International Journal of Research in Management & Social Science
Volume 13, Issue 4(III) October - December 2025
ISSN 2322 - 0899

TO EVALUATE CYBERSECURITY RESILIENCE IN AI-DRIVEN DIGITAL ENTERPRISES: AN INDIAN PERSPECTIVE WITH GLOBAL INSIGHTS

Anika Dave and Drashti Nandu
TYBMS (Management studies), Usha Pravin Gandhi College of Arts, Science & Commerce, Mumbai
[1]anika.yogen@gmail.com and [2]drashtinandu29@gmail.com

ABSTRACT
The rapid rise of Artificial Intelligence (AI) in Indian enterprises is transforming sectors such as IT services, banking, manufacturing, telecom, and healthcare through predictive analytics, automation, and smarter decision-making. However, this growth introduces new cybersecurity threats like data poisoning, adversarial manipulation, model inversion, and deepfake-enabled fraud. India lacks AI-specific security regulations, with CERT-In offering only general guidance. This study examines the AI threat landscape, benchmarks global practices, and proposes a resilience framework covering prevention, detection, response, and recovery. Recommendations include proactive AI security measures, regulatory reforms, and stronger cross-sector collaboration to safeguard India's AI-driven digital economy.

Keywords: *AI, Cybersecurity, Data Poisoning, Adversarial Attacks, AI Governance*

INTRODUCTION
What if the voice on the other end of the line, the face on your video call, or the report generated by your most trusted system wasn't real but a flawless creation of Artificial Intelligence? In an era where AI is shaping industries, powering decision-making, and redefining efficiency, this unsettling scenario is no longer science fiction, it is an emerging cybersecurity reality.

AI has quickly turned into one of the biggest game-changers of the 21st century. It allows for automation, predicting trends, understanding human language, and making quick decisions.

AI is crucial for improving everything from supply chains to detecting fraud. In India, its use is growing fast across various fields like IT services, banking, manufacturing, healthcare, and telecommunications. A report by NASSCOM (2024) suggests that innovations driven by AI could add up to $500 billion to India's economy by 2025. Major companies like Infosys, Wipro, TCS, HDFC Bank, Reliance Jio, and Apollo Hospitals are integrating AI deeply into what they do.

Yet, while AI has many advantages, it also brings about new and complicated cybersecurity risks. Unlike older IT systems, AI models rely heavily on data and can change over time, which makes them vulnerable to:

- Data Poisoning – injecting malicious data into training sets to bias AI outputs.
- Adversarial Attacks – subtly altering inputs to deceive AI systems.
- Model Inversion and Theft – extracting sensitive training data or replicating proprietary AI models.
- Deepfake and Synthetic Media Fraud – creating convincing yet fraudulent content to manipulate public perception or execute financial scams.

These threats are often difficult to notice, expensive to handle, and can cause severe damage, leading to operational issues, financial loss, and harm to reputation without obvious signs of a problem. In India, these dangers are made worse by the lack of specific regulations for AI cybersecurity. While CERT-In provides general warnings, there are no strict guidelines focusing on AI-related weaknesses, resulting in a lack of clear policies.

This study looks into cyber threats related to AI in India, compares how prepared companies are globally, and suggests an AI Cybersecurity Resilience Framework to protect the country's digital economy through technical, organizational, and regulatory strategies.

LITERATURE REVIEW

1. AI in Business Transformation
Artificial Intelligence (AI) is reshaping the way enterprises operate, enabling predictive analytics, automation, and enhanced decision-making. In India, AI adoption is expected to add $500 billion to the GDP by 2025 (NASSCOM, 2024). Large IT service providers such as Infosys, Wipro, and TCS are integrating AI into client solutions for automation and efficiency.

ISSN 2322 - 0899

In sectors such as banking, AI-powered fraud detection systems are helping identify suspicious activities in real-time, while in healthcare, AI-driven diagnostics are reducing diagnosis times. Globally, AI deployment in business is seen as a core driver of competitive advantage (Brynjolfsson & McAfee, 2023).

2. Cyber Threat Landscape for AI Systems
Unlike traditional IT systems, AI models introduce new attack vectors:

- **Adversarial Attacks:** These involve making subtle modifications to input data to mislead AI models (Szegedy et al., 2014). In image recognition systems, for example, attackers can change just a few pixels to cause incorrect classifications.
- **Data Poisoning:** Attackers manipulate training datasets to bias AI decision-making. This is particularly dangerous in financial fraud detection or healthcare diagnostics, where poisoned data could produce harmful predictions (Biggio & Roli, 2018).
- **Model Inversion Attacks:** Hackers attempt to reconstruct sensitive training data from AI models, compromising privacy (Fredrikson et al., 2015).
- **Deepfake & Synthetic Media Fraud:** AI-generated videos and audio can impersonate individuals to manipulate financial transactions or spread misinformation.

3. In India, the Indian Computer Emergency Response Team (CERT-In) has warned about rising incidents of deepfake-enabled frauds targeting corporates and government officials (CERT-In, 2024).

4. The PwC India AI Adoption and Impact Report (2023)**:** reinforces this gap, revealing that 65% of surveyed Indian enterprises have increased AI adoption in the past two years, but only 26% have implemented AI-specific cybersecurity protocols. The report highlights a concerning trend where organizations underestimate AI security risks, focusing more on functional deployment than on securing AI pipelines. It also stresses that AI-related breaches can result in 15-20% higher recovery costs compared to conventional cyber incidents, largely due to the need for retraining models, validating datasets, and restoring compromised AI systems.

This body of literature collectively indicates that while AI is becoming a critical enabler of digital transformation in India, security preparedness is lagging significantly behind adoption rates. There is a pressing need for an integrated approach combining technical safeguards, regulatory clarity, workforce training, and international best-practice alignment to mitigate AI-specific cybersecurity risks.

AI-SPECIFIC CYBER SECURITY THREATS IN ENTERPRISES
AI-driven systems face unique vulnerabilities beyond those of traditional IT infrastructure. For Indian enterprises, the following threats are most critical:

1. Data Poisoning Attacks:
Attackers manipulate training datasets to alter the behavior of AI models. In 2023, an Indian fintech startup reported losses of ₹8.5 crore after its fraud-detection AI was compromised due to poisoned historical transaction data.

2. Adversarial Attacks:
Small, imperceptible changes in AI inputs can mislead systems. For example, altering medical imaging inputs to misclassify tumors could disrupt diagnosis in AI-assisted healthcare systems.

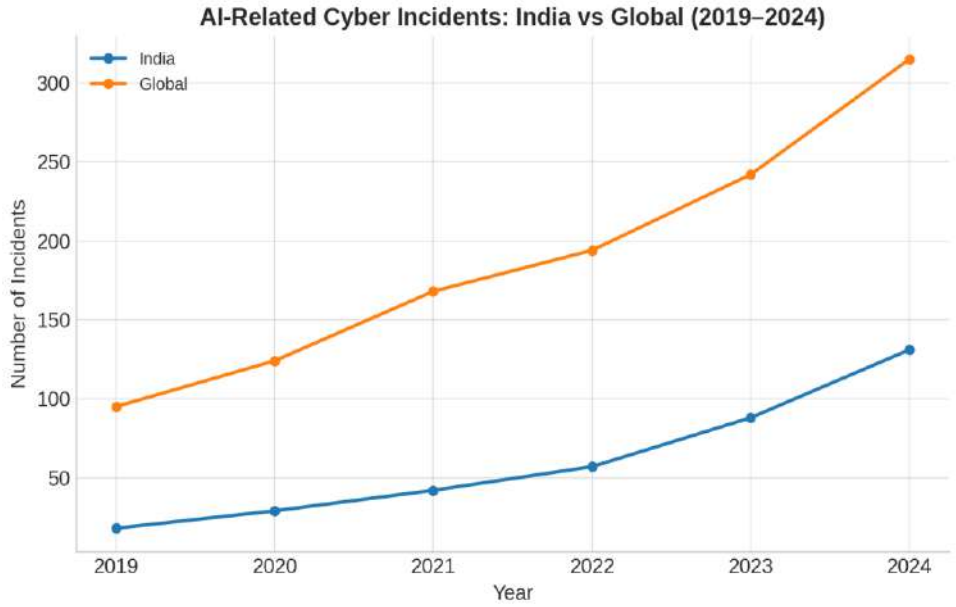
3. Model Inversion & Theft:
Hackers reverse-engineer AI models to extract sensitive training data or replicate proprietary algorithms. This is particularly concerning for IT service providers like Infosys or Wipro, whose AI models often contain client-specific data.

4. AI Supply Chain Vulnerabilities:
Many Indian firms use third-party AI APIs hosted abroad. A breach in any link of this AI supply chain whether a cloud provider or data labelling vendor can compromise the entire enterprise system.

ISSN 2322 - 0899

BUSINESS IMPACTS OF AI CYBER SECURITY BREACHES

AI-Related Cyber Incidents – India vs Global (2019–2024)



Source: Compiled from CERT-In, IBM Security, and NASSCOM reports (2019–2024).

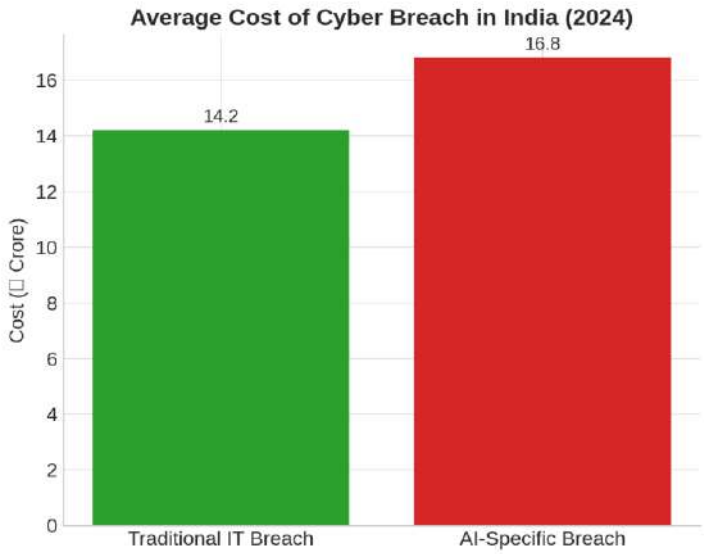
Observation:

- India's AI-related cyber incidents have surged from 18 in 2019 to 131 in 2024 a growth of >600% in just five years.
- The global trend shows higher absolute numbers, but India's rate of growth is significantly faster.
- Notable acceleration is visible after 2021, coinciding with increased AI adoption in finance, telecom, and healthcare sectors.

Implication:

- Rapid AI adoption in India is not matched by proportionate investment in AI-specific cybersecurity, leading to increased vulnerability.
- The steep rise post-2021 indicates that cybercriminals are quickly adapting their tactics to exploit AI systems in emerging economies.

Average Cost of Breach – AI vs Traditional (India, 2024)



IBM Security Cost of a Data Breach Report (2024)

ISSN 2322 - 0899

Observation:

- AI-specific breaches in India cost an average of ₹16.8 crore, compared to ₹14.2 crore for traditional IT breaches at an 18% cost premium.
- The additional cost is driven by factors such as model retraining, data integrity restoration, and AI system downtime.

Implication:

- AI-specific breaches are not only more sophisticated but also more expensive to resolve.
- This cost gap suggests that prevention measures, though initially costly, would be far more economical than remediation after an incident.

BUILDING CYBER SECURITY RESILIENCE IN AI ENTERPRISES

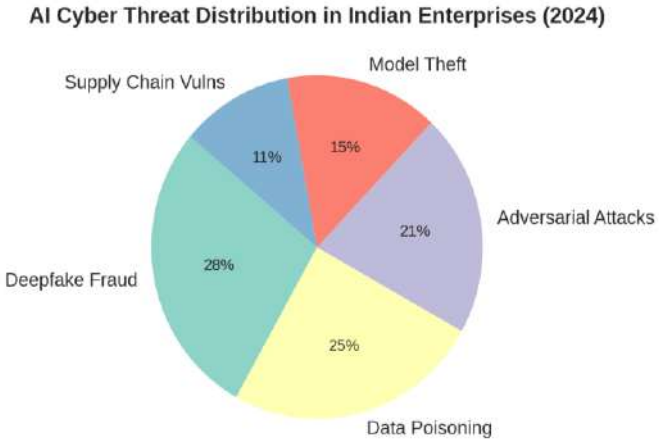
In order to establish cybersecurity resilience in AI businesses, a strategic combination of technology, governance, and culture is necessary. AI models are exposed to particular hazards that are not present in conventional IT systems, such as adversarial manipulation, data poisoning, and model theft, necessitating specific defense strategies. Resilience starts with integrating security-by-design concepts throughout the development of the AI system, making sure that threat modeling, model integrity testing, and adversarial input detection are incorporated before deployment. With this proactive strategy, you may lessen potential weaknesses that may be taken advantage of after implementation.

The development of strong incident identification and response procedures suited to AI settings is another crucial component. This encompasses implementing AI-driven threat detection systems, real-time monitoring of AI outputs for irregularities, and developing incident playbooks that are specific to AI for rapid containment and recovery. The necessity for quick response skills is brought about by the fact that AI systems frequently function in high-stakes sectors such as healthcare, finance, and infrastructure, where even little interruptions can have major domino consequences.

Continuous model validation and retraining are equally crucial. Since AI systems are constantly changing and learning from data over time, they are more vulnerable to assaults if they are not routinely evaluated. Using encrypted model storage, setting up safe data pipelines, and performing routine security audits all help to maintain the reliability and resilience of AI models.

Resilience is an organizational problem as well as a technical one. To aid workers in identifying AI-enabled hazards like manipulated datasets and deepfake scams, AI firms should invest in employee awareness and training programs. Before technological instruments raise any anomalies, this cultural layer of defense is frequently the initial line of detection. Sharing AI threat intelligence across industries and with government organizations like CERT-It can allow for early warning of new threats, and cooperation is also crucial.

Last but not least, regulatory alignment is essential. To improve compliance and gain consumer confidence, AI businesses should stay ahead of the curve when it comes to cybersecurity laws unique to AI by implementing international best practices, such as the NIST AI Risk Management Framework and the EU AI Act. Resilience will ultimately depend on the capacity of AI firms to integrate cutting-edge security technologies, flexible governance structures, expert human monitoring, and cooperative defense strategies in order to turn AI from a possible vulnerability into a safe engine of innovation and competitive advantage.



Source: NASSCOM AI Security Survey (2024).

ISSN 2322 - 0899

AI Cyber Threat Distribution in Indian Enterprises (2024)

Observation:
- Deepfake Fraud (28%) and Data Poisoning (25%) together account for more than half (53%) of AI-specific cyber incidents.
- Adversarial Attacks (21%) remain a significant concern, especially in autonomous systems and image recognition.
- Model Theft (15%) and Supply Chain Vulnerabilities (11%) are emerging threats, often overlooked in current security strategies.

Implication:
- Indian enterprises must prioritize defenses against deepfake and data poisoning attacks.
- There is a risk of underestimating model theft and supply chain risks, which could cause strategic IP loss and long-term security compromises.

CASE STUDY

Deepfake CFO Scam in Hong Kong (2024)

Background
In February 2024, a Hong Kong-based multinational lost HK$200 million (≈ $25 million / ₹207 crore) in a deepfake-enabled fraud. Cybercriminals used AI-generated video and audio to impersonate the UK-based CFO and other executives during a live video conference.

Incident Overview
- Attackers trained deepfake models using public recordings of executives.
- Finance staff attended a virtual meeting believing they were speaking to real leaders.
- The "CFO" instructed urgent fund transfers for a confidential acquisition, leading to 15 transactions before detection.

Impact & Lessons
- Loss: HK$200 million; operations disrupted.
- Lessons: Enforce multi-channel verification, use deepfake detection tools, limit public exposure of executive media, and train staff to identify AI impersonation.

LIMITATIONS
While this research provides valuable insights into the cybersecurity challenges of AI-driven enterprises in India, several limitations must be acknowledged:

1. Data Availability Constraints
- There is no centralized public database of AI-specific cyber incidents in India.
- Many companies do not disclose AI-related breaches due to fear of reputational damage, leading to underreporting.
- Incident statistics often merge AI-specific and traditional IT breaches, making it harder to isolate AI-focused attack trends.

2. Evolving Nature of Threats
- The AI cyber threat landscape changes rapidly, with new attack techniques emerging almost monthly.
- Any findings from this research may require frequent updates to remain relevant.
- Tools and techniques used by cybercriminals evolve faster than corporate cybersecurity measures.

3. Scope Restriction
- This study focuses primarily on enterprise-level AI use cases in sectors such as IT, BFSI, manufacturing, healthcare, and telecom.
- Small and medium-sized enterprises (SMEs) and government institutions, which may have different AI adoption patterns, are not deeply covered.

ISSN 2322 - 0899

4. Limited Empirical Testing

- While the proposed resilience framework is based on industry best practices and case studies, it has not been empirically tested across multiple real-world enterprises in India.
- Further longitudinal studies are needed to validate its effectiveness in diverse industry contexts.

5. Regulatory Gaps

- India currently lacks binding, enforceable AI-specific cybersecurity regulations, meaning organizations are free to interpret guidelines differently.
- Comparative analysis with more mature AI regulatory environments (EU, US) is constrained by these policy differences.

AI-SPECIFIC CYBER SECURITY POLICIES AND REGULATORY LANDSCAPE

The growth of Artificial Intelligence (AI) has revolutionized industries worldwide, but it has also created novel and complicated cybersecurity challenges. Although regulatory frameworks have not kept pace with these developing dangers, AI-driven attacks like data poisoning, adversarial inputs, and deepfake-enabled fraud are becoming more prevalent and sophisticated.

The majority of cybersecurity legislation in India continues to be centered on conventional IT systems. The main legislative framework for cybersecurity and cybercrime prevention is the Information Technology Act of 2000 (and its amendments), but it does not address AI-specific issues. Although the Digital Personal Data Protection Act of 2023 addresses the privacy, consent, and processing of personal data, it places a greater emphasis on user data than on the security of AI algorithms, models, and training datasets. While the Indian Computer Emergency Response Team (CERT-In) has established standards for incident reporting, system hardening, and log retention, these criteria are still quite broad and do not specifically cover attack routes that are unique to AI. Despite the ethical and trust factors surrounding AI being covered by the NITI Aayog's National Strategy on Artificial Intelligence (2018), it is not a legally enforceable framework and provides little advice on specific cybersecurity measures.

Conversely, global events point to a more focused strategy. The European Union AI Act (2024) categorizes AI systems into risk categories of unacceptable, high, limited, and minimal, and it mandates that high-risk systems be subjected to thorough risk assessment, bias testing, and security audits. In the United States, the NIST AI Risk Management Framework (2023) provides organized instructions for identifying, analyzing, and mitigating AI-related risks throughout the system lifecycle. The AI Governance Framework in Singapore places a strong emphasis on accountability, transparency, and the necessity for human supervision in decision-making systems powered by artificial intelligence. These frameworks are proactive efforts to foresee threats unique to AI, not just respond to them.

Despite this progress, India still has a big gap in addressing cybersecurity threats brought about by AI. There is no legal requirement for businesses to implement security-by-design principles in AI development, report AI system vulnerabilities, or routinely test the integrity of AI models in the absence of laws specific to AI.

As a result, businesses are vulnerable to sophisticated attacks that can circumvent conventional security measures. Furthermore, there is no agreed-upon categorization of AI systems based on their possible impact on cybersecurity, and enforcement procedures are ambiguous.

A collaborative approach between regulators, industry participants, and technology developers will be necessary to close this gap. India might develop a national AI security policy that is context-specific by modifying pertinent elements of the NIST AI Framework and EU AI Act. Mandatory AI risk classification, security standards for AI providers, and industry-wide adoption of incident reporting criteria particular to AI should be included in such a framework. India runs the danger of falling behind its worldwide counterparts in safeguarding companies and people from the next wave of AI-enabled cyber threats if it does not take such proactive measures.

RECOMMENDATIONS

The growing sophistication of AI-enabled cyberattacks demands a holistic, multi-stakeholder response that integrates technical safeguards, organizational readiness, legal reform, and public awareness. Based on the findings of this study, the following recommendations are proposed:

1. Establish AI-Specific Cybersecurity Standards

India should develop and enforce national AI security guidelines tailored to the unique vulnerabilities of AI systems. These must address:

ISSN 2322 - 0899

- Model Integrity Testing: Regular adversarial testing to identify weaknesses in AI models before deployment.
- Dataset Security: Protection of training data from poisoning and unauthorized manipulation.
- Algorithm Protection: Encryption of AI models and deployment in secure environments to prevent theft or tampering.

2. Introduce AI Risk Classification and Compliance Framework
Adopt a system similar to the EU AI Act where AI systems are classified as unacceptable, high, limited, or minimal risk. For high-risk systems such as those used in finance, healthcare, or national infrastructure implement mandatory risk assessments, continuous monitoring, and annual third-party audits.

3. Mandate Security-by-Design Principles
Require AI developers and vendors to integrate security measures from the design phase, rather than retrofitting them after deployment. This includes:

- Built-in defense against adversarial inputs.
- Secure update and patching mechanisms.
- Logging and traceability features to detect and investigate breaches quickly.

4. Strengthen Deepfake and AI-Fraud Countermeasures
Invest in AI-driven deepfake detection tools capable of identifying manipulated audio, video, and text in real time. Corporate networks should deploy these tools for executive communications, while social media platforms should make them mandatory for content verification.

5. Improve Incident Detection, Reporting, and Response
Develop AI-specific incident reporting protocols that require enterprises to report AI-related breaches to CERT-In within 24-48 hours. Establish a national AI threat intelligence sharing network to ensure faster cross-industry response and mitigation.

6. Enhance Workforce and Public Awareness

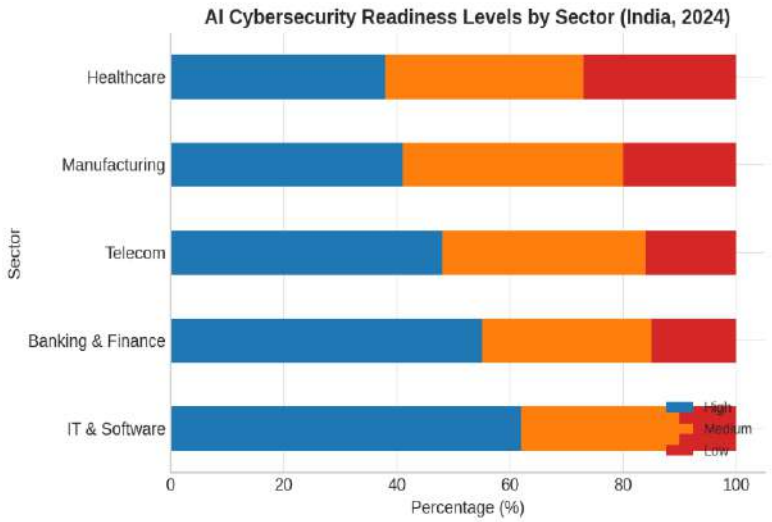
- **Enterprise Level:** Conduct AI-cybersecurity training for employees, especially in finance, IT, and leadership roles.
- **Public Level:** Launch government-led media literacy campaigns to educate citizens about AI-driven scams, deepfake threats, and safe verification practices.

7. Align with International Best Practices
Collaborate with global leaders in AI governance, adopting elements from:

- NIST AI Risk Management Framework (USA) for structured AI risk control.
- Singapore AI Governance Framework for transparency and accountability.
- OECD AI Principles for ethical and trustworthy AI deployment.

ISSN 2322 - 0899

AI Cybersecurity Readiness by Sector (India, 2024)



Source: Based on Industry survey

Observation:

- IT & Software leads with 62% high readiness, reflecting better awareness and access to skilled cybersecurity professionals.
- Banking & Finance (55%) and Telecom (48%) show moderate readiness but face constant targeting from AI-enabled fraud.
- Manufacturing (41%) and Healthcare (38%) have the lowest high-readiness scores, relying heavily on outdated security measures.
- Low readiness levels are highest in healthcare (27%), posing risks for patient safety and data privacy.

Implication:

- AI cybersecurity maturity is uneven across sectors, leaving critical industries like healthcare and manufacturing highly vulnerable.
- Cross-sector knowledge sharing and government-supported readiness programs are essential to close these gaps.

CONCLUSION

The rapid integration of Artificial Intelligence into Indian enterprises is driving innovation, efficiency, and competitive advantage. However, it is also introducing a new generation of cybersecurity threats that are more complex, targeted, and costly than traditional IT risks. This research has shown that AI-specific cyber incidents in India have risen sharply in recent years, with threats such as deepfake fraud, data poisoning, and adversarial manipulation posing significant challenges to business continuity and trust.

While India has made progress in general cybersecurity regulation, the absence of AI-specific frameworks leaves enterprises vulnerable to evolving threats. The findings suggest an urgent need for proactive, AI-focused resilience strategies that combine technical safeguards, organizational readiness, and regulatory reform.

The proposed AI Cybersecurity Resilience Framework centered on prevention, detection, response, and recovery offers a practical approach for Indian enterprises to secure their AI systems.

Ultimately, safeguarding AI in India will require collaboration between policymakers, industry leaders, and technology vendors, as well as alignment with emerging global best practices. By taking a proactive approach today, Indian enterprises can ensure that AI remains a driver of growth rather than a source of systemic risk.

REFERENCES

NASSCOM. (2024). India's AI Adoption Report 2024: Opportunities and Challenges. National Association of Software and Service Companies.

CERT-In. (2022). Cyber Security Directions Under Sub-section (6) of Section 70B of the IT Act, 2000. Indian Computer Emergency Response Team.

ISSN 2322 - 0899

MeitY. (2023). Digital Personal Data Protection Act, 2023. Ministry of Electronics and Information Technology, Government of India.

European Commission. (2024). EU Artificial Intelligence Act. Official Journal of the European Union.

National Institute of Standards and Technology (NIST). (2023). AI Risk Management Framework. U.S. Department of Commerce.

IBM Security. (2024). Cost of a Data Breach Report 2024. IBM Corporation.

Gartner. (2023). Emerging Technologies: AI Security Threats. Gartner Research.

NITI Aayog. (2018). National Strategy for Artificial Intelligence. Government of India.

PwC India AI Adoption and Impact Report (2023)