ISRG Journal of Engineering and Technology (ISRGJET)

ISRG PUBLISHERS

Abbreviated Key Title: ISRG J Eng Technol

ISSN: 3107-5894 (Online)

Journal homepage: https://isrgpublishers.com/isrgjet/

Volume – II Issue-II (March-April) 2026

Frequency: Bimonthly

OPEN ACCESS

Cybersecurity 5.0: From Firewalls to Fully Autonomous Digital Protection

Dr. Alex Mathew[1*] & Westerman Sydney[2]

[1, 2] Bethany College, West Virginia, USA

| **Received:** 05-03-2026 | **Accepted:** 09-03-2026 | **Published:** 14-03-2026

***Corresponding author: Dr. Alex Mathew**

Abstract

The cybersecurity world has experienced a generational shift from system-based monitoring and border protection to wholly autonomous, AI-powered systems. The current paper follows the development of cybersecurity architectures across five generations and concludes with the 5.0 paradigm, characterized by agentic AI, Zero Trust architecture (ZTA), and self-directed incident response. Using the latest empirical sources, the paper will explore how machine learning, adversarial AI modeling, and autonomous response mechanisms are transforming organizational security postures in an age of increasingly sophisticated threats. It has been established that AI-based systems significantly reduce detection latency and breach costs, but at the same time pose governance and adversarial risks that will require preemptive policy-technical efforts (Kshetri, 2025; Adeyemi, 2023).

Keywords: *Cybersecurity 5.0, autonomous threat detection, artificial intelligence, Zero Trust Architecture, agentic AI, adversarial machine learning, incident response, data governance.*

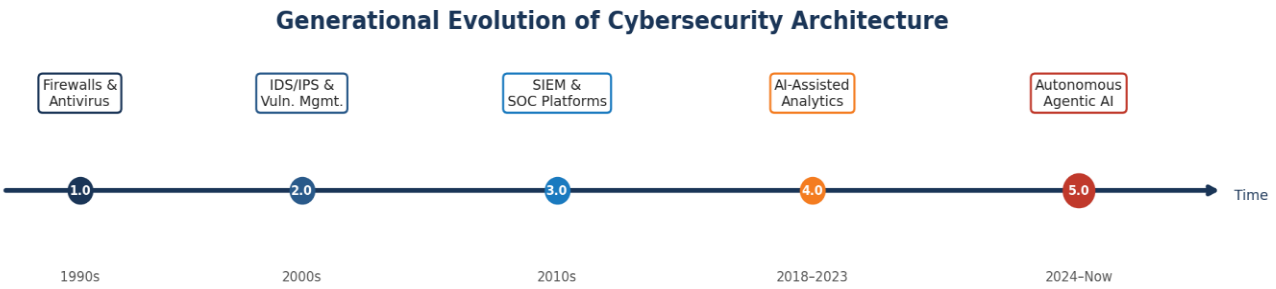
I. Introduction: The Autonomy Imperative

The threat surface organizations face has been fundamentally redefined by digital transformation. The spread of cloud infrastructure, IoT devices, and remote-work infrastructure has erased the traditional network perimeter, making traditional rule-based defenses more pointless than ever (Allioui & Mourdi, 2023). The opponents have currently used AI-based phishing, automated exploit models, and polymorphic malware, which significantly increase the speed of attacks (Olukoya et al., 2025).

Cybersecurity has therefore progressed to the 5th generation, Cybersecurity 5.0, which involves full autonomy and agency, meaning an autopilot and AI systems able to respond to threats in a multi-step fashion (Kshetri, 2025). The paper discusses the architecture, threat vectors, and governance imperatives of this fifth-generation paradigm, supported by figures and a comparative analysis table.

DOI: 10.5281/zenodo.19016855

Figure 1. *Generational evolution of cybersecurity architecture from perimeter-based defenses (Generation 1.0) to fully autonomous agentic AI systems (Cybersecurity 5.0).*



II. The Five Generations of Cybersecurity

Older security systems were based on perimeter defenses, where inbound traffic was implicitly trusted. They are victims of a fixed set of rules and malware signatures, which make them blind to new or zero-day threats. Statistics captured in Gracy (2025) indicate that the frequency and severity of breaches have been steadily increasing over the past 20 years, and this is directly linked to adversaries outpacing signature-based protection through the introduction of automation and AI. The growth of IoT ecosystems has energized these weaknesses, as Allioui and Mourdi (2023) indicate that billions of interconnected devices create an under protected entry point into organizational networks. Hassan et al. (n.d.) explain that institutional failure to govern these flawed pillars of IoT products further increases exposures even where technical constraints are present, as making next-generation autonomous forms of defense is structurally or systemically required.

III. Core Pillars of Cybersecurity 5.0

A. Autonomous Threat Detection

The shift from human-in-the-loop alert triage to automated incident response defines Cybersecurity 5.0 as the trend. According to Ismail et al. (2023), these platforms are characterized by combining AI-based eyes and orchestration, such as the ability to perform containment, isolation, and remediation without analyst attention, with a mean-time-to-respond (MTTR) that has been dramatically reduced. This article confirms a systematic review showing that autonomous tools are always faster and more reliable in processing high-volume threats than human-led processes (Adeyemi, 2023). Faruquee (2025) also illustrates that the integration of AI and cybersecurity has led to quantifiable improvements in business continuity, with the AI security and market worldwide expected to expand to 136.18 billion dollars in 2032, a significant increase from the current 23.12 billion in 2024.

B. Zero Trust Architecture

Zero-trust architecture: It is a network architecture that implements the principle of continuously authenticating each access request, regardless of its origin within the network. As described by Paul et al. (2024), the combination of ZTA and AI analytics can be defined as a synergistic model where the ML model will consider the indicators of behavior, such as posture, the abnormalities of a session, and geolocation data, and use them to make a real-time adaptive access-control decision. This substitutes the unchanging guidelines of policy with probabilistic risk scoring, empowering organizations to respond to credential theft and insider attacks that would have otherwise gone undetected by existing controls (Paul et al., 2024).

C. Self-Healing Systems

Kshetri (2025) explains agentic AI in cybersecurity as systems that can independently, goal-oriented, conviction, and multi-step task execution, i.e., threat investigation, vulnerability-focused, and cross-platform coordination, without human direction. Greeshma and Thankachan (n.d.) go further and present the Cyber Guardians 2.0 vision, stating that proactivity, transparency, and trustworthiness are design properties rather than compliance properties that should be designed as fundamental.

Table 1 *Comparative Analysis of Cybersecurity Generations across Key Operational Dimensions*

Dimension	Legacy (Gen 1–2)	SIEM (Gen 3)	AI-Assisted (Gen 4)	CS 5.0 (Gen 5)	Regulatory Alignment	Cost Impact
Threat Detection	Signature-based, static rules	Rule correlation & log analysis	ML anomaly detection + human review	Real-time behavioral AI, self-learning	Reactive compliance checks	High breach costs ($4.88M avg.)
Response Time	Hours–days (manual)	Hours (analyst-driven)	Minutes (semi-automated)	Seconds (bounded autonomy)	Audit trails required	~$1M savings (AI-assisted)
Zero-Day Coverage	None	Very limited	Partial (heuristic)	High (behavioral + predictive)	NIST CSF alignment	Reduced dwell time

DOI: 10.5281/zenodo.19016855

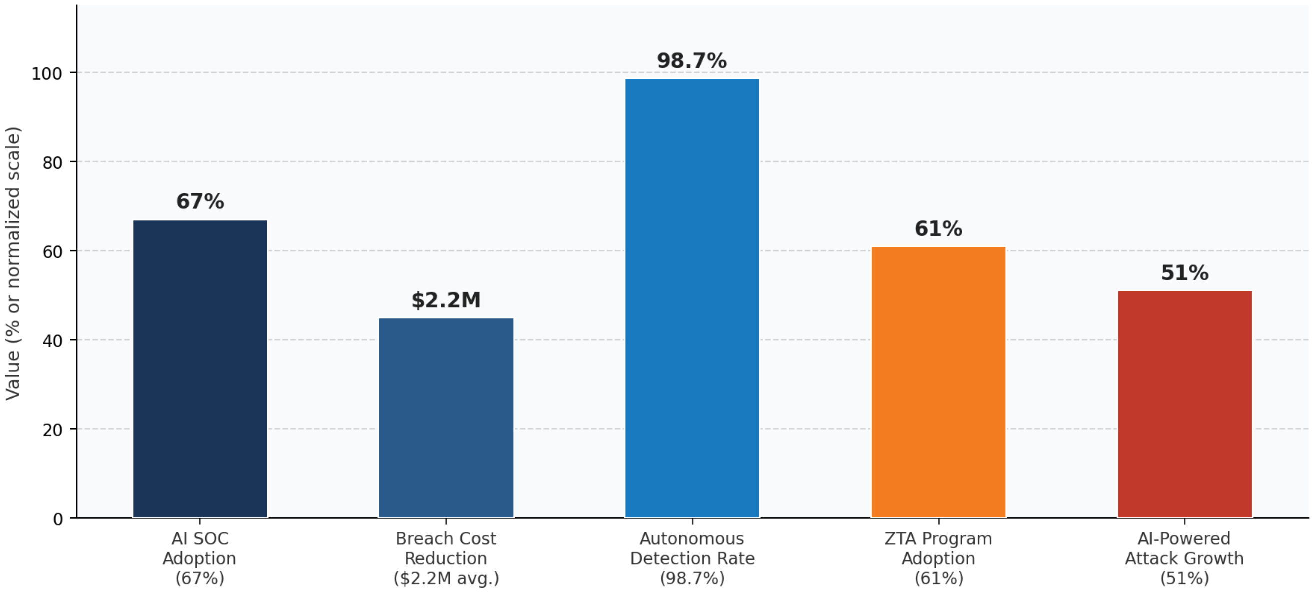
Dimension	Legacy (Gen 1–2)	SIEM (Gen 3)	AI-Assisted (Gen 4)	CS 5.0 (Gen 5)	Regulatory Alignment	Cost Impact
Human Oversight	Full (analyst-driven)	High (SOC analysts)	Moderate (triage support)	Strategic only (irreversible acts)	EU AI Act Art. 14	Lower headcount cost
Adversarial AI Resilience	None	Minimal	Partial (adversarial training)	Continuous model auditing & hardening	NIST AI RMF	$2.2M avg. savings (IBM, 2024)

Note. CS 5.0 = Cybersecurity 5.0. Data synthesized from IBM (2024, 2025); Adeyemi (2023); Paul et al. (2024); Kshetri (2025); Olukoya et al. (2025).

IV. The Threat Landscape: Why Autonomy Is Necessary

The two-sided nature of AI has sparked an arms race in cybersecurity: autonomous defenses use the same tools as adversaries. Olukaya et al. (2025) provide an account of how AI-assisted threats, such as AI-generated spear-phishing, automated exploitation, and deepfake social engineering, have increased exponentially in both magnitude and complexity. Sarker (2023) goes on to show that engineered adversarial examples can be used to abuse intrusion detection models into classifying malicious activity as benign. Adversarial ML is suggested by Ijiga et al. (2024) as both an attack and a defense, and it is demonstrated that training AI on simulated adversarial examples significantly increases model resistance. The very defense of the AI is now a major operation frontier of Cybersecurity 5.0.

Figure 2. Key Performance Metrics: AI-Driven Cybersecurity (2024–2025)



Sources: IBM (2024, 2025); Adeyemi (2023); Paul et al. (2024); Olukoya et al. (2025)

***Figure 2.** Key performance metrics for AI-driven cybersecurity (2024–2025): AI SOC adoption, average breach cost reduction, autonomous detection rate, Zero Trust Architecture program adoption, and AI-powered attack growth rate. Sources: IBM (2024, 2025); Adeyemi (2023); Paul et al. (2024); Olukoya et al. (2025).*

V. Benefits and Risks

Autonomous AI security implementations raise legal requirements. Jain (2024) emphasizes that the EU AI Act has placed systems with access revocation or network isolation in the high-risk category, and that such systems must be transparent or audited for bias and must include signed human oversight procedures. Hassan et al. (n.d.) point out that governance structure should include issues of accountability and explainability of AI-generated decisions. The current consensus on responsible deployment is in the principle of bounded autonomy, which allows AI agents to perform reversible actions without the need for human authorization on the use of irreversible actions (Greeshma & Thankachan, n.d.), as well as alignment, which is to ISO/IEC 27001:2022 and to the NIST AI Risk Management Framework (Jain, 2024).

VI. Conclusion

The paradigm shifts in the organization's digital protection, Cybersecurity 5.0, is not a traditional capability upgrade but a categorical one. A combination of agentic AI, Zero Trust Architecture, and autonomous incident response has never been as capable of neutralizing threats as quickly as machines can (Adeyemi, 2023; Kshetri, 2025; Faruquee, 2025). Nevertheless, the weaknesses of adversarial AI and regulatory requirements require no less strategic consideration (Ijiga et al., 2024; Sarker, 2023). Formal safety assurances for autonomous agents under adversarial distribution shifts, scale-independent explainability systems, and international interoperability standards for multi-agent security systems should be research priorities in the future.

DOI: 10.5281/zenodo.19016855

References

1. Adeyemi, D. S. (2023). Autonomous response systems in cybersecurity: A systematic review of AI-driven automation tools. Communication in Physical Sciences, 9(4), 878–898. https://journalcps.com/index.php/volumes/article/view/696/709
2. Allioui, H., & Mourdi, Y. (2023). Exploring the full potential of IoT for better financial growth and stability: A comprehensive survey. Sensors, 23(19), 8015. https://doi.org/10.3390/s23198015
3. Faruquee, Q. M. A. A. (2025). Scaling the significance of AI and cybersecurity implementation for efficient business management. Theseus. https://www.theseus.fi/handle/10024/909151
4. Gracy, S. S. (2025). A global analysis of data breaches from 2004 to 2024. arXiv preprint arXiv:2502.05205. https://arxiv.org/pdf/2502.05205
5. Greeshma, K. V., & Thankachan, N. M. (n.d.). CyberGuardians 2.0: A vision for proactive, transparent, and trustworthy AI in cybersecurity. ResearchGate. https://www.researchgate.net/profile/Greeshma-K-V/publication/398410296
6. Hassan, S. A. D. H., Rasheed, A. A., Mousa, A. A., Hussein, Z. A., & Ambudkar, B. (n.d.). HighTech and innovation. Academia.edu. https://www.academia.edu/download/125958293/11.pdf
7. Ijiga, O. M., Idoko, I. P., Ebiega, G. I., Olajide, F. I., Olatunde, T. I., & Ukaegbu, C. (2024). Harnessing adversarial machine learning for advanced threat detection: AI-driven strategies in cybersecurity risk assessment and fraud prevention. Journal of Science and Technology, 11(1), 1–24. https://www.researchgate.net/profile/Godslove-I-Ebiega/publication/380459159
8. Ismail, B. I., Abdul, S., Khan, S. M., Sattar, S. A., & Muhammad, S. (2023). AI for cyber security: Automated incident response systems [SSRN Working Paper No. 5477114]. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5477114
9. Jain, A. (2024). AI and privacy: Shifting from 2024 to 2025. Cloud Security Alliance Blog. https://cloudsecurityalliance.org/blog/2025/04/22/ai-and-privacy-2024-to-2025-embracing-the-future-of-global-legal-developments
10. Kshetri, N. (2025). Transforming cybersecurity with agentic AI to combat emerging cyber threats. Telecommunications Policy, 49(6), 102976. https://doi.org/10.1016/j.telpol.2025.102976
11. Olukoya, D., Amoran, S. O., Lawal, O., Altawati, M., Ibiyeye, S. O., Ibiyeye, A. O., & Onwuegbuchi, O. C. (2025). Data security and governance in the age of AI-enabled attacks. ResearchGate. https://www.researchgate.net/profile/Didunoluwa-Olukoya/publication/399192999
12. Paul, E. M., Mmaduekwe, U., Kessie, J. D., & Dolapo, M. (2024). Zero trust architecture and AI: A synergistic approach to next-generation cybersecurity frameworks. International Journal of Science and Research Archive, 13(2), 4159–4169. https://doi.org/10.30574/ijsra.2024.13.2.1738
13. Sarker, I. H. (2023). Multi-aspects AI-based modeling and adversarial learning for cybersecurity intelligence and robustness: A comprehensive overview. Security and Privacy, 6(5), e295. https://doi.org/10.1002/spy2.295

DOI: 10.5281/zenodo.19016855