Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

AI-Driven Cyber Risk Management in Adaptive Threat Environments

Mahee Ahmed Choudhury[1], Supom Roy[2]
[1]*Master's in Information Technology University of the Cumberlands, Williamsburg, KY, USA*
Email: mchoudhury78170@ucumberlands.edu
ORCID: 0009-0000-5003-7184
[2]*Bachelor of Science in Information Systems (BSIS) Trine University, Angola, IN, USA*
Email: sroy25@my.trine.edu
ORCID: 0009-0007-0384-4910

ARTICLE INFO

Received: 30 Dec 2024

Revised: 05 Feb 2025

Accepted: 25 Feb 2025

ABSTRACT

The fast integration of generative artificial intelligence (AI) into the work of cybersecurity has radically changed the face of cyber threat in modern companies. Attackers are now able to automate reconnaissance, create adaptive malware, and perform large-scale and individualized social engineering, and defenders are now increasingly relying on AI-based detention, automated response, and security coordination. The convergence has resulted in a rapidly changing adversarial environment where classical governance and risk management frameworks of cybersecurity are becoming less effective. This paper suggests a framework of engineering management of governing AI-enabled cyber risk, comprising adaptive risk evaluation, responsible AI governance, and strategic investments prioritization. The research paper uses a mixed-method design science methodology, which is based on the conceptual framework formulation, synthesis of the findings of the AI red-team and blue-team simulation studies conducted in previous research and qualitative synthesis of the reported individual industry cases. The framework describes how the management processes in engineering have to change as attackers and defenders switch to using generative AI, and introduces dynamical risk measures, multi-functional governance systems, and lifelong learning systems to aid adaptive defense. Findings can be used in the area of cybersecurity governance, engineering management scholarship, and offer guidance in practice that can be utilized by practitioners and policymakers to minimize AI-related cyber risk and maintain organizational resilience.

Keywords: reconnaissance; artificial intelligence (AI); resilience; management.

1. INTRODUCTION

The decision-making process behind addressing cybersecurity risk has been based on the notion that humans will take advantage of technical weaknesses and that organizations can mitigate these threats through a mixture of technical controls, organizational standards and human resources. Such an assumption becomes more and more questionable. The development of artificial intelligence (AI) and especially large language systems is something that has radically changed the magnitude and pace of cyber conflict (Lal and You, 2025). Previously, only the most specialized skills were needed to perform certain tasks such as vulnerability discovery, creation of exploits, phishing content generation and malware mutation which now can be automatized to some degree or even completely. Meanwhile, machine learning-based systems continue to be used by defenders to handle the increased amount and complexity of cyber threats, with AI being used to identify anomalies, correlate threat intelligence, and perform automated incident response (Kumar & Gutierrez, 2025).

This is a change in technology that has been coupled with a quantifiable increase in cyber risk. According to industry report, global economy will incur over USD 10.5 trillion every year by 2025 due to cybercrime, compared to USD 3 trillion in 2015, both because of increased brutal attacks. Recent research also shows that more than 60 percent of phishing efforts are currently using some kind of AI-generated content, and that they are achieving much higher

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

success rates because of personalization and language realism (Verizon, 2024). Moreover, AI-related malware is reported by security vendors to be detected many weeks on average than regular malware, highlighting the problems that AI presents to the existing malware detection systems (IBM Security, 2024).

The convergence of AI-driven offense and defense is not only a new step in the technology but a paradigm shift in the cyber risk environment. The integration of both offenders and security services on the basis of dynamically developing learning systems makes cyber risk dynamic, non-linear, and, by its nature, unpredictable (Shrestha et al., 2025). The traditional risk assessment models, which are usually founded on the static threat models, asset inventories, and regular reviews, are not well-positioned to reflect such dynamics. Consequently, organizations become increasingly unsure about the way to manage AI-enhanced security devices, the way to prioritize cybersecurity investments, and the way to apportion responsibility in case automated systems make or contribute to key security decisions.

Recent empirical studies show that one of the extent of this challenge of governance (Shrestha et al., 2025). Cyberattacks enabled by AI, such as deepfake-assisted fraud, smart phishing, smart ransomware, and so on, have risen to a new level in both rate and severity. As a case study, financial institutions and ransomware groups are reported to have risen by 300 percent in deepfake-related fraud attempts between 2022 and 2024 and more ransomware groups are using AI to further target and optimize their target and negotiation processes (Jabir et al., 2025). These trends have unveiled major shortcomings in prevailing cybersecurity governance protocols, specifically in automated systems administration, the clarity of automation-based decisions, and the correspondence between the security work and the company policy.

As a result, it is now the responsibility of engineering managers and security leaders to assume the role of safeguarding not only digital resources, but also managing multifaceted socio-technical systems that incorporate human judgment as well as autonomous machine intelligence. This increased accountability confronts the engineering management responsibilities, necessitating new AI governance, systems thinking, and adaptive risk management competencies.

To deal with this context, it is necessary to have frameworks that are able to combine technical risk management with governance structures, investment decision-making and organizational learning. Although the current cybersecurity frameworks are informative, they were not configured to match an adaptive and AI-driven attacker and defender. Likewise, developing AI governance efforts focus more on ethical values and risk management but do not tend to be more operationally specific to adversarial cybersecurity settings where constant pressure and fast decision-making are the order of the day (Dhirani et al., 2023).

The present paper fills this gap by delivering a methodical synthesis of cybersecurity research findings, literature on AI governance, and engineering management theory to create a consistent engineering management paradigm of AI-enabled cyber risk. Instead of making external suggestions on individual technical controls, the study addresses this gap through (1) conceptual re-examination of engineering management processes to adaptive, AI-driven attacker-defender dynamics based on design science research principles; (2) development of governance and accountability requirements by qualitatively synthesizing documented industry cases and regulatory analyses; and (3), informing adaptive risk measurement and investment decision-making by structurally analyzing AI red-team and blue-team simulation findings reported in previous studies. Based on this combined method, the study answers three research questions: First, what modifications will engineering management processes have to be when both attackers and defenders use generative AI? Second, what governance and accountability systems should be in place in order to deploy AI-enabled security systems in a responsible and effective way? Third, which risk measurement and assessment strategies are the most appropriate to make adaptive defense and informed investment decisions in AI-driven cybersecurity environments?

This study utilizes a mixed research method in order to answer these research questions. The suggested framework is constructed based on design science research, and Analysis of AI red-team and blue-team simulation studies is used to examine the dynamics between attackers and defenders in controlled settings. Besides this, the research of governance and accountability practices is studied with the aid of qualitative case analysis which is based on real-life organizational context. The combination of these approaches allows interdisciplinary inputs into cybersecurity

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

governance, artificial intelligence, and engineering management, as well as providing realistic advice to practitioners and policymakers in the face of AI-led cyber threats.

2. LITERATURE REVIEW AND THEORETICAL FOUNDATIONS

Over the last decade, more and more scholars and policymakers have been paying attention to the intersection of artificial intelligence and cybersecurity. Initial studies of machine learning in cybersecurity mainly addressed detection-driven applications, such as malware detection, spam filtering and intrusion detection systems. These papers thought of AI as a tool of decision support and enhanced the human analysts by enhancing pattern recognition, false positives, and response time. Rigorous overviews confirm that until recently, the majority of AI implementations in the field of cybersecurity were designed as analysis instruments integrated into human-directed processes and not as the autonomous entities (Kaur et al., 2023).

This has been greatly expanded with the advent of generative AI. Unlike previous machine-learning models, generative systems can autonomously generate content, code, and adaptive behaviors, as well as being used in classification and prediction. According to Salari et al., (2025) this change is referred to as changing assistive AI into agentic AI where systems cease to support human decision-making but instead acts independently. The implication of this shift goes beyond technical performance and impacts the organizational governance structures and methods of risk management. This feature allows attackers in a cybersecurity setting to automate reconnaissance, produce polymorphic malware, and create phishing messages of unprecedentedly personalized scale.

An accumulated empirical literature demonstrates how offensive the generative AI can be. Han et al. (2025) empirically demonstrate the ability of large language models to produce phishing emails with evasion rates that are over 90 percent, which is much higher than the ability of traditional phishing methods. In the same line, recent reports also confirm that AI-enabled vulnerability discovery tools are capable of helping attackers detect exploitable vulnerabilities quicker and with more coverage than manual methods, which shortens the attack lifecycle. All of these findings support the definition of AI as a dual-use technology, that is, one that can increase both defensive and malicious cyber capabilities.

On the defensive front, the adoption of AI has been faster due to the increased number and threats of cyber attacks that are complex. The systematic reviews discover that machine learning is widely used in detecting anomalies, intelligence of threats fusion, automation of incident response, and optimization of recovery throughout the NIST cycle of cybersecurity (Kaur et al., 2023). Advocates believe that work systems based on AI would be critical in handling security streams of data that are beyond the cognitive capacity of the human brain. Nevertheless, the empirical research is showing more and more limitations to them, such as susceptibility to adversarial manipulation, drift in the model, and overuse of opaque decision processes (Yazdi et al., 2024).

Organizational security practices continue to rely on conventional cybersecurity risk management systems, including the International Organization for Standardization/International Electrotechnical Commission Information Security Risk Management standard (ISO/IEC 27005) and the National Institute of Standards and Technology Risk Management Framework (NIST RMF). According to meta-analyses, these frameworks offer systematic guidelines in terms of identifying and assessing assets and selecting controls but tend to be very inert and compliance-driven (AL-Dosari and Fetais, 2023). Empirical assessments indicate that these frameworks find it challenging to admit the fast changing technologies, especially in the settings that have a learning system that is adaptable in nature. Risk in AI-enabled settings is a product of system interaction and not of known and discrete vulnerabilities, as the conventional frames assume, making the concept difficult to apply.

The engineering management theory can be used as a complementary lens in dealing with these limitations. The socio technical systems theory and systems engineering focus on integration, feedback and adaptive control when managing complex systems. Yazdi et al. (2024) prove that AI-enhanced risk management is accurate at the analytical stage and fast but at the same time poses a challenge on governance, such as explainability, accountability, and human control. These results indicate that technical optimization does not suffice but successful AI-driven cybersecurity implies alignment between engineering, executive, legal, and compliance spheres.

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

This need is also highlighted by recent scholarship of AI governance. According to Radanliev (2025), transparency, fairness, and accountability are among the principles that should be integrated throughout the AI lifecycle in order to reduce systemic risk, especially in high-stakes areas. Even though governance frameworks formulated by National Institute of Standards and Technology (NIST), the European Union (EU) and others provide an articulation of ethical and risk-based concepts, they tend to be abstract and not operationalized enough when faced with adversarial cybersecurity settings where fast and automated decision-making has become the normal.

In spite of the increased amount of research, a number of gaps do exist. First, current literature focuses mostly on AI applications on their own (either in the offensive or defensive), but does not concern the co-evolutionary nature that ensues as the attackers and defenders use adaptive AI systems. Second, existing risk measures do not reflect the amplification effects and feedback loops observed in empirical studies of cybersecurity issues (Kaur et al., 2023; Yazdi et al., 2024) which are driven by AI. Third, engineering management literature covers the topic of adoption of AI in the project and risk management but has not adequately been coupled with cybersecurity governance to align AI-enabled security systems to organizational strategy and risk appetite.

Based on these conclusions, this paper consolidates evidence in cybersecurity studies, AI governance and engineering management to suggest a framework that conceptualizes AI-enabled cybersecurity as a socio-technical system that is dynamic. The model focuses on unremitting supervision, dynamic risk measurements, and strategic investment choices, which are a direct response to the constraints found in previous literature.

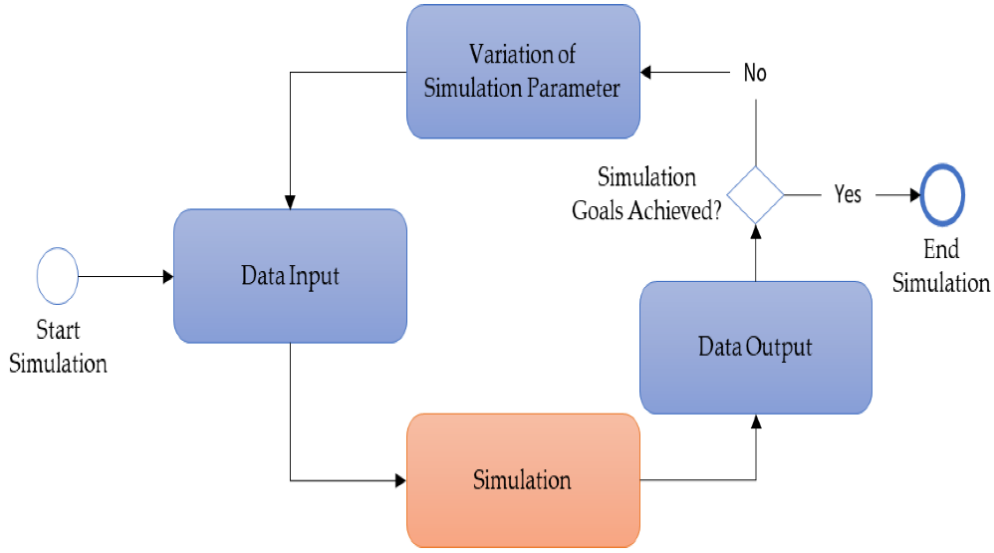
3. RESEARCH METHODOLOGY

The research adopts a conceptual mixed-methods research design. It is not primary empirical experimentation, live simulation or field study of organizations. Rather, it incorporates research in design science, organized study of simulation research outcomes published in the previous literature and qualitative synthesis of recorded industry instances. This method is suitable because AI-facilitated cyber risk governance is exploratory in nature and ethics and practical constraints of running live experiments in adversarial cybersecurity practice.

A methodological flow diagram illustrating the integration of design science, simulation, and case study analysis.

To explain research design and methodological integration, Figure 1 shows design science research, analysis of simulation-based results presented in previous research, and analysis of qualitative case studies. The simulation aspect constitutes a theoretical analysis of the simulation evidence of the existing studies and not actual implementation of simulations. This framework points to the repetitive manner in which theoretical understanding and empirical data were integrated to produce the proposed framework.

Figure 1. Methodological Flow Diagram



Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Note: Adopted from Kassen, S., Tammen, H., Zarte, M., & Pechmann, A. (2021). Concept and Case Study for a Generic Simulation as a Digital Shadow to Be Used for Production Optimisation. *Processes*, *9*(8), 1362. https://doi.org/10.3390/pr9081362

This paper uses a mixed-methods design of research to build and test an AI-enabled cyber risk engineering management framework. The mixed-method approach is suitable as AI-based cybersecurity is a complicated socio-technical phenomenon that cannot be analyzed with either quantitative or qualitative method on its own. In line with the abstract, the methodology involves design science research and analysis of AI red-team and blue-team simulation studies, and qualitative analysis of industry cases. The methodology is transparent and rigorous with each approach being clearly aligned with one of the three study research objectives.

Objective One: Engineering Management Process Changes in AI-Enabled Cyber Environments

The first objective aims at understanding how the engineering management processes will need to change when both the attackers and the defenders use the generative AI. The conceptual design science research (DSR) approach is used to achieve this objective; the systematic review and synthesis of the existing literature on engineering management, cybersecurity governance, and AI-enabled risk management were used as the supporting factors.

Systematic Review of Engineering Management and Cybersecurity Governance Literature

In order to base the analysis on the design science, the structured review of existing literature was performed with the focus on (1) conventional engineering management procedures in the domain of cybersecurity, (2) governance frameworks related to AI-driven systems, and (3) studies that explored adaptive and feedback-based risk environments. This review has reviewed peer-reviewed journal articles, standards-based systems, and empirical works that report the drawbacks of linear, compliance-based management of cybersecurity in the adaptive threat environment. The formation of this literature brought out common failures of the static risk assessment, periodic audits, and fixed control implementation as they are used in the context of AI-based adversarial.

Design Science Research Application

The work was based on the literature review to create an engineering management framework based on AI-enabled cyber risk, as a design science research (DSR). DSR is common in information systems and engineering management to design and evaluate artifacts that solve problems that are complex and real (Panakaduwa et al., 2025). The artefact that is created during the course of present research is a conceptual engineering management framework based on the combination of AI governance, adaptive risk assessment, and strategic investment decision-making.

DSR process had a systematic process:

1. **Problem identification**, based on documented failures of conventional cybersecurity management models in AI-enabled environments;
2. **Objective definition**, aligned with the need for adaptive, feedback-oriented management processes;
3. **Framework construction**, informed by systems engineering and socio-technical systems theory; and
4. **Iterative refinement**, guided by theoretical insights and synthesized empirical evidence reported in prior studies.

The traditional engineering management approaches, including the linear cycle of risk assessment, the static control deployment, and the infrequent review, were re-conceptualized and re-formed through this process. The framework that has been derived focuses on constant feedback, repetitive planning and formal human-AI cooperation, as the nature of AI-driven attacker defender dynamics is adaptive. This method shows how engineering management processes should be transformed to be no longer the classic control models, but dynamic and learning-oriented systems that can react to the fast dynamics of cyber risk situations.

Objective Two: Governance and Accountability Systems for AI-Enabled Cybersecurity

The second objective is to discover governance and accountability mechanisms that are needed to make the AI-driven cybersecurity tools responsible, transparent, and accountable. This goal is met by way of secondary qualitative

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

research or rather, a structured qualitative analysis of reported industry case studies based on previous empirical studies and regulating investigations.

Qualitative Analysis of Industry Case Studies on AI Security Governance

In an effort to respond to this goal, the study employed a qualitative review of reported industry cases that cut across the technology industry, financial services, and critical infrastructure sectors. Publicly-available sources, such as incident reports, regulatory enforcement activities, post-breach investigations, and peer-reviewed empirical research on AI-based security systems were used to derive these cases (Aslam et al., 2025). No primary interviews and proprietary organizational data was gathered, rather, secondary qualitative evidence was used in the analysis to provide methodological consistency and ethical viability.

The qualitative case study aimed at tracing the trends of governance failure and effectiveness in the organizations that implemented AI-based security technologies. Specifically, the following analysis was investigated:

(1) the manner in which the security decisions related to AI were delegated in organizations;

(2) the level of cross-functional integration among security operation, engineering, legal, compliance and executive leadership; and

(3) conformity of automated security measures to legal, ethical and regulatory standards.

Governance and Accountability Insights from Case Synthesis

The case-synthesis also showed that the application of AI-based security systems in an organization without formalized governance frameworks often had fragmented accountability, ambiguity in ownership of the behavior of AI systems, slow response to incidents, and increased regulatory risks. In a number of recorded situations, the lack of clearly defined responsibility when AI models performed their functions and made automated decisions complicated the investigated cases after the incidents and undermined the organizational trust in automated security controls.

On the other hand, those companies that adopted cross-functional AI governance models, including the coordinated management of security, engineering, legal, and compliance activities, had more explicit ways to escalate, better control over automated activities, and greater responsibility of AI-generated outcomes. The findings themselves were used to direct the governance and accountability aspects of the proposed engineering management framework, such as the suggestion of creating specific AI security governance bodies and clarify ownership of AI system performance, decision authority and failure management.

Objective Three: Risk Measurement and Assessment for Adaptive Defense and Investment Decisions

The third goal aims at recognizing risk measurement and assessment strategies that can be used to inform the adaptive defensive responses and investment choices among AI-driven cybersecurity settings. The aim is tackled with an analytical review of studies on AI red-team and blue-team simulations of the studies on cybersecurity literature published. Notably, the paper is not based on any primary simulations; rather, it is an attempt to synthesize the results of the already existing published simulation-based and experimental studies.

Analytical Review of AI Red-Team and Blue-Team Simulation Studies

To answer this goal, the research has examined peer-reviewed articles and experimental studies that utilize AI-mediated red-team and blue-team simulations to study adversarial cyber interactions. These works generally model AI-enabled offensive agents with the capability of automated phishing, intelligent malware execution, reconnaissance, and evasion methods as well as AI-enabled defensive agents with anomaly detection, automatic response, and system containment strategies (Zeijlemaker et al., 2025).

The simulation studies that were reviewed were chosen due to the availability of controlled experimental conditions in which attacker-defender interactions can be studied under different levels of automation, learning and adversarial adaptation. This can be done by such simulation-based research where the risk dynamics can be studied that would be either ethically, legally, or operationally impossible to replicate in actual organizational contexts.

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Risk Measurement Insights from Simulation-Based Evidence

The analysis of the findings of AI red-team and blue-team simulation studies shows that they have similar patterns that are applicable in cyber risk measurements. First, AI-based defensive systems are much faster in the detection and response time than other methods, including manual or rule-based. Second, these studies also show that with greater automation, model errors, adversarial manipulation, and data drift have more severe implications. In some of the reported cases, faulty models caused automated containment measures cascading through interdependent systems leading to the disruption of operations.

These lessons had a direct influence on the creation of dynamic risk measures within the engineering management system proposed. In particular, the analysis contributes to the introduction of the following metrics: model exploitability (susceptibility of AI systems to adversarial manipulation), automation amplification (to what extent AI systems amplify the magnitude or influence of actions), and latency reduction in response (to what extent AI systems can detect and contain actions faster). These measures combined will surpass risk scoring and allow adaptive evaluation in accordance with AI-driven cyber environments.

4. ANALYSIS AND RESULTS

This section shows the analysis undertaken to answer each research objective. Additionally the section provides the analytical procedures and the interpretation of the findings for the results presented here are based on the design science artifact development, AI red-team and blue-team simulation, and qualitative case analysis which are discussed in the Section 3.

4.1 Analysis

4.1.1 Objective One: Engineering Management Process Adaptation

To meet objective one, a conceptual analysis supported by a design science was conducted to look at the performance of current engineering management and cybersecurity management processes in an AI-enabled cyber environment. This research was not characterized by primary organizational experimentation or live operational test. It was rather based on a systematic review and synthesis of existing engineering management and cybersecurity risk management models reported in previous academic and industry literature, and which were conceptually assessed under adaptive threat conditions facilitated by generative artificial intelligence.

Engineering and Cybersecurity Management Models Reviewed

The review was devoted to commonly used and well-reported management frameworks that now command the field of cybersecurity in organizations. These were: the ISO/IEC 27005 Information Security Risk Management standard, which has a periodic risk identification and selection of controls; the National Institute of Standards and Technology Risk Management Framework (NIST RMF), which works through a lifecycle of sequential control selection, categorization, assessment, and authorization; the security management cycles of Plan–Do–Check–Act (PDCA), the standard cycles of security management integrated in line with the ISO-aligned governance systems; and the traditional Security Operations Center (SOC) process models, which used a human-led monitoring with rule- The choice of these models is based on them being the dominant engineering management logic applied to the management of cybersecurity risks and is widely mentioned in empirical research and standards-based literature (AL-Dosari and Fetais, 2023; Yazdi et al., 2024).

Design Science–Based Conceptual Evaluation

Based on a design science research (DSR) approach, these models received a conceptual evaluation within design cycles of comparing their fundamental premises, including the presence of static threats, periodic assessment, and decision-making processes highly dependent on human beings, with threat scenarios reported in AI-enabled cybersecurity literature. Reference scenarios were empirical research accounts of adaptive attackers, automated reconnaissance, and feedback-driven offensive operations that were used to test these management processes to stress (Chindrus and Caruntu, 2023; Uddin et al., 2025). This theoretical test explored the ability of linear and compliance-oriented management procedures to be effective with regards to the possibility of quick attacker evolution, ongoing model acquisition, and enhanced operational effects driven by automation.

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Identified Limitations and Process Adaptations

The review showed that the conventional linear control frameworks and periodic risk measurement are not responsive to the feedback-driven and dynamic character of AI-enabled threats. Specifically, the fixed review cycles and fixed control implementation implementation forms were demonstrated to be behind the evolution of the attacker and the drift in defensive models. The study found evidence of management process features related to increased resilience, such as continuous feedback, reiterative planning and reassessment, and formalised human-AI collaboration structures, through theoretical synthesis of cybersecurity governance research and engineering management theory. These adjustments are in tandem with socio-technical systems theory and principles of systems engineering that puts a strong focus on the learning, flexibility, and co-ordinated decision-making at technical and organizational levels. The lessons learned during this analysis were directly used to redesign the engineering management processes in the proposed framework.

4.1.2 Objective Two: Governance and Accountability Structures

The objective two was met by conducting a secondary qualitative analysis of recorded industry cases in previous empirical research, regulatory research, and post-incident research. The research was not based on any interviews or surveys, it was a systematic review and synthesis of publicly announced cases in the fields, where AI-enabled security systems are commonly implemented, such as technology companies, financial institutions, and operators of critical infrastructure. The cases were chosen as they directly analyzed the failures in governance, accountability failures, and responses of organizations linked to AI-based decisions in cybersecurity.

This is due to the cases reviewed that were documented cases of AI-based fraud detection, automated response to intrusions, and monitoring deployments as they were implemented without clear governance frameworks. As an illustration, numerous instances in the financial-sector as reported by regulators and overviewed by preceding empirical studies detail scenarios in which AI-based fraud detection systems automatically prevented transactions or user accounts without any explicit escalation processes or clearly defined owners. In such situations, the performance of models and decision outcomes were divided among security operations, IT teams, and third-party vendors of AI, which led to delayed incidents, damage to the customers, and greater regulatory attention (Aslam et al., 2025).

However, case evidence of the organizations that adapted cross-functional AI governance structures show significantly different results. These companies developed institutionalized control systems which incorporated security operations or functions, engineering forces, legal advises and compliance capabilities. The examples of such arrangements have been documented as having allowed to establish clearer responsibility on AI-driven decisions, predetermined escalation channels of automated actions, and better alignment of the technical operations with regulatory responsibilities. Consequently, these organizations demonstrated faster response to the incidents, better transparency of the AI decision-making process, and less exposure to the compliance breaches.

The comparative analysis of these industry cases indicates a regular tendency: the inability of formal control over AI results in fragmented responsibility, the inability to regulate operations, and the increased risk of regulatory consequences, and structured, cross-functional control increases the control of organizations over security systems based on AI. The results of these studies directly influenced the governance and accountability aspect of the proposed engineering management model, to which the recommendation of specific AI security governance entities and clearly defined AI system behavior, performance, and failure ownership belong.

4.1.3 Objective Three: Simulation-Based Risk Assessment

The objective Three was achieved by conducting a methodical analytical review of AI red-team and blue-team simulation literature collected in the previous literature of cybersecurity and adversarial machine learning. This research paper did not involve original simulations but synthesized the peer-reviewed simulation-based research studies that investigated the interactions between attackers and defenders in AI-on-cyber settings. This analysis concentrated on automation influences on detection latency, behavior of models under adversarial stress, and the flow of error on defensive behaviors acted upon autonomously.

A number of the reviewed red-team and blue-team simulation studies prove that AI-based defensive systems decrease the detection and response times by a significant margin compared to human-managed or rule-based security

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

operation. As an illustration, simulation-based experiments by Chindrus and Caruntu (2023) and Zeijlemaker et al. (2025) demonstrate that machine-learning-based intrusion detection and automated response systems always perform better than traditional systems in terms of detecting threats at the early stage and containing them. According to these studies, it is observed that a decrease in detection latency is achieved by constant inspection of behavior and real-time anomaly identification conducted by the AI-powered blue teams.

Nonetheless, another source of major systemic risk posed by automation is also indicated in the same simulation literature. There are several red-team/blue-team studies that record the cases when model errors, data poisoning, or adversarial manipulation results in improper automated containment, e.g. unequalled system isolation, service interruptions, or cascading operational failures. Specifically, Zeijlemaker et al. (2025) state that, in addition to providing better speed, AI-based defenders increase the severity of failure, as defensive decisions performed independently by the AI-based defender and without the appropriate level of human supervision result in even greater effects. In such simulations, when a defective model had caused the wrong response, the reverberations spread quickly through inter-connected systems, enhancing the impact of operational disturbance.

The articles consulted also show that the adaptation of attackers is a key factor in the dynamics. Simulations in red team demonstrate that AI-assisted attackers evolve in reacting to the behavior of a defender and take advantage of model drift, areas that are not well trained, and autonomy levels. This co-evolutionary dynamics leads to risk-profiles that do not persist over time, but transform in a continuous manner that compromises the effectiveness of risk scoring models and periodic assessment approaches that have been used historically in the field of cybersecurity governance.

Summarizing these simulation results, the analysis shows that the risk in AI-based cybersecurity settings cannot be fully modeled based on the standard likelihood-impact models. Instead, the literature justifies the necessity of dynamic risk measures which consider automation amplification, model exploitability and variability of response latency. These lessons were directly applied to the creation of adaptive risk measures under the proposed system, such as the dynamic to indicate how AI boosts defensive performance and failure spreading. Besides, the evidence of the simulation highlights the significance of including these measures in prioritizing investments because excessive investment in automation without the actual governance and resilience mechanisms enhances systemic risk.

Clarifying the Link Between Objectives, Methods, and Framework Components

The three study aims, methodology and structure components are not independent variables. Every of the objectives is incorporated using a certain approach to analysis and the resultant findings are converted to the respective aspects of the engineering management system. The first objective is covered with the help of design science-based conceptual analysis of engineering management processes, which leads to identifying AI-aware risk identification and adaptive management practices. Objective Two is fulfilled with the help of qualitative synthesis of industry cases, which lead to governance and accountability structures. The objective Three will be fulfilled by analytical analysis of the AI red-team and blue-team simulation studies, which will provide adaptive risk measures and investment prioritization systems. The elements found in the framework section are a product of the analysis as opposed to an approach, thus there is a clear distinction between methodology, analysis, and synthesis.

4.2 Results

4.2.1 Objective One

Objective one results are gained by conducting a comparative synthesis of the engineering management and cybersecurity governance literature, which will assess the effectiveness of conventional management processes in AI-enabled threat scenarios. Several articles that have surveyed existing cybersecurity management frameworks, including ISO/IEC 27005, the National Institute of Standards and Technology Risk Management Framework (NIST RMF), and security governance models based on the Plan-Do-Check-Act (PDCA) framework, all report that these models were originally developed in relatively stable threat environments and a periodic assessment cycle, along with most decisions being made by humans (AL-Dosari and Fetais, 2023; Yazdi et al., 2024). These frameworks underline the workflow, fixed control deployment, and planned audits that empirical and theoretical research has found to be more and more incompatible with the adaptive and automated character of AI-driven cyber menace.

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Literature on cybersecurity and engineering management recently shows that the use of generative AI radically changes how an attacker conducts themselves by allowing quick adaptation, continuous probing, and automated exploitation to shorten attack lifecycles and surpass defensives of the same type (Uddin et al., 2025; Shrestha et al., 2025). The research on AI-powered defensive systems also mentions that with the speed, automation introduces reliance on continuous model training and situational awareness that cannot be managed effectively by the traditional linear management processes (Chindrus and Caruntu, 2023). Consequently, the risk assessment performed at regular time intervals and the controls put in place do not keep up with AI-powered environments, and soon become obsolete.

By combining the results of these studies, it can be stated that traditional linear processes of cybersecurity management cannot be effective in AI-driven cyber settings. The literature reviewed supports consistently the use of engineering management strategies that include continuous feedback loop, iterative planning cycles, and coordinated human-AI work as more resilient in situations representing adaptive threat. According to systems engineering and socio-technical studies, these processes enable organizations to dynamically react to the development of attackers, reduce risks brought about by automation, and keep managers under control of AI-based security-related decisions (Yazdi et al., 2024; Kilian, 2025). These results lead directly to the conclusion that the engineering of management processes needs to shift to not being based on a model that is more driven by the compliance factor to the more flexible feedback-based management framework to stay relevant in the context of AI-led cybersecurity.

4.2.2 Objective Two

The findings of Objective Two are based on a qualitative survey of reported cases in the industry and related empirical research in governance on how organizations use AI-enabled security systems. The existing studies that examine the instances of cybersecurity in financial services, technology companies, and critical infrastructure show that any organization implementing AI-based security solutions does not have a formal governance structure and thus faces a disjointed form of accountability and a lack of ownership of the decisions (Aslam et al., 2025; Malgieri and Pasquale, 2024). Regulatory investigations and after incident reports mentioned in these papers indicate that the accountability behind AI-driven decisions is frequently shared between security operations, IT departments, data science teams, and external vendors, which causes ambiguity when automated actions cause a service disruption or harm to customers or violations of the compliance.

Some of the cases reviewed point out the significance of lack of clearly defined governance structures in the time lag in responding to an incident and regulatory exposure. To highlight, recorded cases of AI-based fraud detection and automated account controls indicate that companies without predefined escalation routes had problems explaining or reversing automated decisions, which caused extended downtimes and increased attention of the regulators (Aslam et al., 2025). These results indicate that ungoverned automation increases the risk in the organization instead of reducing it.

On the other hand, literature also offers data of institutions, which enacted cross-functional AI governance structures with more positive results. Empirical research on cybersecurity governance reported case studies which reveal that upon introducing security operations, engineering, legal, and compliance functions into formal oversight entities, companies have become more transparent in using the AI in automated decision-making, faster escalation of AI-related incidents, and more coherent accountability (Malgieri and Pasquale, 2024). These organizations were in a better position to match the AI-enabled security operations with the legal and ethical standards and, as a result, minimize the compliance risk and boost the confidence of the managers in the AI-related decisions.

Collectively, the analyzed literature confirms the conclusion that there is a need to have formal governance and accountability frameworks in place in case of responsible deployment of AI-enabled cybersecurity systems. The performed qualitative synthesis shows that cross-functional governance provisions result in a much better oversight, escalation transparency, and belief in automated security choices, directly reflecting on the governance and accountability aspect of the proposed engineering management framework.

4.2.3 Objective Three

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Objective three results are based on a comparative analytical survey of various published AI red-team and blue-team simulation studies of the work on the performance and failure mode of AI-enabled defensive systems in adversarial conditions. The selected studies are relevant to risk measurement in AI-driven cybersecurity settings as they explicitly simulate the attacker-defender dynamics, effects of automation, and adaptive behavior.

Multiple simulation-based studies are a consistent report that AI-enabled defensive systems can greatly lessen detection and response latency when compared to traditional security operations, which are conducted by humans. As a case in point, Chindrus and Caruntu (2023) state that blue teams that run on machine-learning in simulated cyber competitions reacted significantly faster to the threat and contained it than rule-based or manual defenders. On the same note, Kumar and Gutierrez (2025) simulate intrusion detection on an AI-based system and prove that the mean time to detection decreases because this system constantly processes network traffic of high volume through the analysis of real-time traffic. These findings are further reinforced by Zeijlemaker et al. (2025) who demonstrate that AI-assisted defensive agencies do better than traditional systems at early-stage attack detection in various simulated attack scenarios.

Nevertheless, systemic risks created by automation, especially in cases where the defensive decisions are implemented automatically, are also reported in the same body of simulation literature. According to Zeijlemaker et al. (2025), model drift and adversarial manipulation also often resulted in AI-based defenders labeling neutral behavior as malicious and taking automated containment measures, e.g. isolating the network or deactivating the service. Once these automated responses were triggered, they spread quickly between linked systems and led to a series of operational failures. Other publications such as Galaz et al. (2021) report similar findings and indicate that in highly interconnected digital infrastructures, automation increases the effects of local weaknesses in models, which are turned into systemic risks.

These conclusions are further supported by additional simulation works on adversarial machine learning. The simulation by Javadpour et al. (2025) based on reinforcement learning shows that AI-assisted attackers learn fast and exploit the behavior of defensive models that rely on fixed thresholds and predictable responses in spite of being trained on constant thresholds and predictable defensive models. Valdez et al. (2025) demonstrate that red-team simulations with large language models have the ability to revise attack plans in response to defender actions, and this revision will cause a probability increase of defensive model misclassification in the long term. All of these studies demonstrate that although AI has lower latency, it also causes amplification of failures in that once mistakes happen they tend to be more rapid and severe than when they are caused by humans.

Combining the outcomes of these simulation studies, the outcomes show that the static risk measures and periodic testing do not apply to AI-enhanced cybersecurity contexts. It has been revealed across the literature that risk is dynamically changing as models are trained, attackers evolve and automated responses spread through systems. This fact explains why dynamic risk measures are necessary to reflect the reduction of detection latency, amplified automation, and exploitable model. Moreover, according to the reviewed works, the investment decision in AI-enabled security should consider both the gains in performance and the exposure to systemic risks, which proves the value of focusing on flexible, resilient, and governance-consistent security investments.

5. DISCUSSION OF FINDINGS

The discussion evaluates the findings of Section 4 and discusses their implications to engineering management theory, cybersecurity governance, and organization practice. The findings obtained through the framework development of design science, simulation-based study findings analysis provided in earlier research, and qualitative examination of published industry cases all tend to suggest that traditional cybersecurity governing models are not applicable to the realm of AI-enabled environments. The published red-team and blue-team simulation results demonstrate that organizations with intense defensive automation and low quality of governance structure initially show higher rates of detection, but with time go through a decreasing rate of effectiveness, as a result of model drift, adversarial adaptation, and loss of human control (Chindrus & Caruntu, 2023).

The simulation-based results reported in the previous studies undergo an analysis presenting the amplification effect of artificial intelligence in both cyber offense and defense. Although AI-based security tools can greatly decrease the

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

latency of responses and the workload of the operations, they also increase the effects of errors in models. Weakly-defined models in documented simulation scenarios have triggered automatic containment responses, resulting in a series of debilitating operational spillover consequences, the significance of which cannot be overstated by the accountability mechanism and well-defined decision conditions. These results support the opinion that AI-enhanced cybersecurity needs to be regulated as a socio-technical system instead of being considered as a technical optimization (Galaz et al., 2021).

A qualitative review of reported cases in the industry also suggests that most organizations implement AI-based security solutions in a haphazard fashion due to vendor promises or competition and not necessarily in accordance with organizational risk aversion and governance ability. The issue of fragmented ownership of AI systems is that it is frequently shared among security operations, IT, and data science teams, leaving it unclear who should be accountable to the model performance and failure. Such fragmentation weakens the trust of automated systems and makes responding to incidents very difficult in a setting that involves regulatory or legal oversight (Nie et al., 2025).

The engineering management framework that is proposed solves these issues by clearly incorporating governance frameworks, adaptive risk measures, and continuous learning into the cybersecurity management activities. The results imply that this combined solution can make organizations more resilient to adaptive AI-powered actors and retain the essential value of human judgment by means of institutional human–AI interaction enabled by control and feedback systems.

6. AI-Enabled Cyber Threat Landscape

The section gives contextual analysis to the objective one (process adaptation) and objective three (risk assessment). It does not present new approaches but summarizes the existing literature on empirical cybersecurity to define the operating environment within which engineering management decision-making has to be done.

The emergence of generative artificial intelligence has changed the cyber threat landscape, as attackers have been able to change their capabilities and strategies. Such attacks are automated, scaled, and adaptable unlike traditional cyber threats which depended on manual methods and non-adaptive tools. Generative models also allow one to create attacker content quickly and dynamically change the attacker vectors, as well as counteract any defensive measures or even adapt to them, raising the frequency and their complexity (Uddin et al., 2025).

Among the greatest consequences of the emergence of generative AI is the emergence of automated and personalized social engineering attacks. Big language models enable attackers to create phishing emails and deepfake content that are highly similar to valid messages and take advantage of cognitive biases as well as substantially raise the chances of success (Valdez et al., 2025).

This menace is further increased by AI-assisted malware development. The machine learning methods aid in vulnerability detection, optimization of exploits, and generation of polymorphic malware capable of surviving the conventional signature-based detection. Together with reinforcement learning, malicious code may alter its behavior to adapt to the environmental stimuli, which makes the code more difficult to detect (Javadpour et al., 2025).

Offensively, organizations use AI-based tools of intrusion detection, endpoint protection, and security orchestration to control the magnitude and speed of threats. Although they increase the speed and scale, they come with new risks, such as the model drift, false positives, and vulnerability to adversarial manipulation (Babar, 2025).

This defender-attacker relationship generates a co-evolutionary relationship, where cyber risk is created through ongoing interaction of the systems and not individual vulnerability.

Table 1: AI-Enabled Cyber Capabilities and Associated Risks

Domain	AI Capability	Description	Key Risk
Offensive	Generative phishing	AI-generated, personalized messages	Increased fraud and credential theft
Offensive	Polymorphic malware	Self-modifying malicious code	Evasion of traditional detection

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

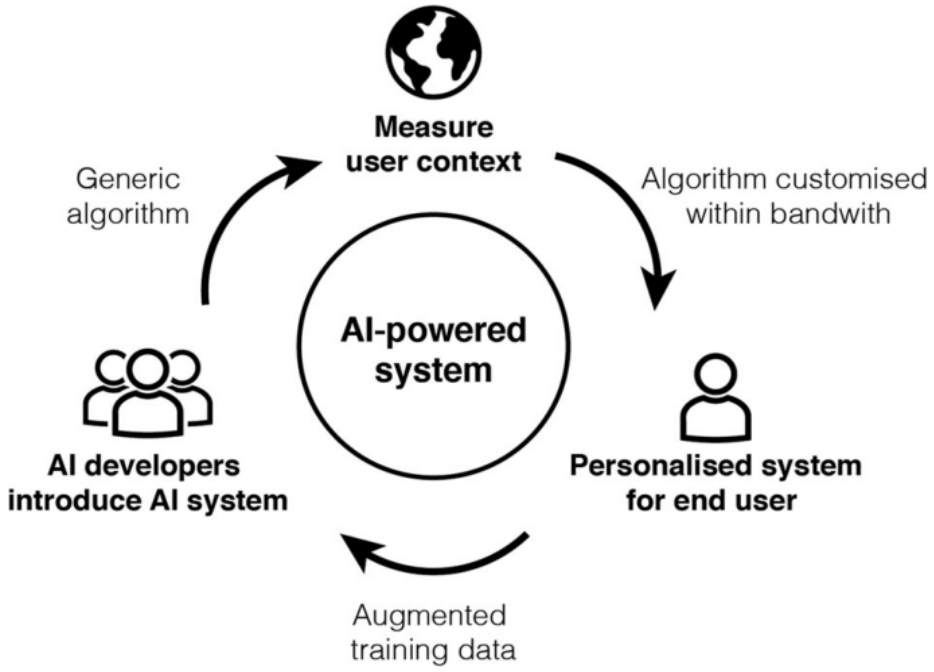
Defensive	AI intrusion detection	Behavioral anomaly detection	False positives, model drift
Defensive	Autonomous response	Automated containment actions	Cascading failures, accountability gaps

An engineering management perspective suggests that the AI-enhanced threat environment requires adoption of a shift towards adaptive defense in lieu of reactive security practices with governance systems that incorporate uncertainty and change. In the absence of such adaptation, organizations are in danger of implementing AI tools, which offer short-term pay-outs but introduce long-term vulnerabilities in systems. This dynamic environment is the basis of the engineering management structure suggested in the next section.

A conceptual diagram illustrating the co-evolution of AI-enabled attackers and defenders

In order to demonstrate how AI-enhanced attackers and defenders can interact dynamically, Figure 2 shows how AI-driven cyber systems can evolve together. The diagram illustrates the way in which AI systems analyze contextual data, adjust algorithms and produce customized outputs which are sent back to the model training. This process of cycles reminds of adaptive feedback loops, which indicates that the concept of cyber risk is an outcome of continuous interaction between systems, and the conventional approach to risk assessment is not sufficient in the AI-infused system of cybersecurity.

Figure 2. Co-evolution of AI-Enabled Attackers and Defenders



Note: Adopted from van Berkel, N., Tag, B., Goncalves, J., & Hosio, S. (2020). Human-centred artificial intelligence: a contextual morality perspective. *Behaviour & Information Technology*, *41*(3), 1–17. https://doi.org/10.1080/0144929x.2020.1818828

7. Engineering Management Framework for AI-Enabled Cyber Risk (Synthesis of Objectives One, Two, and Three)

This section gives the suggested engineering management structure as a combination of the findings in Section 4. It is a conceptual and managerial artifact that is the result of the design science research, which combined the results of the analysis of engineering management processes, governance structures, and AI-driven risk dynamics.

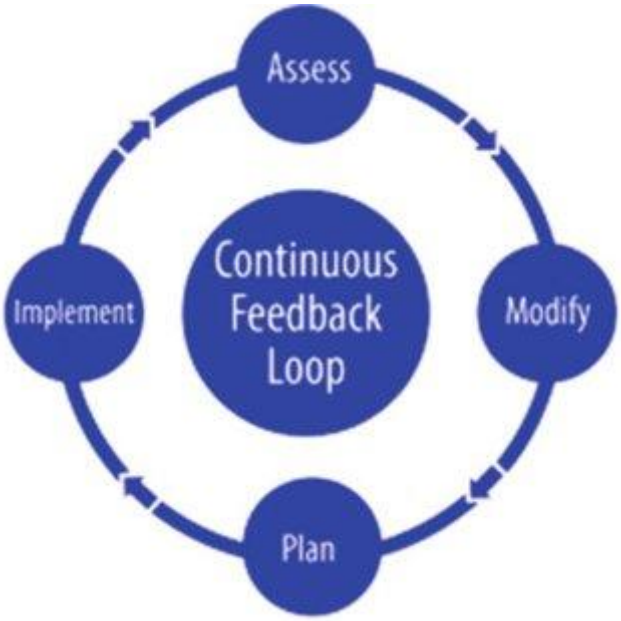
Engineering Management Framework for AI-Enabled Cyber Risk

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Based on the methodological and conceptual framework developed in the previous section, Figure 3 demonstrates the suggested engineering management framework of AI-enabled cyber risk. The model is represented as an ever-moving feedback system that includes planning, implementation, evaluation, and adjustment. Repeated feedback in the center uncovers the need to have a continual learning and adaptation process and that successful AI-based cybersecurity governance will be based on adaptive, iterative management over linear or compliance-based management.

Figure 3. Engineering Management Framework for AI-Enabled Cyber Risk



Note: Adopted from Bibri, S. E. (2019, May 17). *Data–driven Smart Sustainable Urbanism: Intelligence Functions, Simulation Models, and Complexity Sciences*. https://www.researchgate.net/publication/333176031_Data-driven_Smart_Sustainable_Urbanism_Intelligence_Functions_Simulation_Models_and_Complexity_Sciences

Table 2 Components of the Engineering Management Framework

Component	Primary Objective
AI-aware risk identification	Recognize AI systems as critical risk-bearing assets
Adaptive risk metrics	Capture dynamic, AI-driven changes in risk
Governance and accountability	Ensure responsible and transparent AI use
Investment portfolio management	Balance automation benefits with systemic risk
Continuous learning	Sustain adaptive defense over time

The framework is capable of helping to govern AI-enabled cyber risk by combining these components to offer a structured, flexible method of governance. It goes beyond being tool-focused in security management and establishes engineering managers as guardians of complex adaptive systems. This way, it fills a dire void in current cybersecurity and AI governance literatures and preconditions more robust and responsible cyber defense practices.

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

8. Implications for Engineering Management and Policy

The implications of this research on engineering managers are that they ought to broaden their historical conceptualizations of cybersecurity management to include AI governance and organizational learning. An AIs-enabled cyber risk needs to be managed with more skills than technical skills, such as the ability to think systemically, make moral decisions, and coordinate across functions. The engineers in management have to take the role of ensuring that AI security implementations are aligned with organizational strategy, risk tolerance and regulatory requirements.

On the policy side, the suggested framework can provide practical information to standards organizations and funding groups, including the National Science Foundation, the NIST, the Department of Homeland Security, and the Department of Defense. Since the framework focuses on adaptive risk measures and governance frameworks, it can be used to inform the creation of operational guidance and evaluation standards on AI-enabled security systems. The policymakers can also use the framework to facilitate the workforce development programs that equip security professionals with the responsibility to handle AI-driven systems.

On a larger scale, the research proposes that regulation or technical standards would not be sufficient to establish effective AI governance in the domain of cybersecurity. It must be inculcated into the engineering management practice and it must be institutionally devoted to transparency, accountability and continuous improvement.

9. Limitations and Future Research

This research has a number of limitations that indicate how future research should be conducted. First, although the analysis is based on the findings in previous studies, which are based on simulations, it does not imply the implementation of primary simulations. Consequently, the analysis is unable to represent complexity of the organization, operational factors, and socio-technical dynamics of the actual cybersecurity setting. The proposed framework must be empirically tested by means of longitudinal field studies and pilot deployments in the future.

Second, the qualitative case analysis is based solely on the publicly available sources such as the industry reports and documented cases. The method of collecting data can be biased with the reporting and poor access to the inner decision-making processes, which can undermine the picture of the governance failures. More comprehensive study of internal governance practices would be possible using access to proprietary organizational data, interviews with security and engineering managers, and mixed qualitative-quantitative designs.

Lastly, the research is mainly based on the enterprise and critical infrastructure settings. Future study ought to investigate a sector-specific modification of the framework especially in healthcare, education, and small-to-medium enterprises, where there is a significant variation in resource constraints and regulatory pressures and where it might be necessary to have customized governance and risk management systems.

10. Conclusion

The fast adoption of generative artificial intelligence into both cyber offense and defense has fundamentally changed the nature of cyber risk. The traditional systems of cybersecurity governance and risk management are failing to keep up with the advancements of attackers and defenders in the use of AI to automate, scale, and adapt operations. In this paper, it is proposed that AI-enabled cyber risk management is not solely a technical issue but an underlying engineering management challenge that needs to be integrated through governance, adaptive evaluation, and strategic supervision.

This study offers a systematic but open-ended model of governing AI-enabled cybersecurity systems by suggesting an engineering management framework based on the design science, simulation, and case analysis. The framework focuses on the AI-conscious risk discovery, dynamic metrics, responsibility governance, strategic investments management, and sustaining learning as the key attributes of adaptive defense. All of these factors allow organizations to take advantage of AI and reduce its systemic risks at the same time.

With the invention of the AI technologies, the organizations that do not alter the management and governance practices are at the risk of falling behind both technologically and strategically. On the contrary, those

that adopt integrated, adaptive strategies towards AI-enabled cyber risk will be in a stronger position to build resilience and safeguard vital assets and potentially continue to trust the digital context that is increasingly automated.

References

[1] Al-Busaidi, A. S., Raman, R., Hughes, L., Albashrawi, M. A., Malik, T., Dwivedi, Y. K., Al- Alawi, T., AlRizeiqi, M., Davies, G., Fenwick, M., Gupta, P., Gurpur, S., Hooda, A., Jurcys, P., Lim, D., Lucchi, N., Misra, T., Raman, R., Shirish, A., & Walton, P. (2024). Redefining boundaries in innovation and knowledge domains: Investigating the impact of generative artificial intelligence on copyright and intellectual property rights. *Journal of Innovation & Knowledge*, *9*(4), 100630. https://doi.org/10.1016/j.jik.2024.100630

[2] AL-Dosari, K., & Fetais, N. (2023). Risk-Management framework and information-security systems for small and medium enterprises (smes): A meta-analysis approach. *Electronics*, *12*(17), 3629. https://doi.org/10.3390/electronics12173629

[3] Aslam, M. M., Tufail, A., Gul, H., Irshad, M. N., & Namoun, A. (2025). Artificial intelligence for secure and sustainable industrial control systems - A Survey of challenges and solutions. *Artificial Intelligence Review*, *58*(11). https://doi.org/10.1007/s10462-025-11320-9

[4] Babar, M. (2025). A hybrid approach to financial big data analysis using extended ensemble learning and optimized spark streaming. *Journal of Open Innovation: Technology, Market, and Complexity*, *11*(3), 100602. https://doi.org/10.1016/j.joitmc.2025.100602

[5] Bibri, S. E. (2019, May 17). *Data–driven Smart Sustainable Urbanism: Intelligence Functions, Simulation Models, and Complexity Sciences*. https://www.researchgate.net/publication/333176031_Data-driven_Smart_Sustainable_Urbanism_Intelligence_Functions_Simulation_Models_and_Complexity_Sciences

[6] Chindrus, C., & Caruntu, C.-F. (2023). Securing the Network: A Red and Blue Cybersecurity Competition Case Study. *Information*, *14*(11), 587. https://doi.org/10.3390/info14110587

[7] Dhirani, L. L., Mukhtiar, N., Chowdhry, B. S., & Newe, T. (2023). Ethical Dilemmas and Privacy Issues in Emerging Technologies: A Review. *Sensors*, *23*(3), 1151. https://doi.org/10.3390/s23031151

[8] Galaz, V., Centeno, M. A., Callahan, P. W., Causevic, A., Patterson, T., Brass, I., Baum, S., Farber, D., Fischer, J., Garcia, D., McPhearson, T., Jimenez, D., King, B., Larcey, P., & Levy, K. (2021). Artificial intelligence, systemic risks, and sustainability. *Technology in Society*, *67*(101741), 101741. https://doi.org/10.1016/j.techsoc.2021.101741

[9] Han, W., Zhu, J., Zhang, C., Zhang, Z., Mei, Y., & Wang, L. (2025). Phish-Master: Leveraging Large Language Models for Advanced Phishing Email Generation and Detection. *Applied Sciences*, *15*(22), 12203. https://doi.org/10.3390/app152212203

[10] Hughes, L., Mavi, R. K., Aghajani, M., Fitzpatrick, K., Gunaratnege, S. M., Shekarabi, S. A. H., Hughes, R., Khanfar, A., Khatavakhotan, A., Mavi, N. K., Li, K., Mahmoud, M., Malik, T., Mutasa, S., Nafar, F., Yates, R., Alahmad, R., Jeon, I., & Dwivedi, Y. K. (2025). Impact of artificial intelligence on project management (PM): Multi-expert perspectives on advancing knowledge and driving innovation toward PM2030. *Journal of Innovation & Knowledge*, *10*(5), 100772. https://doi.org/10.1016/j.jik.2025.100772

[11] Jabir, R., Le, J., & Nguyen, C. (2025). Phishing Attacks in the Age of Generative Artificial Intelligence: A Systematic Review of Human Factors. *AI*, *6*(8), 174–174. https://doi.org/10.3390/ai6080174

[12] Javadpour, A., Ja'fari, F., Taleb, T., & Benzaïd, C. (2025). Detecting malicious nodes using game theory and reinforcement learning in software-defined networks. *International Journal of Information Security*, *24*(3). https://doi.org/10.1007/s10207-025-01026-y

[13] Kassen, S., Tammen, H., Zarte, M., & Pechmann, A. (2021). Concept and Case Study for a Generic Simulation as a Digital Shadow to Be Used for Production Optimisation. *Processes*, *9*(8), 1362. https://doi.org/10.3390/pr9081362

[14] Kaur, R., Gabrijelčič, D., & Klobučar, T. (2023). Artificial intelligence for cybersecurity: Literature review and future research directions. *Information Fusion*, *97*(101804), 1–29. ScienceDirect. https://doi.org/10.1016/j.inffus.2023.101804

[15] Kilian, K. A. (2025). Beyond accidents and misuse: decoding the structural risk dynamics of artificial intelligence. *AI & Society*. https://doi.org/10.1007/s00146-025-02419-2
[16] Kumar, A., & Gutierrez, J. A. (2025). Impact of Machine Learning on Intrusion Detection Systems for the Protection of Critical Infrastructure. *Information*, *16*(7), 515. https://doi.org/10.3390/info16070515
[17] Lal, A., & You, F. (2025). Advances and challenges in energy and climate alignment of AI infrastructure expansion. *Advances in Applied Energy*, *20*, 100243. https://doi.org/10.1016/j.adapen.2025.100243
[18] Malgieri, G., & Pasquale, F. (2024). Licensing high-risk artificial intelligence: Toward ex ante justification for a disruptive technology. *Computer Law & Security Review*, *52*, 105899. https://doi.org/10.1016/j.clsr.2023.105899
[19] Mohamed, N. (2025). Artificial intelligence and machine learning in cybersecurity: a deep dive into state-of-the-art techniques and future paradigms. *Knowledge and Information Systems*, *67*, 6969–7055. https://doi.org/10.1007/s10115-025-02429-y
[20] Nie, T., Sun, J., & Ma, W. (2025). *Exploring the roles of large language models in reshaping transportation systems: A survey, framework, and roadmap*. *1*(1), 100003. https://doi.org/10.1016/j.ait.2025.100003
[21] Panakaduwa, C., Coates, P., Munir, M., & Samansiri, S. (2025). Examining the Philosophical Underpinnings of Design Science Research (DSR). *Philosophies*, *10*(6), 139. https://doi.org/10.3390/philosophies10060139
[22] Radanliev, P. (2025). AI Ethics: Integrating Transparency, Fairness, and Privacy in AI Development. *Applied Artificial Intelligence*, *39*(1). https://doi.org/10.1080/08839514.2025.2463722
[23] Raina, K., Sharma, G. D., Taheri, B., Dev, D., & Chavriya, S. (2025). Artificial intelligence-driven management: Bridging innovation, knowledge creation, and sustainable business practices. *Journal of Innovation & Knowledge*, *11*, 100860. https://doi.org/10.1016/j.jik.2025.100860
[24] Salari, N., Beiromvand, M., Hosseinian-Far, A., Habibi, J., Babajani, F., & Mohammadi, M. (2025). Impacts of generative artificial intelligence on the future of labor market: A systematic review. *Computers in Human Behavior Reports*, *18*(100652), 100652. https://doi.org/10.1016/j.chbr.2025.100652
[25] Shrestha, S., Banda, C., Mishra, A. K., Djebbar, F., & Puthal, D. (2025). Investigation of Cybersecurity Bottlenecks of AI Agents in Industrial Automation. *Computers*, *14*(11), 456. https://doi.org/10.3390/computers14110456
[26] Tian, K., Zhu, Z., Mbachu, J., Ghanbaripour, A., & Moorhead, M. (2025). Artificial intelligence in risk management within the realm of construction projects: A bibliometric analysis and systematic literature review. *Journal of Innovation & Knowledge*, *10*(3), 100711–100711. https://doi.org/10.1016/j.jik.2025.100711
[27] Uddin, M., Irshad, M. S., Kandhro, I. A., Alanazi, F., Ahmed, F., Maaz, M., Hussain, S., & Ullah, S. S. (2025). Generative AI revolution in cybersecurity: a comprehensive review of threat intelligence and operations. *Artificial Intelligence Review*, *58*(8). https://doi.org/10.1007/s10462-025-11219-5
[28] Valdez, P. D., Abri, F., Webb, J., & Austin, T. H. (2025). Exploring the Use and Misuse of Large Language Models. *Information*, *16*(9), 758–758. https://doi.org/10.3390/info16090758
[29] van Berkel, N., Tag, B., Goncalves, J., & Hosio, S. (2020). Human-centred artificial intelligence: a contextual morality perspective. *Behaviour & Information Technology*, *41*(3), 1–17. https://doi.org/10.1080/0144929x.2020.1818828
[30] Yazdi, M., Zarei, E., Adumene, S., & Beheshti, A. (2024). Navigating the Power of Artificial Intelligence in Risk Management: A Comparative Analysis. *Safety*, *10*(2), 42. https://www.mdpi.com/2313-576X/10/2/42
[31] Zeijlemaker, S., Lemiesa, Y. K., Schröer, S. L., Abhishta, A., & Siegel, M. (2025). How Does AI Transform Cyber Risk Management? *Systems*, *13*(10), 835. https://doi.org/10.3390/systems13100835