INDIAN JOURNAL OF LEGAL REVIEW [IJLR – IF SCORE – 7.58]

VOLUME 5 AND ISSUE 11 OF 2025

APIS – 3920 - 0001 *(and)* ISSN - 2583-2344

Published by

Institute of Legal Education

https://iledu.in

BLACK BOX, RED FLAGS: NAVIGATING LEGAL FAULT LINES IN THE AGE OF AI AND CYBERSECURITY

AUTHOR – VIKAS KABEER, LLM SCHOLAR AT MVN UNIVERSITY, PALWAL

BEST CITATION – VIKAS KABEER, BLACK BOX, RED FLAGS: NAVIGATING LEGAL FAULT LINES IN THE AGE OF AI AND CYBERSECURITY, *INDIAN JOURNAL OF LEGAL REVIEW (IJLR),* 5 (11) OF 2025, PG. 184-188, APIS – 3920 – 0001 & ISSN - 2583-2344

Abstract

Introduction of Artificial Intelligence (AI) in cybersecurity systems is a revolutionary advantage in the areas of threat detection, fraud prevention, and response to a given incident. Nonetheless, it has also brought with it dicey ethical and legal issues. Malicious individuals grow to use AI to carry out complex cyberattacks, including deep fake frauds, machine-code malware, and data poisoning attacks, bringing new liability, privacy, governance, and jurisdiction concerns. This article describes the legal risks of AI-based cybersecurity threats and discusses the emergent changes in regulations in different major jurisdictions, such as the European Union, the United States, and India. It discusses key legal frameworks including EU Artificial Intelligence Act, NIS2 Directive, DPDP Act of India or U.S Executive Orders of AI. Internal governance processes that an agency should implement, as detailed in the article, are the board level oversights, adversarial testing, and incident reporting ones. Lastly, it pinpoints the need to address the emerging legal confusion like autonomous offensive AI and quantum-enabled cryptanalysis. The article gives practical suggestions to the legal practitioners, policy- makers, and business executives involved in a cross-sectoral and comparative approach to the problem of AI and its laws of cybersecurity.

Keywords: Cybersecurity, AI Regulation, Legal Liability, Offensive AI, Quantum Security

I. Introduction

The emergence of Artificial Intelligence (AI)/cyber security is a revolutionary era of regulation in the digital domain. AI is a two-edged weapon, it may become a great savior of the cyber world and also used as a weapon of the cyberattack as well. These two points of intersection bring forth legal and ethical challenges that are unprecedented in nature, such as how to assign liability to algorithm-related booboos or how to formulate international standards concerning the use of AI-enhanced cyber warfare. ChatGPT creates the pressing need to keep up with the growing implementation of AI in cyber-defense infrastructure, finance, critical infrastructure, and biometrics authentication that requires not only technical regulation but also effective legal governance.[367]

II. AI-Compelled Cyber Threats: A Budding Venture of Legal Hazard

Its introduction into the cyber systems has created new classes of risks that have complicated current legal solutions. Artificial intelligence-based social engineering is one of them. Such as, generative AI has been used in deepfake technology to impersonate voices and video manifestations to fool employees into sending huge amounts of money in situations they were tricked into. An interesting example was the one of the British engineering company Arup in which deepfake voice that sounded like

[367] Nat'l Inst. of Standards & Tech. (NIST), AI Risk Management Framework, Version 1.0 (Jan. 2023), https://www.nist.gov/itl/ai-risk-management-framework.

a senior executive resulted in a 20-million-pound fraud.[368]

Automatic development of malware has become possible through large language models (LLM); this makes it possible to generate polymorphic code the adversaries can use to bypass detection mechanisms. Moreover, security systems can become vulnerable to model poisoning where the compromised data is injected into the AI training dataset prompting the system to retrospectively complete malware as legitimate trafficking.[369] Privacy risks are also caused by data leakage via model inversion in which the attackers can mine sensitive training data in AI models.[370]

These powers push the legal structures to excess. Accountability emerges when AI algorithms decide or bring situational actions. The size and complexity of threats caused by AI demand a renewed concept of responsibility, proof, and diligence.[371]

III. The Law and Adverse Repercussions

The applicable issues legal to the utilization of AI in cybersecurity are the application of product liability, compliance of data retention, the bias associated with algorithms, and transnational enforcement.

A subject of critical importance is product liability in case of failure in an AI-based security system. According to the European Union, Artificial Intelligence Act, high-risk AI systems should comply with special safety and resilience requirements. Under Article 15, there are forced risk-management and security measures (European Commission, 2024a). Companies that offer non-compliant AI can be fined between 7 percent of the total annual global turnover and 35 million Euros.[372] Negligence claims might come up in the common-law jurisdiction in case no sufficient AI defenses were maintained.

The cybersecurity obligations are tight on the data processors and controllers on laws, particularly in the EU General Data Protection Regulation (GDPR), and India Digital Personal Data Protection Act, 2023. Article 32 of the GDPR demands organizations to adopt technical and organizational protection measures and the requirement of breach notification comes with preferable security measures.[373] [374]

Bias and discrimination in AI occur in cases like when AI tools discriminate or profile or unfairly target the users. Financial systems on its part, AI-powered credit scoring or fraud detection can lead to discrimination of zip, income or gender. The U.S law tackles such risks in the form of the Equal Credit Opportunity Act whereas the Indian regulators like the RBI have also issued warning against the opaque credit-scoring models.[375]

Jurisdiction across borders is a tricky issue to hang on, particularly when attacks are to be made through the AI that is found in another country as servers. The Budapest Convention on Cybercrime and Mutual Legal Assistance Treaties provide options of cross-border enforcement although their applicability is weak because there is no harmony in their adoption.[376]

IV. Global Legal and regulatory developments

Adopted in August 2024, the EU Artificial Intelligence Act will be the first legislation in the world that is dedicated to AI. Its categories AI systems according to risk levels and renders lifecycle obligations to providers of systems that

[368] Julia Kollewe, UK Engineering Firm Arup Falls Victim to £20m Deepfake Scam, *The Guardian* (May 17, 2024), https://www.theguardian.com.
[369] E.U. Agency for Cybersecurity (ENISA), Threat Landscape Report 2024 (2024), https://www.enisa.europa.eu/publications/enisa-threat-landscape-2024.
[370] Int'l Org. for Standardization, ISO/IEC 42001:2023, Artificial Intelligence Management Systems (2023).
[371] Ctr. for Sec. & Emerging Tech., Breaking Down the Biden AI Executive Order (2024), https://cset.georgetown.edu/publication/breaking-down-the-biden-ai-eo.
[372] The Nat'l L. Rev., The EU AI Act: Sanctions and Liability, Mar. 2024, https://www.natlawreview.com/article/eu-ai-act-sanctions-and-liability.
[373] Digital Personal Data Protection Act, No. 22 of 2023, INDIA CODE, https://indiacode.nic.in.
[374] Regulation (EU) 2016/679, of the European Parliament and of the Council of 27 Apr. 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data (General Data Protection Regulation), 2016 O.J. (L 119) 1.
[375] Securities & Exch. Bd. of India, Consultation Paper on the Use of Artificial Intelligence and Machine Learning in the Securities Market (May 2025), https://www.sebi.gov.in.
[376] Convention on Cybercrime, Nov. 23, 2001, Europ. T.S. No. 185 (Budapest Convention).

present high risks to lifecycle. Serious incidents must be reported pursuant to Article 26 and due to Article 56, industry codes of conduct may be established.[377]

The NIS2 Directive, which comes into force on October 2024, revises the system of guaranteeing cybersecurity of key and important entities, including digital infrastructure as well as AI service providers. It introduces an obligation to notify about the incident within 24 hours, the implementation of the supply-chain security and board governance.[378]

In the U.S., one related to AI safety and resilience as was in the newly released Executive Order 14110 (October 2023). It orders CBPP and NIST to create cybersecurity requirements and promotes AI implementation in the protection of federal infrastructure (White House, 2023; DHS, 2024). Such activities are done within the framework of NIST AI Risk Management, which provides a way to mitigate risk management through a four-functions scheme: Map, Measure, Manage, and Govern.[379]

India is in the process of coming up with its own Digital India Act that will be replacing the Information Technology act, 2000. Even in the draft form, it is projected to deal with AI-generated deepfakes, cybersecurity duties and algorithmics. Simultaneously, CERT-In Advisory and the Master Directions of the RBI on non-bank financial entities already have timelines on incident reporting and requirements of governance of cybersecurity frameworks.[380]

V. Mechanisms of Governance at Organizational Level

Companies will have to combine governance systems that comply with the regulations of both cybersecurity and AI. On the leadership level, directors are supposed to take responsibility to ensure that AI is deployed under their monitoring.[381]

One should implement an AI-Security Development Life Cycle (SDLC). This involves carrying out adverse testing, threat analysis, and design principles of secure by design following the ISO/IEC 42001:2023.[382] Organizations are also expected to match their incident response strategies to the risks associated with AI, when they also need to comply with various laws such as the GDPR, DPDP Act, and the AI Act.

High-risk AI systems will need verification by third party auditors and conformity checks. Article 43 of the EU AI Act establishes an obligation in regard to third-party assessment. The performance, security, and transparency of AI is also promoted through independent validation of performance, security, and transparency in voluntary framework, including that of NIST RMF.[383]

Sectoral Focus

Financial In the financial sector, AI has prevalently been applied in fraud detection, e-KYC (Know Your Customer), and automatic lending. Financial institutions need to have Chief Information Security Officers, IT audits should be executed at least four times a year, and a board-level responsibility should be in place. The 2025 consultation paper by SEBI suggests standards of transparency, fairness, and explainability with respect to the use of AI models in the process of securities trading.[384]

Use of AI in cybersecurity is gaining popularity even in critical infrastructure operators which will pose threats in case of intrusion by an

[377] Regulation (EU) 2024/865, of the European Parliament and of the Council of 1 Aug. 2024 on Artificial Intelligence (AI Act), 2024 O.J. (L 865) 1.
[378] Int'l Bar Ass'n, The NIS2 Directive: A New Era of Cybersecurity Regulation (2024), https://www.ibanet.org.
[379] Nat'l Inst. of Standards & Tech. (NIST), AI Risk Management Framework, Version 1.0 (Jan. 2023), https://www.nist.gov/itl/ai-risk-management-framework.
[380] CERT-In Directions, Ministry of Electronics & Info. Tech., Gov't of India, Directions on Information Security Practices, Procedures, Prevention, Response, and Reporting of Cyber Incidents (Apr. 28, 2022), https://www.cert-in.org.in.
[381] Reserve Bank of India, Master Direction – Information Technology Framework for the NBFC Sector (Jan. 2024), https://www.rbi.org.in.
[382] Int'l Org. for Standardization, ISO/IEC 42001:2023, Artificial Intelligence Management Systems (2023).
[383] Nat'l Inst. of Standards & Tech. (NIST), AI Risk Management Framework, Version 1.0 (Jan. 2023), https://www.nist.gov/itl/ai-risk-management-framework.
[384] Securities & Exch. Bd. of India, Consultation Paper on the Use of Artificial Intelligence and Machine Learning in the Securities Market (May 2025), https://www.sebi.gov.in.

adversary. The 2025 directive of the Biden administration imposes on the U.S. Department of Energy the establishment of AI red-team environments and securing cloud clusters utilized in the defense of critical infrastructure.[385]

Content providers and social media platforms have a challenge to deal with synthetic media which is gaining the potential. Component 7: Transparency must be given to AI systems that produce such materials as images, videos, or texts through AI systems (Article 52 of the AI Act). The UK is also setting new polices of prosecuting the destructive creation of deepfakes, particularly in non-consensual sexual images.[386]

VI. Trends and New Issues in Enforcement

New enforcement trends have shown a tendency of moving towards individual accountability and real time compliance. The NIS2 Directive issued by the EU permits the national regulators to even suspend or impose fines on the executives who do not adhere to the cybersecurity regulations.[387] SEC requires companies to report on their cyber governance process, which involves AI risk management, in the listed companies in the United States. The DPDP Act of India allows the fine between 15 crore Indian rupees and 250 crore Indian rupees for the breach of data protection laws.[388]

The risks of AI are rapidly expanding as AI expands as well. International law has a problem with concerns about autonomous offensive AI where computers are allowed to make their choices to conduct a cyber action independently. Experts can find pieces of advice on cyber warfare in the Tallinn Manual 3.0, yet it cannot be applied to the setting when AI acts as a sovereign.[389]

Quantum computing is another thing creeping at the horizon. Soon compilers will find that AI-optimized quantum algorithms will break current cybersecurity measures, prompting the change to post-quantum encryption standards. The governments can initiate obligatory quantum resilience testing as a certification procedure of AI relative to cybersecurity.[390]

VII. Recommendations

Considering these challenges, organisations have to combine the legal compliance with internal risk governance. The boards are to order AI risk registers and to implement monitoring protocols to comply with NIS2 Directive or RBI rules. Technical teams are advised to conform to the ISO and NIST requirements regarding adversarial testing and secure development. Regulatory civilized to ensure that legal terms used in AI laws, cybersecurity, and data protection laws are cross-referenced and lead to greater harmonization of legal terms in these laws and encourage collaboration between people and privately owned organizations to address the menace of deepfakes and phishing through the use of AI.

VIII. Conclusion

AI is ushering in a new age of cybersecurity as well as making it stronger and more complex. The related problems of legal issues, which vary form product liability to cross state enforcement need to be addressed in multidimensional governance. With frameworks being developed, including the EU AI Act, the NIS2 Directive and the Digital India Act, organizations must match their technical operations with legal requirements as they develop. With the establishment of an AI-domain of transparency, accountability, and moral design, the possibilities of using AI to enhance cybersecurity will not threaten the individual or jeopardize national security.

[385] Exec. Order No. 14,110, Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, 88 Fed. Reg. 75,201 (Oct. 30, 2023).
[386] Julia Kollewe, UK Engineering Firm Arup Falls Victim to £20m Deepfake Scam, *The Guardian* (May 17, 2024), https://www.theguardian.com.
[387] Int'l Bar Ass'n, The NIS2 Directive: A New Era of Cybersecurity Regulation (2024), https://www.ibanet.org.
[388] Digital Personal Data Protection Act, No. 22 of 2023, INDIA CODE, https://indiacode.nic.in.
[389] Michael N. Schmitt et al., *Tallinn Manual 3.0 on the International Law Applicable to Cyber Operations* (NATO Cooperative Cyber Def. Ctr. of Excellence 2023).
[390] Ctr. for Sec. & Emerging Tech., Breaking Down the Biden AI Executive Order (2024), https://cset.georgetown.edu/publication/breaking-down-the-biden-ai-eo.

References:

1. CERT-In Directions, Ministry of Electronics & Info. Tech., Gov't of India, Directions on Information Security Practices, Procedures, Prevention, Response, and Reporting of Cyber Incidents (Apr. 28, 2022), https://www.cert-in.org.in.
2. Ctr. for Sec. & Emerging Tech., Breaking Down the Biden AI Executive Order (2024), https://cset.georgetown.edu/publication/breaking-down-the-biden-ai-eo.
3. U.S. Dep't of Homeland Sec., AI Cybersecurity Guidance Under Executive Order 14110 (2024), https://www.dhs.gov/ai.
4. E.U. Agency for Cybersecurity (ENISA), Threat Landscape Report 2024 (2024), https://www.enisa.europa.eu/publications/enisa-threat-landscape-2024.
5. Regulation (EU) 2024/865, of the European Parliament and of the Council of 1 Aug. 2024 on Artificial Intelligence (AI Act), 2024 O.J. (L 865) 1.
6. Directive (EU) 2022/2555, of the European Parliament and of the Council of 14 Dec. 2022 on Measures for a High Common Level of Cybersecurity Across the Union (NIS2 Directive), 2022 O.J. (L 333) 80.
7. Regulation (EU) 2016/679, of the European Parliament and of the Council of 27 Apr. 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data (General Data Protection Regulation), 2016 O.J. (L 119) 1.
8. Digital Personal Data Protection Act, No. 22 of 2023, INDIA CODE, https://indiacode.nic.in.
9. Int'l Bar Ass'n, The NIS2 Directive: A New Era of Cybersecurity Regulation (2024), https://www.ibanet.org.
10. Int'l Org. for Standardization, ISO/IEC 42001:2023, Artificial Intelligence Management Systems (2023).
11. Nat'l Inst. of Standards & Tech. (NIST), AI Risk Management Framework, Version 1.0 (Jan. 2023), https://www.nist.gov/itl/ai-risk-management-framework.
12. Reserve Bank of India, Master Direction – Information Technology Framework for the NBFC Sector (Jan. 2024), https://www.rbi.org.in.
13. Securities & Exch. Bd. of India, Consultation Paper on the Use of Artificial Intelligence and Machine Learning in the Securities Market (May 2025), https://www.sebi.gov.in.
14. Julia Kollewe, UK Engineering Firm Arup Falls Victim to £20m Deepfake Scam, *The Guardian* (May 17, 2024), https://www.theguardian.com.
15. Exec. Order No. 14,110, Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, 88 Fed. Reg. 75,201 (Oct. 30, 2023).
16. Convention on Cybercrime, Nov. 23, 2001, Europ. T.S. No. 185 (Budapest Convention).
17. The Nat'l L. Rev., The EU AI Act: Sanctions and Liability, Mar. 2024, https://www.natlawreview.com/article/eu-ai-act-sanctions-and-liability.
18. Michael N. Schmitt et al., *Tallinn Manual 3.0 on the International Law Applicable to Cyber Operations* (NATO CCDCOE 2023).