novations with presumptive mens rea for uncensored large language models; evidentiary bulwarks via cryptographic provenance standards and national AI forensics laboratories; synchronized DPDP-IT Act protocols for breach notifications; recalibrated Section 79 safe harbours contingent on proactive moderation duties; augmented CERT-In imperatives for real-time threat intelligence fusion; mandatory impact assessments for high-risk deployments in critical information infrastructure; and multilateral consonance through Budapest Convention ratification. This forward-looking edifice, if operationalized via an omnibus IT Amendment Bill, 2026, would equilibrate deterrence with due process, transmuting vulnerabilities into prophylactic resilience without imperiling privacy sanctuaries or entrepreneurial verve.

At its crux, the central implication resonates with unyielding clarity: robust, clarifying reforms constitute not merely desirable prophylactics but existential imperatives to deter AI-driven phishing and to bolster India's cybersecurity governance apparatus in 2024–2025 and beyond. Absent such legislative alacrity, the Digital India imprimatur risks subversion by adversarial AI, with cascading erosions to economic sovereignty, public trust, and national security. Conversely, emulating pragmatic hybrids from comparator jurisdictions the EU's prohibitory-risk taxonomy, Singapore's regulator-agile audits, the UK's duty-of-care mandates, and the US's graduated criminality positions India as a vanguard in AI governance, harmonizing technological sovereignty with global interoperability. Judicial capacity augmentation, public-private sandboxes, and ethical AI advisories under MeitY aegis would further entrench this trajectory, ensuring that cybersecurity jurisprudence evolves conterminously with AI's dual-use ontology. In sum, these reforms proffer a bulwark against phishing's generative apocalypse, fortifying the rule of law in an era where silicon sentinels guard democratic ramparts.