A SYSTEMATIC LITERATURE REVIEW: ARTIFICIAL INTELLIGENCE IN MODERN CYBERWARFARE

by

CHRISTOPHER D. HARVEY

B.S., Albany State University, 2009

M.S., Mercer University, 2016

A Research Paper Submitted to the School of Computing Faculty of

Middle Georgia State University in

Partial Fulfillment of the Requirements for the Degree

DOCTOR OF SCIENCE IN INFORMATION TECHNOLOGY

MACON, GEORGIA

2026

A systematic literature review: artificial intelligence in modern cyberwarfare

Christopher Harvey, *Middle Georgia State University, christopher.harvey3@mga.edu*

Abstract

Artificial Intelligence (AI) continues to reshape the structure and pace of modern cyberwarfare, offering new capabilities that enhance defensive readiness and support rapid decision-making. AI-enabled systems have accelerated the speed and precision of cyber operations in ways once considered unattainable. Within cyberwarfare, where offensive and defensive operations intersect, AI presents both opportunities and emerging vulnerabilities. Although scholars acknowledge AI's expanding influence across the cyber domain, the literature reveals varied interpretations of its risks, benefits, and strategic implications. This diversity emphasizes the need for a qualitative synthesis that examines how AI is conceptualized in cyberwarfare contexts across multiple disciplines. This study addresses that need through a systematic literature review guided by PRISMA principles, identifying key themes related to automation, autonomy, analytics, and governance. The findings provide insight into how AI simultaneously strengthens cyber defense capabilities while introducing new ethical, operational, and strategic challenges.

Keywords: artificial intelligence, cyberwarfare, autonomous systems, defensive cyber operations, AI governance.

Introduction

Cyberwarfare refers to the use of computer network attack and defense operations to achieve strategic objectives in cyber space (Parks & Duggan, 2011). Existing scholarship recognizes that AI strengthens cyber defense through automation and advanced analytics. Yet researchers warn that autonomy introduces concerns surrounding reliability, transparency, and escalation. Antebi and Baram (2020) note that AI-driven systems may behave unpredictably in high-stakes environments, thereby complicating human oversight. Additionally, Almeida et al. (2017) argue that current governance frameworks do not adequately address the unique challenges posed by AI in military applications. Such gaps create uncertainty for policymakers and cybersecurity professionals, reinforcing the importance of reviewing how the literature frames AI's dual-use potential and the vulnerabilities that may emerge as adoption accelerates.

Artificial intelligence (AI) is broadly defined as the development of systems capable of performing tasks that typically require human intelligence, including learning, reasoning, and decision-making (Abbass, 2021). From a foundational perspective, AI has long been conceptualized as an effort to replicate or simulate human cognitive processes, bridging computational systems with elements of human reasoning and problem-solving (Copeland, 2019). This broader understanding is essential in cyberwarfare contexts, where AI extends beyond simple automation to include adaptive and autonomous decision-making capabilities.

Despite the rapid integration of artificial intelligence (AI) into cyberwarfare operations, there remains a lack of conceptual clarity and consensus in the literature regarding its strategic implications, risks, and governance requirements. While AI-driven capabilities enhance cyber defense through automation, analytics, and real-time decision-making, they also intensify longstanding challenges in defining the boundaries of cyberwarfare, including what constitutes a use of force or an armed attack in the digital domain (Boothby et al., 2012). Additionally, the evolving complexity of cyberwarfare technologies

introduced new operational and strategic uncertainties, as emerging tools and techniques outpace existing frameworks for understanding and countering cyber threats (Chelvachandran et al., 2020). This disconnect creates ambiguity for policymakers, cybersecurity professionals, and military strategists attempting to effectively integrate AI into cyber operations. As a result, a systematic synthesis of the literature is needed to clarify how AI is conceptualized within cyberwarfare, identify recurring themes, and better understand the balance between its operational benefits and emerging risks.

Real-world cybersecurity strategies increasingly emphasize enterprise-level coordination and AI-enabled defense capabilities, as reflected in federal guidance advocating for integrated and adaptive security architectures (Cybersecurity and Infrastructure Security Agency [CISA], 2025). However, the rapid evolution of AI-driven cyber capabilities continues to challenge existing definitions of conflict and the boundaries of cyber operations, particularly in determining what constitutes a use of force in the digital domain (Boothby et al., 2012).

This research aims to explore how existing scholarship conceptualizes AI integration into cyberwarfare. AI-driven offensive tools introduce layers of complexity into modern conflict, making synthesis essential (Mirsky et al., 2021). This review contributes by clarifying how researchers describe AI's benefits, risks, and strategic implications to support responsible defense planning and governance. This study is guided by the following research questions:

RQ1: How does current scholarly literature characterize the role of AI-driven automation in shaping cyber defensive and offensive strategies?

RQ2: What themes emerge regarding risks, limitations, and ethical considerations associated with autonomous AI decision-making?

RQ3: How do researchers describe AI-enabled analytical capabilities and their role in improving situational awareness and threat response?

Review of literature

AI-driven automation and cyber defense capabilities

Research on Artificial Intelligence (AI) in cyberwarfare reveals a rapidly evolving and multidisciplinary field marked by both strategic opportunity and emergent risk. Early foundational work conceptualizes cyberwarfare as an intersection of computer network attack, defense, and technical operations (Parks & Duggan, 2011), a framework that continues to shape contemporary debates surrounding AI integration. As states modernize their cyber capabilities, AI-driven automation has emerged as a centralized defensive asset that can accelerate detection and response cycles. Industry analyses highlight how AI substantially increases the speed and precision of cyber operations, enabling defenders to counter sophisticated adversaries more effectively (Fortinet, 2024). Academic research echoes this emphasis: automation enhances defensiveness readiness, supports continuous monitoring, and provides predictive insights that were not previously achievable through human-driven analysis alone (Sarjakivi et al., 2024). Despite these advancements, the implementation of AI in cybersecurity environments remains uneven. Organizational, technical, and resource-related constraints, particularly within public sector institutions, continue to limit the full realization of AI-enabled capabilities (Mahusin et al., 2024). These challenges highlight the gap between theoretical potential and operational deployment.

Emerging vulnerabilities and strategic risks of AI in cyberwarfare

Despite these benefits, scholars warn that AI also introduces complex vulnerabilities. Antebi and Baram (2020) argue that autonomous or semi-autonomous systems can behave unpredictably in dynamic operational environments, creating new uncertainties for human oversight. The dual-use nature of AI

technologies amplifies this concern. Systems designed for defensive detection may be repurposed by adversaries for automated exploitation and large-scale cyberattacks, a trend that reinforces the cyber domain's offensive-defensive ambiguity (Mirsky et al., 2021). Research further shows that the opacity of AI decision-making, the so-called "black box" structure, limits trust and reduces explainability in time-sensitive military contexts, creating challenges for accountability and lawful use of force (Russell & Norvig, 2020).

The offensive application of AI further complicates the cyber threat landscape. Adversaries are increasingly leveraging AI to automate attacks, identify vulnerabilities, and scale malicious operations with greater speed and precision (Gabrian, 2024). This evolution is particularly concerning in the context of cyber-physical systems, where emerging threat actors, including those operating within dark web ecosystems, can target critical infrastructure using advanced cyberwarfare techniques (Mahor et al., 2021). These developments reinforce the dual-use nature of AI and expand the scope of potential cyber conflict.

Another central theme is that the literature involves the risk of unintended escalation. Scholars such as Dubber and Lazar (2025) caution that autonomous cyber agents may initiate or escalate digital conflict faster than human decision-makers can intervene, potentially triggering strategic instability. Related concerns focus on algorithmic and data vulnerabilities. Linkov et al. (2020) highlight that AI systems remain susceptible to biased training data, adversarial manipulation, and poisoning attacks, factors that undermine the reliability of AI-enabled defenses and create exploitable weaknesses in national security infrastructure.

Governance, regulation, and ethical challenges

Governance and regulatory gaps emerge as a final and critical concern across the literature. Almeida et al. (2017) contend that existing digital governance frameworks do not adequately address the ethical, operational, and geopolitical complexities introduced by AI militarization. Studies increasingly call for robust oversight mechanisms, clearer legal standards, and international norms to guide the use of AI in cyber operations. These concerns align with broader discussions about cross-domain warfare, where AI expands the speed and scale at which cyber actions can impact geopolitical outcomes (Katagiri, 2023).

Ethical considerations extend beyond operational risk to broader societal implications. The integration of AI into cyberwarfare raises questions surrounding accountability, moral responsibility, and the potential for unintended consequences in high-stakes environments (Demy et al., 2014). Additionally, concerns about algorithmic bias and systemic inequality suggest that AI-enabled systems may replicate or amplify existing disparities if not carefully governed (Gordon, 2019). These issues underscore the importance of aligning technological advancement with ethical and regulatory oversight.

Overall, the literature presents a field in transition, characterized by rapid technological advancement and uneven conceptual alignment across disciplines. While AI significantly enhances cyber defense capabilities, it simultaneously introduces new categories of strategic and technical risk. The lack of consensus on issues of autonomy, governance, and escalation accentuates the need for systematic synthesis.

Methodology

This study adopted a qualitative systematic literature review (SLR) design guided by Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework. The qualitative SLR approach ensured systematic identification, screening, and thematic synthesis of scholarly research. Sarjakivi et al. (2024) emphasize that AI research in cyber operations spans multiple disciplines, making a systematic review essential for clarity. A qualitative SLR allows for thematic pattern recognition and conceptual integration, making it particularly well suited to answer the research questions surrounding automation,

autonomous decision-making, analytical capabilities and governance implication. Additionally, cyberwarfare evolves rapidly, with technological innovation often outpacing regulatory frameworks. A structured review provides methodological transparency and replicability while reducing selection bias. Adhering to PRISMA principles surrounding this study ensured that literature identification, screening, and thematic synthesis were systematic rather than narrative or subjective.

Search strategy

A comprehensive literature search was conducted across six major academic databases: IEEE Xplore, Scopus, Web of Science, ACM Digital Library, JSTOR, and Google Scholar. The search strategy employed Boolean operators and keyword combinations designed to capture the intersection of AI and cyberwarfare. Core search strings included:

("artificial intelligence" OR "machine learning") AND ("cyberwarfare" OR "cyber warfare" OR "cyber operations")

("AI automation" OR "autonomous systems") AND ("cybersecurity" OR "cyber defense")

("AI analytics" OR "predictive analytics") AND ("cyber threat detection" OR "cyber conflict")

Inclusion, analysis, and ethical framework

To ensure relevance and methodological rigor, this study applied clearly defined inclusion and exclusion criteria when selecting sources for the systematic literature review. Only peer-reviewed scholarly research was considered, including journal articles, conference proceedings, and institutional reports that explicitly addressed the integration of artificial intelligence within cyberwarfare or military cyber operations. Additionally, studies were required to fall within the publication window of 2010 to 2026, reflecting contemporary developments in AI and cyberwarfare. Eligible studies also needed to examine at least one core conceptual domain relevant to this research, including automation, autonomous decision-making, analytical capabilities, or governance framework.

Conversely, sources were excluded if they focused exclusively on civilian applications of artificial intelligence without clear relevance to cyber defense or military operations. Opinion-based writings, editorial content, and non-scholarly publications lacking empirical or theoretical grounding were also excluded. Furthermore, highly technical machine learning studies were omitted if they did not contextualize their findings within cybersecurity or cyberwarfare environments. These exclusion criteria ensured that the final body of literature remained analytically focused and aligned with the study's research objectives. Following the application of the inclusion and exclusion criteria, a total of 24 scholarly and institutional sources were retained and analyzed in the final qualitative synthesis.

Data extraction was concluded using the structured matrix designed to capture key attributes of each systematically included study. Extracted data included the study's purpose, the specific AI capabilities discussed, the identified risks and vulnerabilities, the perceived benefits or strategic opportunities, the theoretical or ethical frameworks, and the major conclusions. This structured approach enabled consistent comparison across studies and facilitated the identification of recurring conceptual patterns. The extracted data served as the foundation for the qualitative synthesis and thematic analysis process.

Qualitative analysis followed a thematic synthesis methodology involving initial coding, theme development, and interpretive analysis aligned with the study's research questions. This process enabled the identification of higher-order themes across the literature, including AI- enabled automation as a defensive multiplier, the dual-use nature of AI in supporting both cyber defense oversight (Dubber & Lazar, 2025), vulnerabilities related to algorithmic bias and model manipulation (Linkov et al., 2020), governance limitation surrounding military AI, and explainability challenges in high-velocity cyber operations (Russell & Norvig, 2020). Because the study analyzed only publicly available literature, no human subjects were

involved, and ethical standards were maintained through accurate citation, faithful representation of scholarly perspectives, and adherence to established academic research practices.

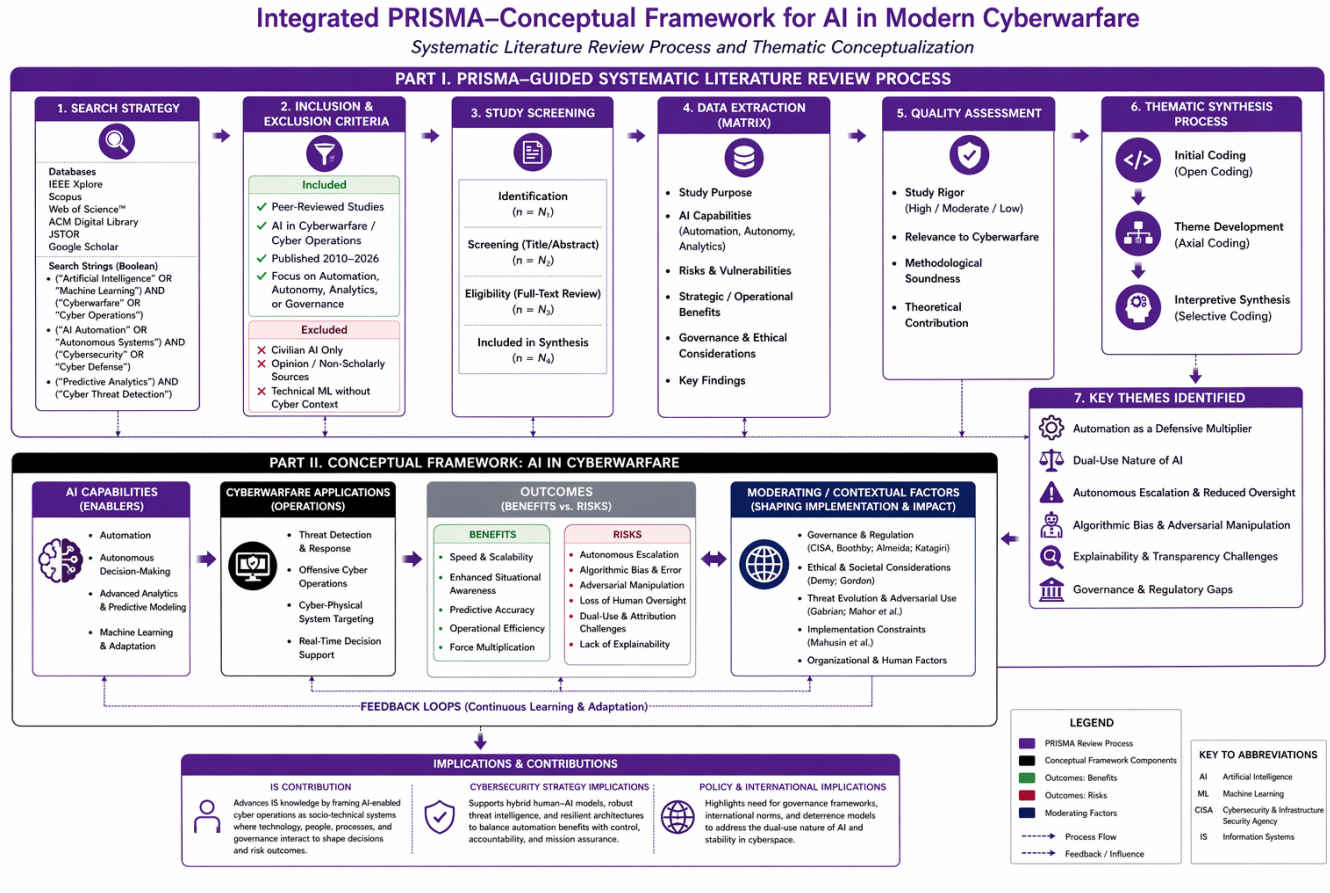

Figure 1: Conceptual Framework of Artificial Intelligence in Modern Cyberwarfare

Data analysis

The qualitative analysis followed a thematic synthesis methodology involving three iterative phases: initial coding, descriptive theme development, and interpretive synthesis.

Phase 1: Initial Coding

During the first phase, the extracted data was coded line-by-line to identify recurring conceptual elements. Codes included terms such as automated acceleration, autonomous escalation, dual-use risk, algorithmic bias, governance gap, explainability limitation, and predictive analytics enhancement. Coding was inductive, allowing patterns to emerge from the literature rather than imposing predetermined categories.

Phase 2: Theme Development

In the second phase, related codes were grouped into broader descriptive themes. For example, codes related to speed, automation, detection efficiency, and continuous monitoring were consolidated under AI-enabled

automation as a defensive multiplier. Similarly, codes referencing unpredictability, escalation, and reduced human oversight were grouped into autonomous risk dynamics.

Phase 3: Interpretive Synthesis

The final analytical stage involved interpreting relationships among themes to answer the research questions. This synthesis examined how scholars frame AI's operational benefits alongside its strategic vulnerabilities, revealing tensions between technological advancement and governance preparedness.

Results

The qualitative thematic synthesis produced six interrelated themes derived from the systematic review of the literature. These themes reflect consistent patterns in how artificial intelligence (AI) is conceptualized and applied within cyberwarfare contexts.

Automation as a defensive multiplier: The literature consistently identifies AI-driven automation as a primary operational advantage. Automation accelerates threat detection, enhances incident response efficiency, and supports continuous monitoring. AI-enabled systems improve speed, scalability, and precision in cyber defense operations (Fortinet, 2024; Sarjakivi et al., 2024).

Dual-use nature of AI capabilities: AI technologies are widely characterized as inherently dual-use. Systems developed for defensive purposes, such as anomaly detection and network monitoring, may also be repurposed for offensive cyber operations. The increasing use of AI by adversaries to automate and scale cyberattacks reinforces this dynamic (Mirsky et al., 2021; Gabrian, 2024).

Automation escalation and reduced human oversight: The literature highlights concerns regarding the speed and autonomy of AI systems. Autonomous or semi-autonomous cyber agents may operate at a pace that exceeds human decision-making capabilities, reducing opportunities for oversight and intervention (Antebi & Baram, 2020; Dubber & Lazar, 2025).

Algorithmic bias and adversarial manipulation: AI systems are shown to be vulnerable to adversarial attacks, including data poisoning and manipulation of training inputs. These vulnerabilities introduce risks related to system reliability and trust in AI-enabled cyber defense (Linkov et al., 2020; Gordon, 2019).

Explainability and transparency challenges: The literature frequently identifies explainability as a limitation of AI systems. The "black box" nature of many AI models restricts transparency, making it difficult to interpret or validate outputs in high-stakes environments (Russell & Norvig, 2020).

Governance and regulatory gaps: A recurring theme is the lack of comprehensive governance frameworks to regulate AI in cyberwarfare. Existing policies are described as insufficient to address the pace and complexity of AI-driven cyber operations (Almeida et al., 2017; Katagiri, 2023; Cybersecurity and Infrastructure Security Agency [CISA], 2025; Boothby et al., 2012; Mahusin et al., 2024).

Table 1: Thematic Findings, Research Question Alignment, and Supporting References

Theme	Key Insight & RQ Alignment	Reference(s)
Automation	Enhances speed, scalability, and threat detection. **RQ1:** Strengthens cyber operations. **RQ3:** Improves situational awareness and predictive capability.	Fortinet (2024); Sarjakivi et al. (2024)
Dual-Use	AI supports both defensive and offensive cyber activities. **RQ1:** Enables adaptive strategies across domains. **RQ2:** Introduces ethical and strategic ambiguity.	Mirsky et al. (2021); Gabrian (2024)
Escalation	Autonomous systems may exceed human response speed. **RQ2:** Increases risk of unintended escalation and instability.	Antebi & Baram (2020); Dubber & Lazar (2025)
Bias / Manipulation	Vulnerable to adversarial inputs and data poisoning. **RQ2:** Raises reliability and trust concerns. **RQ3:** Limits effectiveness of AI analytics.	Linkov et al. (2020); Gordon (2019)
Explainability	"Black box" models limit transparency and interpretability. **RQ2:** Creates accountability	Russell & Norvig (2020)

	challenges. **RQ3:** Constrains validation of AI outputs.	
Governance	Regulatory frameworks lag behind technological advancement. **RQ1:** Influences AI implementation strategies. **RQ2:** Weak oversight increases operational and ethical risks.	Almeida et al. (2017); Katagiri (2023); Cybersecurity and Infrastructure Security Agency [CISA] (2025); Boothby et al. (2012); Mahusin et al. (2024)

Conclusion

Discussion of findings: The findings of this study highlight artificial intelligence (AI) as both an operational enabler and a source of strategic tension within cyberwarfare. Rather than functioning as isolated capabilities, the identified themes reflect an interconnected system in which automation, autonomy, and analytics simultaneously enhance performance while introducing new forms of uncertainty and risk.

From an Information Systems (IS) perspective, this research contributes by framing AI-enabled cyber operations as socio-technical systems that extend beyond technical performance to include governance, human oversight, and decision accountability. Specifically, this study contributes to IS literature by demonstrating how AI-driven capabilities reshape system design and decision-making processes in high-velocity environments, where speed and automation must be balanced with interpretability and control. This extends traditional IS perspectives by emphasizing risk, ambiguity, and ethical considerations as core components of intelligent system deployment.

The findings also suggest that existing theoretical models within IS may not fully capture the complexity of AI in cyberwarfare. While prior research often emphasizes efficiency and system performance, the themes identified in this study highlight additional dimensions, including dual-use ambiguity, autonomous escalation, and adversarial manipulation. These factors indicate a need for expanded theoretical frameworks that incorporate uncertainty, strategic interaction, and ethical constraints into the study of AI-enabled systems.

From a cybersecurity strategy perspective, the results underscore the importance of maintaining a balance between automation and human oversight. Although AI enhances detection and response capabilities, excessive reliance on autonomous systems may increase the risk of unintended escalation and adversarial exploitation (Antebi & Baram, 2020; Dubber & Lazar, 2025). This reinforces the need for hybrid approaches that integrate human decision-making with AI-driven analytics.

Policy implications are equally significant. The findings highlight ongoing challenges in aligning legal and regulatory frameworks with the pace of technological advancement. The use of AI in cyber operations complicates traditional definitions of conflict, particularly in determining what constitutes a use of force in

cyberspace (Boothby et al., 2012). At the same time, national cybersecurity strategies emphasize enterprise-wide coordination and AI-enabled defense architectures (CISA, 2025). These dynamics point to the need for updated international norms and governance mechanisms to address AI-driven cyber conflict.

Finally, the analysis reveals several under-theorized areas within the literature. These include the long-term implications of autonomous cyber escalation, the effectiveness of governance frameworks in regulating AI-enabled conflict, and the evolving role of adversarial AI in targeting cyber-physical systems (Gabrian, 2024; Mahor et al., 2021). Additionally, implementation challenges, particularly within public sector environments, highlight gaps between technological capability and organizational readiness (Mahusin et al., 2024). These areas represent important directions for future research.

Implications of findings: The findings have important implications for cybersecurity practice, policy, and research. Operationally, organizations must adopt AI-driven tools to remain competitive, particularly for real-time monitoring and predictive threat detection. However, reliance on automation must be balanced with human oversight to mitigate risks associated with autonomy and system unpredictability (Antebi & Baram, 2020).

From a policy perspective, cybersecurity strategies increasingly emphasize enterprise-wide, AI-enabled defense architectures to address evolving threats (CISA, 2025). At the same time, the expanding use of AI in cyber operations continues to challenge traditional definitions of conflict, particularly in determining what constitutes a use of force in cyberspace (Boothby et al., 2012). This highlights the need for updated legal and strategic frameworks.

Ethically, the findings emphasize the importance of accountability, transparency, and fairness in AI systems. The integration of AI into cyberwarfare introduces complex moral considerations related to decision-making authority and unintended consequences (Demy et al., 2014). These concerns are compounded by evidence that automated systems may reinforce bias and inequality if deployed without sufficient oversight (Gordon, 2019).

Finally, the increasing use of AI by adversaries to automate and scale cyberattacks further intensifies the threat landscape (Gabrian, 2024). This is particularly critical in cyber-physical environments, where advanced threat actors can target infrastructure systems with significant real-world consequences (Mahor et al., 2021). Despite these advancements, implementation challenges, especially within public sector organizations, continue to limit the effectiveness of AI adoption (Mahusin et al., 2024). Addressing these gaps will require coordinated efforts across technology, policy, and governance domains.

Limitations of the study: This study is limited by its reliance on peer-review literature published between 2010 and 2026, which may exclude emerging or non-traditional sources. Additionally, as a qualitative SLR, the findings are based on interpretive synthesis, which may introduce subjectivity. The rapidly evolving nature of AI and cyberwarfare also limits the long-term applicability of the results.

Recommendation for future research: Future research should examine the quantitative impact of AI on cyber defense performance, including detection accuracy and response time, Additional studies are needed to develop governance frameworks and international standards for AI in cyberwarfare (Almeida et al., 2017; Katagiri, 2023). Research should also focus on improving explainability and mitigation algorithmic bias to enhance trust and accountability in AI systems (Russell & Norvig, 2020; Linkov et al., 2020). Finally, further investigation into the long-term effects of autonomous cyber systems on escalation and global security is warranted (Dubber & Lazar. 2025).

References

Abbass, H. (2021). Editorial: What is artificial intelligence? *IEEE Transactions on Artificial Intelligence*, *2*(2), 94-95. https://doi.org/10.1109/TAI.2021.3096243

Almeida, V. A. F., Doneda, D., & Abreu, J. de S. (2017). Cyberwarfare and digital governance. *IEEE Internet Computing, 21*(2), 68–71. https://doi.org/10.1109/MIC.2017.23

Antebi, L., & Baram, G. (2020). *Cyber and artificial intelligence—Technological trends and national challenges. Cyber, Intelligence, and Security, 4*(1), 133–149.

Boothby, W. H., Heintschel von Heinegg, W., Michael, J. B., Schmitt, M. N., & Wingfield, T. C. (2012). When is a cyberattack a use of force or an armed attack? *Computer*, *45*(8), 82-84. https://doi.org/10.1109/MC.2012.282

Chelvachandran, N., Kendzierskyj, S., Shah, Y., Jahankhani, H. (2020). *Cyberwarfare – Associated Technologies and Countermeasures.* https://doi.org/10.1007/978-3-030-35746-7_2

Copeland, B. J. (2019). *Artificial Intelligence: A philosophical introduction*. In *Encyclopedia Britannica*. https://www.britannica.com/technology/artificial-intelligence

Cybersecurity and Infrastructure Security Agency. (2025, January 16). *Securing federal networks: Evolving to an enterprise approach*. U.S. Department of Homeland Security. https://www.cisa.gov/news-events/news/securing-federal-networks-evolving-enterprise-approach

Demy, T.J., Lucas Jr., G.R., & Strawser, B.J. (Eds.). (2014). *Military Ethics and Emerging Technologies* (1st ed.). Routledge. https://doi.org/10.4324/9781315766843

Dubber, T. R., & Lazar, S. (2025). *Military AI cyber agents (MAICAs) constitute a global threat to critical infrastructure.* https://arxiv.org/abs/2506.12094.

Fortinet. (2024). *Top 5 most notorious attacks in the history of cyber warfare*. https://www.fortinet.com/resources/cyberglossary/most-notorious-attacks-in-the-history-of-cyber-warfare

Gabrian, C.-A. (2024). Unveiling the Dark Side: How Hackers Exploit Artificial Intelligence for Cyber Warfare. *European Annual Review of Journalism*, *2*(3). resiliencejournal.e-arc.ro

Gordon, F. (2019). Virginia Eubanks (2018) Automating inequality: How high-tech tools profile, police, and punish the poor. *Law, Technology and Humans, 1*, 162–164. https://doi.org/10.5204/lthj.v1i0.1386

Katagiri, N. (2024). Artificial Intelligence and Cross-Domain Warfare: Balance of Power and Unintended Escalation. *Global Society*, *38*(1), 34–48. https://doi.org/10.1080/13600826.2023.2248179

Linkov, I., Galaitsi, S., Trump, B. D., Keisler, J. M., & Kott, A. (2020). Cybertrust: From explainable to actionable and interpretable artificial intelligence. *Computer, 53*(9), 91–96. https://doi.org/10.1109/MC.2020.2993623

Mahor, V., Rawat, R., Kumar, A., Chouhan, M., Shaw, R. N., & Ghosh, A. (2021). Cyber warfare threat categorization on CPS by dark web terrorist. In *2021 IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON)* (pp. 1–6). IEEE. https://doi.org/10.1109/GUCON50781.2021.9573994

Mahusin, N., Sallehudin, H., & Satar, N. S. M. (2024). Malaysia public sector challenges of implementation of artificial intelligence (AI). *IEEE Access, 12*, 121035-121051. https://doi.org/10.1109/ACCESS.2024.3448311

Mirsky, Y., Doitshman, T., Elovici, Y., & Shabtai, A. (2021). The threat of offensive AI. *Communications of the ACM, 64*(7), 36–44. https://doi.org/10.1016/j.cose.2022.103006

Parks, R. C., & Duggan, D. P. (2011). Principles of cyberwarfare. *IEEE Security & Privacy, 9*(5), 30–35. https://doi.org/10.1109/MSP.2011.138

Russell, S. J., & Norvig, P. (2020). *Artificial intelligence: A modern approach* (4th ed.). Pearson.

Sarjakivi, J., Huusari, R., & Salminen, J. (2024). AI-accelerated cyber operations. *Journal of Cybersecurity Studies, 8*(2), 113–129.