JOURNAL OF CYBER GOVERNANCE AND INTELLECTUAL PROPERTY

Volume 1 Article 25

Digital Intermediaries as Frontiers of National Security: The role of AI in Cyber Threat Detection and State Surveillance

Anshika Saxena

CHRIST (Deemed to be University)

Recommended Citation:

Saxena (2026) "Digital Intermediaries as Frontiers of National Security: The role of AI in Cyber Threat Detection and State Surveillance" Journal of Cyber Governance and Intellectual Property, Vol. 1, Article 25. (DOI)

ABSTRACT

The rapid expansion of India's digital ecosystem has ably transformed the status of "digital intermediaries" as critical actors in "national security" governance. This paper is a mere attempt at the examination of this nexus of digital intermediaries (and their AI systems) and India's national security ecosystem. AI, in contemporary times, plays a vital role in cyber threat detection and state surveillance. These digital intermediaries use systems like machine learning for phenomena like anomaly detection and behavioral analytics against the legal and constitutional frameworks of our nation which govern interception, data access and privacy of the juristic identities. Through the doctrinal analysis of statutes like, DPDP Act, 2034 and IT Act, 2000, leading judgements, contemporary research papers and other relevant documents, this paper evaluates the legal gaps which occur at the location of algorithmic accountability and oversight which shall arise when intermediaries deploy totally opaque AI for reasons, such as "security". This analysis tries to establish that while AI might materially improve threat detection or speed, it simultaneously amplifies certain other risks such as automated profiling, rights towards infringement of state surveillance, when used without a clear statutory framework. This paper argues for the absence of a "mother law" governing AI systems and urges the need for a standardized regulatory framework including judicial authorization while high-risk deployments, statutory limits on government exemptions in data laws and obligation-based liability rules for digital intermediaries that must balance national security with fundamental rights.

KEYWORDS

Digital intermediaries, Artificial intelligence, Cybersecurity, Surveillance, Statutory governance

INTRODUCTION

India's transition into a digitally governed society has significantly altered the relationship between technology, governance and national security. There has been a noticeable unprecedented concentration of data within both state institutions and private intermediaries due to widespread adoption of technologies like UPI, DigiLocker, etc. In this evolving digital atmosphere, the digital intermediaries no longer operate as neutral channels of information but as active gatekeepers of national security. Hence, the prior national security threats have evolved from isolated cyber security incidents to sophisticated multi-vector campaigns. These may include supply-chain compromises, AI-assisted phishing, Malicious Deepfake content, Financial frauds, Cyber intrusions, and more.

These digital intermediaries hold in capability to sense and detect, while flagging individuals or content online, blocking content, etc. However, when these capabilities are aided with AI models, the risks and liabilities swell. These risks can be observed easily when India's regulatory landscape with the condition of recent laws like DPDP Act[1], 2023, Telecommunication Act, 2023[2], etc. provide expansive surveillance powers and state exemptions to a large extent, while AI-specific governance remains hushed.

AI has significantly enhanced India's state surveillance and cyber defense capabilities, especially against modern cyber-attacks that conventional rule-based systems might've struggled to address. However, there exists no comprehensive AI legislation governing the development or accountability of AI systems. Hence, the legal framework governing AI-enabled cybersecurity and state surveillance remains deeply inadequate.

Therefore, this paper analyses the role of digital intermediaries as frontline actors of national security, while critically examining AI's role in simultaneously empowering cyber defence and state surveillance in India. It further evaluates the inadequacies of India's existing legislative regime in regards with AI governance and argues for the urgent need of a comprehensive AI governance framework.

LITERATURE REVIEW

[1] Digital Personal Data Protection Act, No. 22 of 2023 (India).
[2] Telecommunications Act, No. 44 of 2023 (India).

The first part of the literature is about the status of digital intermediaries in India. Some time back people who studied this topic thought of intermediaries as a medium to pass on information and their responsibility was limited by certain protections[3] under Section 79 of the Information Technology Act, 2000[4]. The courts also said that intermediaries do not have to remove content just because someone complains about it as seen in the case of Shreya Singhal v. Union of India[5]. However now intermediaries are not just passive, they actually manually check the content. Remove it, if necessary, they use algorithms to find threats and also give data to the government when asked. This change has made scholars think that the law should also change and intermediaries should be more responsible for their actions. There is no clear theory on how to do this and it is a big problem when intermediaries are used for security purposes[6].

The second part is about using Artificial Intelligence in procuring cyber threats[7]. Many studies have shown that Artificial Intelligence is better than methods in finding malware, ransomware and other threats. Artificial Intelligence can find these threats accurately and quickly and it also reduces "false alarms"[8]. But there are also some problems with using Artificial Intelligence like it can make decisions that're not transparent and it can be biased. These problems are not just technical, but legal because Artificial Intelligence is used to make decisions about what to block or what to investigate. Many scholars are worried that Artificial Intelligence is fast. It is not very transparent and this creates a problem between doing things quickly and following the right process especially when big intermediaries use Automated Intelligence tools.

The third part is about Artificial Intelligence and surveillance and privacy. Some scholars have written about how India's using facial recognition, predictive policing and social media monitoring but it does not have a good legal framework to control these things[9]. They say that

---

[3] *An Analysis of Intermediary Liability in India and the European Union*, Manupatra Articles. https://articles.manupatra.com/article-details/AN-ANALYSIS-OF-INTERMEDIARY-LIABILITY-IN-INDIA-AND-THE-EUROPEAN-UNION (last accessed 12 May, 2026).

[4] Information Technology Act, No. 21 of 2000, § 79 (India).

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

[6] *State Surveillance and the Right to Privacy in India*, Manupatra PDF. http://docs.manupatra.in/newsline/articles/Upload/62E569B9-547B-4144-80AC-03A9EBFC45DA.pdf (last accessed 12 May, 2026).

[7] Siva Vignesh S.K.V. & Nagarjun D.N., *Legal Challenges of Artificial Intelligence in India's Cyber Law Framework: Examining Data Privacy and Algorithmic Accountability via a Comparative Global Perspective*, 6(6) IJFMR (2024). https://www.ijfmr.com/research-paper.php?id=31347 (last accessed 12 May, 2026).

[8] *The Surveillance Law Landscape in India and the…*, Global Network Initiative / CCG PDF. https://globalnetworkinitiative.org/wp-content/uploads/2023/07/CCG-June-15.pdf (last accessed 12 May, 2026).

[9] *State Surveillance and the Right to Privacy in India*, Manupatra PDF. http://docs.manupatra.in/newsline/articles/Upload/62E569B9-547B-4144-80AC-03A9EBFC45DA.pdf (last accessed 12 May, 2026).

Artificial Intelligence surveillance is being used by the government without oversight and this is a big problem. Many scholars think that India's surveillance system does not meet the standards set by the court in the case of K.S. Puttaswamy v. Union of India[10] which're legality, necessity and proportionality. They also worry that Artificial Intelligence surveillance can target some groups more than others and it can stop people from speaking out[11].

The fourth part is about the laws and regulations related to data and Artificial Intelligence. Some commentators have written about the Digital Personal Data Protection Act, 2023 and the new telecommunications and interception rules. They say that these laws give the government a lot of power to access data. While some people think this is necessary for security others think it can be used to spy on people. They also say that India does not have a regulatory statute for Artificial Intelligence and this creates a big problem[12].

The main problem with the existing literature is that the technical and legal aspects are not well connected. The technical people show how useful AI is. Instead, they critique the use of Artificial Intelligence. They do not understand the technical details. This paper tries to draw a nexus and "connect" these two fields. It says that when digital intermediaries use Artificial Intelligence for security they become like public actors and their technological choices have big consequences[13]. It argues that India's challenge is not whether to use Artificial Intelligence but how to control its use and make sure that it is fair and transparent.

RESEARCH METHODOLOGY

This paper adopts a primarily doctrinal legal methodology, supplemented by comparative policy analysis. The doctrinal component comprises Indian statutes, rules and leading judgments in regards with data interception, data protection, intermediary liability and privacy. This analysis is predominantly focused on textual interpretation, legislative intent and judicial standards for surveillance and privacy.

[10] *Justice K.S. Puttaswamy (Retd.) v. Union of India*, (2017) 10 SCC 1 (India).
[11] *AI, Surveillance and Privacy in India: Human Rights in the Age of Technology*, Oxford Human Rights Hub. https://ohrh.law.ox.ac.uk/ai-surveillance-and-privacy-in-india-human-rights-in-the-age-of-technology/ (last accessed 12 May, 2026).
[12] *Cyber and AI Governance Law in India: A 2026 Legal Roadmap*, India Law Shield. https://www.indialawshield.com/Cyber-and-AI-Governance-Law-in-Indiablog1.html (last accessed 12 May, 2026).
[13] *Strengthening AI Governance Through Techno-Legal Framework*, PSA. https://psa.gov.in/CMS/web/sites/default/files/publication/AI-WP_TechnoLegal.pdf (last accessed 12 May, 2026).

The primary sources utilized for this research include statutes, court judgements, state rules and official procurement documents. The secondary sources include peer-reviewed articles and papers, technical reports, think tank analyses and news reported on the agendas of cyber threats and state surveillance.

In order to address the doctrinal gaps, the paper uses normative legal arguments rooted deep in the constitutional principles (like that of, legality, necessity, proportionality and procedural fairness) to justify the regulatory interventions.

DIGITAL INTERMEDIARIES AS SECURITY AGENTS

AI inculcated digital intermediaries are active agents when it comes to the national security of India's digital ecosystem. Certain actions like Data aggression, i.e. retaining metadata and logs, operating ML/AI channels to flag malicious content, responding to state requests for data, real-time interception and many more are taken care by such digital intermediaries. This form of reconfiguration creates a "quasi-public" security architecture, where private entities exercise de facto policy functions which were earlier dealt with by the state. This raises questions of delegation liability, and accountability against such private intermediaries[14].

Now, the existing Indian regulations treat such intermediaries in an unsystematic way. Section 79 of the IT Act[15] made an attempt at conditions on due-diligence and compliance methods however, the recent AI enforcements by these digital intermediaries expect proactive content moderation and cooperation. These policy-laden decisions and hence can't be left solely to private governance. The law is required to recognize the intermediaries' role and impose certain duties proportionate to the public-security functions they perform.

AI in Cyber Threat Detection

It has already been established by many studies that AI-based systems are central to modern cyber defense by enabling real-time anomaly detection and automated incident triage across networks and cloud platforms. However, despite its benefits, using AI for cyber detection raises legally salient technical limits. Many high-performing AI models are "black boxes", making post-hoc justification of automated actions difficult. Now, these AI systems work by reflecting historical pre-saved datasets and may produce skewed detection outcomes by

[14] *An Analysis of Intermediary Liability in India and the European Union*, Manupatra Articles. https://articles.manupatra.com/article-details/AN-ANALYSIS-OF-INTERMEDIARY-LIABILITY-IN-INDIA-AND-THE-EUROPEAN-UNION (last accessed 12 May, 2026).

[15] Information Technology Act, No. 21 of 2000, § 79 (India).

generating a disproportionate scrutiny of particular networks. Another loophole which arises due to lack of statutory regulations of AI operators is that the attackers can craft inputs to evade and poison models. This would lead to false negatives and attributions which have both legal and diplomatic consequences across the globe. These automation bias possess much severe risks which may not surface to the superficial levels at the prima facie, when executed without human intervention[16].

From a legal standpoint, these limitations implicate due process. In case AI outputs are a party to investigative measures, the affected people must have avenues to contest findings. This ends up requiring documented audits and hence requiring human intervention again at the end of the day. Even the mega industries which deploy AI intermediaries are required to meet best industry practices like robust testing and standards[17]. They are held responsible and face conditional liabilities in cases of negligence in implementation results in rights violation or data breaches, etc. Regulatory frameworks must therefore translate technical safeguards into legal obligations that ensure both national security efficacy and constitutional protections.

AI in State Surveillance

The digital intermediaries with AI operators function by scaling identification, profiling and predictive functions across vast fragments of data present. One major example of this includes “face recognition” for instant identification. Another would be graph analytics for risk scoring. These intermediaries help both investigations and surveillance. These are a testimony to how state agencies leverage private data and vendor analytics to generate investigative leads to increase situational awareness.

Many surveillance needs operate under very broad statutory powers with administrative rules that lack granular safeguards. The safeguards, which are the need of the hour while handling AI intermediaries in cases of state security[18]. This framework would help tackle non-transparency which occurs at the instances of emergency data access and algorithmic profiling without independent authorization. AI’s opacity and error rates may exacerbate the risk of wrongful targeting. For example, failure of proper facial recognition can lead to

[16] Siva Vignesh S.K.V. & Nagarjun D.N., *Legal Challenges of Artificial Intelligence in India’s Cyber Law Framework: Examining Data Privacy and Algorithmic Accountability via a Comparative Global Perspective*, 6(6) IJFMR (2024). https://www.ijfmr.com/research-paper.php?id=31347 (last accessed 12 May, 2026).

[17] *Cyber and AI Governance Law in India: A 2026 Legal Roadmap*, India Law Shield. https://www.indialawshield.com/Cyber-and-AI-Governance-Law-in-Indiablog1.html (last accessed 12 May, 2026).

[18] *An Analysis of Surveillance and Data Protection with Reference to the Right to Privacy*, IJIRL PDF. https://ijirl.com/wp-content/uploads/2022/08/AN-ANALYSIS-OF-SURVEILLANCE-AND-DATA-PROTECTION-WITH-REFERENCE-TO-THE-RIGHT-TO-PRIVACY.pdf (last accessed 12 May, 2026).

misidentification , wrongful detention and much worse consequences even, affecting marginalized communities due to algorithmic biases.

Evidentiary and fairness concerns arise out of reliance on intermediary-provided analytics outputs as "admissible evidence" in the criminal proceedings. Courts are mandated to probe algorithmic provenance before accepting AI-derived inferences as reliable. The absence of a statutory or a mandate model adds on to such admissibility shortcomings[19]. Remedies shall therefore aim to restrict intrusive surveillance at the design and decision levels. Hence, only through combined procedural and substantive safeguards can AI-enabled surveillance be reconciled with constitutional protections.

INTERMEDIARY LIABILITY AND REGULATORY GAPS

The ecosystem of digital intermediaries with AI entities has evolved rapidly though, unevenly. This produced gaps that are problematic for algorithmic security activities. The rise of proactive moderation and security-oriented AI blurs the line between neutral hosting and investigatory actions. This technical shift exposes the lacuna of existing safe-harbor formulations as they do not adequately apply to the harms arising from the algorithmic decision making nor do they prescribe the required governance standards for high-risk functions like surveillance and cyber threat detection[20].

These procedural gaps give birth to substantive lacunas as well. The data protection laws and interception rules contain exemptions for state agencies that permit expansive state oversight, without commensurate oversight. They undermine transparency while making proper judicial review difficult. Also, there is no "unified" legal standard defining what counts as "high-risk" AI deployment.

The liability rules hence remain unsettled. Courts have grappled with platform responsibility in content moderation and copyright contexts, but there is scant jurisprudence on platform liability for erroneous security actions that harm users or third parties. Without clear duty-of-care norms, these digital intermediaries face regulatory uncertainty. Overly broad immunities often leave victims without a remedy. Excessive liability might chill security innovation and cooperation with state agencies.

[19] *AI, Surveillance and Privacy in India: Human Rights in the Age of Technology*, Oxford Human Rights Hub. https://ohrh.law.ox.ac.uk/ai-surveillance-and-privacy-in-india-human-rights-in-the-age-of-technology/ (last accessed 12 May, 2026)

[20] An Analysis of Intermediary Liability in India and the European Union, Manupatra Articles. https://articles.manupatra.com/article-details/AN-ANALYSIS-OF-INTERMEDIARY-LIABILITY-IN-INDIA-AND-THE-EUROPEAN-UNION (last accessed 12 May, 2026).

Possible options for better governance:[21]

- Statutory AI-risk classification: designate AI systems used for surveillance and critical infrastructure defence as "high-risk" and subject them to mandatory pre-deployment risk assessments and independent audits.
- Explainability and audit trails: require model cards, data provenance documentation and logging sufficient for judicial or independent review while protecting sensitive security details.
- Judicial authorisation for intrusive surveillance: restore and expand warrant or review mechanisms for bulk access or live biometric identification, limiting exemptions to narrowly defined emergencies.
- Liability and shared responsibility: impose obligation-based liability on intermediaries that deploy high-risk AI like duty of care, mandatory reporting of model failures, paired with safe-harbour only where reasonable technical and governance safeguards are demonstrably in place.
- Redress and oversight institutions: create an independent AI oversight body with powers to audit, sanction and require remedial measures; ensure affected persons can access review and remedies.

DISCUSSION

Through the vision of weight scales, national security and rights require "legal realism" about operational needs and technological constraints. The policymakers, at stake, must acknowledge that AI materially strengthens cyber defense and can be used for protecting the digital infrastructure. However, at the same time, an unregulated integration of AI in the digital intermediaries risks undermining constitutional guarantees. The result is a governance model in which private platforms may effectively become the first line of security screening, even though the Constitution requires that such intrusive powers be exercised under clear legal standards and meaningful oversight[22]. Hence, the regulatory designs must enable effective defenses and ensure accountability for apparent errors and biases.

[21]AI Governance Through Techno-Legal Framework, PSA. https://psa.gov.in/CMS/web/sites/default/files/publication/AI-WP_TechnoLegal.pdf (last accessed 12 May, 2026).
[22]Cyber and AI Governance Law in India: A 2026 Legal Roadmap, India Law Shield. https://www.indialawshield.com/Cyber-and-AI-Governance-Law-in-Indiablog1.html (last accessed 12 May, 2026)

One major concern is that the current framework encourages over-compliance and opaque decision-making. The 2021 intermediary rules[23] strengthened platform obligations, but they also produced uncertainty about the level of proactive monitoring expected from intermediaries. This uncertainty can push intermediary platforms toward defensive takedowns and broad content restriction, particularly where orders or requests from the State are involved. The practical effect might produce a chilling environment for speech, criticism and dissent, especially when algorithmic systems have the power to flag content at scale without transparent criteria. Studies on digital governance in India has repeatedly noted that such opacity can suppress lawful expression while leaving users without a meaningful appeal mechanism[24].

Drafting a narrowly tailored definition of "surveillance" and "cyber defence" in law while addressing AI operators is important to avoid overreach. Reforming Data law exemptions when the state claims bulk access powers, to require judicial and parliamentary oversight is also required[25]. These shall make our data protection and interception laws more fit to India's federal law enforcement architecture and cyber threat environment.

Another important issue is intermediary liability. Indian law still largely treats intermediary immunity as a "safe-harbour doctrine" tied to due diligence and notice-based compliance[26]. That model is increasingly inadequate for AI-mediated security functions because the intermediary is no longer just hosting third-party content; it is actively deciding what is flagged, suppressed, traced or escalated. A more realistic approach is obligation-based liability: the law should ask whether the intermediary used reasonable safeguards, maintained audit trails, tested its systems for bias and robustness and ensured human review for high-risk decisions. This approach preserves innovation while preventing total immunity from becoming a shield for harmful automated conduct[27].

---

[23] Rule 3(1)(b), Intermediary Liability, and the Burden of "Reasonable Efforts," NLS Forum on Internet Law & Policy. https://forum.nls.ac.in/ijlt-blog-post/rule-31b-intermediary-liability-and-the-burden-of-reasonable-efforts/ (last accessed 12 May, 2026).

[24] Indian Intermediary Liability Regime, CIS India. https://cis-india.org/internet-governance/files/indian-intermediary-liability-regime (last accessed 12 May, 2026).

[25] National Data Governance Framework Policy, MeitY/Invest India summary. https://www.investindia.gov.in/team-india-blogs/national-data-governance-framework-policy-accelerating-digital-government (last accessed 12 May, 2026).

[26] Legal Challenges of Artificial Intelligence in India's Cyber Law Framework, IJFMR. https://www.ijfmr.com/research-paper.php?id=31347 (last accessed 12 May, 2026).

[27] Digital Democracy and Media Ethics: Legal Dimensions of Free Speech and Expression, RSIS. https://rsisinternational.org/journals/ijriss/articles/digital-democracy-and-media-ethics-legal-dimensions-of-free-speech-and-ex.. .(last accessed 12 May, 2026).

Overall, the discussion shows that India does not need to choose between security and rights. What it needs is a structured model in which AI-enabled intermediaries operate under law, not outside it. The present framework should be reoriented toward judicial oversight, narrowly tailored state access, mandatory technical safeguards and enforceable duties of care. Only then can AI strengthen cyber defence and state surveillance without converting intermediaries into opaque instruments of verification.

CONCLUSION

In the current scenario, digital intermediaries and their embedded systems have become inseparable components of India's contemporary national security infrastructure. The tasks range from real-time cyber threat detection to large-scale surveillance. Yet, this functional indispensability does not and must not translate into "unchecked authority". As per this paper's suggests, while AI does enhance detection speed, scale and predictive capability, it also magnifies risks. These risks directly hamper constitutional guarantees of privacy and due process. Hence, in the absence of tailored legal constraints and institutional oversight, the delegation of quasi-policing functions to intermediaries risks unaccountable forms of surveillance under the mantle of "national security".

Procedural safeguards are equally essential. Judicial authorization shall be the norm for intrusive operations such as live biometric identification or bulk metadata access, with narrowly circumscribed emergency exceptions subject to after-the-fact review. Data-protection exemptions that currently permit broad agency access should be reformed to require parliamentary oversight or judicial sign-off and periodic review to prevent mission creep and institutionalisation of dragnet practices.

In sum, India can retain the operational benefits of AI-driven cyber defence while preventing the normalization of unchecked surveillance, but only through coherent statutory frameworks, calibrated procedural checks and enforceable intermediary duties. The choice is not between security and liberty; it is about structuring governance so that security-enhancing technologies operate within democratic and constitutional bounds. Implementing the recommendations outlined here will help ensure that intermediaries act as responsible partners in national security rather than unaccountable agents of technological coercion.