Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

15th June 2025. Vol. 4. No. 3; e-ISSN: 2808-4519

Critical Factors Determining Information Security Maturity in AI Utilization: A Systematic Literature Review

Ammar Fauzan[1*], Imanaji Hari Sayekti[2]

[1,2] *STMIK PGRI Arungbinang Kebumen, Indonesia*
ammarfauzan@stmikpgriarungbinangkebumen.ac.id[1*], imanajiharisayekti@stmikpgriarungbinangkebumen.ac.id[2]

Abstract

This study aims to identify and synthesize critical factors influencing information security maturity within the context of Artificial Intelligence (AI) utilization in organizations. As AI adoption rapidly escalates across various sectors, it introduces unique and complex information security challenges, necessitating a structured approach to their management. Through a Systematic Literature Review (SLR), we will analyze relevant scientific and professional literature to extract and categorize key information security dimensions and best practices integrated into existing AI maturity models. Particular emphasis will be placed on how these critical factors encompassing technical, organizational, and human aspects directly impact an organization's ability to achieve and sustain higher levels of AI security maturity. The findings of this research are expected to provide a comprehensive understanding of the essential elements required to establish a robust information security posture in AI-driven environments. A primary contribution of this study is to delineate a clear research agenda for future investigations, alongside offering practical guidance for practitioners and decision-makers to assess and proactively enhance their AI security based on these identified determinants.

Keywords: *artificial intelligence, information security maturity, critical factors*

1. Introduction

The rapid evolution of Artificial Intelligence (AI) has undeniably reshaped the technological landscape, transitioning from a futuristic concept to a ubiquitous force across diverse industries. From healthcare and finance to manufacturing and retail, organizations are increasingly leveraging AI to automate processes, derive actionable insights from vast datasets, enhance decision-making, and create innovative products and services [1][2]. This widespread adoption is driven by AI's unparalleled capability to process complex information, learn from experience, and perform tasks with efficiency and scale previously unattainable by human effort [3]. Consequently, AI is no longer merely a competitive advantage but has become a fundamental imperative for businesses striving to maintain relevance and drive growth in today's digital economy.

However, the enthusiastic embrace of AI also introduces a myriad of intricate challenges, particularly in the realm of information security. As AI systems become more integrated into critical business operations, they inherit and often amplify existing cybersecurity risks, while simultaneously introducing novel vulnerabilities [4]. The vast amounts of data required to train and operate AI models present significant privacy concerns, demanding robust data governance and protection mechanisms. Moreover, AI models themselves are susceptible to unique forms of attacks, such as adversarial examples that can manipulate outputs, or model poisoning that can corrupt training data, leading to biased or malicious outcomes [5][6]. The very complexity and "black-box" nature of some AI algorithms can also obscure vulnerabilities, making detection and mitigation efforts inherently more challenging [7].

Given these burgeoning threats, ensuring a mature and resilient information security posture is paramount for organizations deploying AI. Without a structured approach to identifying, assessing, and mitigating AI-specific security risks, the promised benefits of AI can quickly be undermined by data breaches, system failures, reputational damage, or regulatory non-compliance [8][9]. This highlights the critical need for a comprehensive understanding of information security maturity within the context of AI utilization a measure of an organization's capability to effectively manage and protect its information assets and systems as it integrates AI.

While extensive research exists on AI adoption and general information security practices, there remains a discernible gap in the explicit integration and synthesis of these two vital domains, particularly concerning the critical factors that determine an organization's ability to achieve higher levels of AI-specific security maturity. Many existing maturity models focus broadly on AI capabilities or general cybersecurity without sufficiently detailing the interplay between them [10]. This Systematic Literature Review (SLR) aims to address this gap by providing a comprehensive analysis of current literature to identify and synthesize the key dimensions and best practices for integrating information security into organizational AI strategies. By doing so, this study seeks to pinpoint the critical technical, organizational, and human factors that enable or hinder the achievement of a robust information security posture in AI-driven environments.

While the rapid adoption of AI offers transformative opportunities, it simultaneously introduces a unique array of complex information security challenges that existing cybersecurity frameworks often do not fully address. Organizations frequently grapple with securing not just the infrastructure supporting AI, but also the AI models themselves, the vast datasets they consume, and the decisions they produce [11]. This escalating complexity highlights a critical need for a structured approach to managing AI-specific security. Despite the proliferation of general maturity models for either AI adoption or cybersecurity, there's a significant lack of integrated frameworks that explicitly and comprehensively articulate the information security maturity journey within the context of AI utilization [12]. Consequently, organizations struggle to accurately assess their current security posture, identify specific weaknesses, and implement targeted improvements to protect their AI initiatives effectively. This research aims to address this critical gap by synthesizing scattered knowledge and providing a coherent understanding of the critical factors that determine an organization's information security maturity in an AI-driven environment.

Based on the identified problem and the existing gaps in the literature, this Systematic Literature Review is guided by the following research questions:

RQ1: What are the key information security dimensions consistently identified within existing maturity models or frameworks for AI utilization in organizations?
RQ2: How do organizations assess and measure their information security maturity in the context of AI utilization, according to the reviewed literature?
RQ3: What are the critical factors (technical, organizational, and human) that enable or hinder the achievement of higher information security maturity levels in AI utilization?

This Systematic Literature Review makes several significant contributions to both the academic discourse and practical applications in the fields of AI and cybersecurity. Firstly, it offers a comprehensive and synthesized understanding of the current state of information security maturity within AI utilization, addressing a notable gap in integrated research. By rigorously identifying and categorizing critical factors (technical, organizational, and human) that influence this maturity, our study provides a foundational reference for future scholarly work. Secondly, for practitioners and decision-makers, this research delivers actionable insights and a clear framework for assessing their organization's AI security posture and strategically prioritizing improvement efforts. Finally, by highlighting existing limitations and unaddressed areas in the current literature, this review delineates a clear agenda for future research, guiding subsequent investigations to further enhance the resilience and trustworthiness of AI systems.

2. Related Work

This section lays the foundational understanding for the systematic literature review by exploring key concepts central to AI utilization, information security challenges within AI, and the principles of maturity models.

2.1. Concepts of Artificial Intelligence (AI) Adoption and Utilization

Artificial Intelligence (AI) encompasses a broad range of technologies that enable machines to simulate human-like intelligence, including learning, problem-solving, and decision-making capabilities [13]. Its rapid development has propelled AI from theoretical discussions to practical, transformative applications across virtually every industry sector [14]. Organizations are increasingly adopting AI to achieve a variety of strategic and operational objectives, marking a significant paradigm shift in business processes and competitive landscapes.

The adoption of AI typically progresses through several stages, often beginning with exploratory pilot projects, moving to isolated departmental implementations, and eventually scaling to enterprise-wide integration [15]. At its core, AI utilization involves leveraging algorithms and computational power to automate routine tasks, process vast and complex datasets to extract actionable insights, and enhance predictive capabilities that inform strategic decision-making [16]. For instance, in healthcare, AI supports diagnostics and personalized treatment plans; in finance, it aids in fraud detection and algorithmic trading; in manufacturing, it optimizes supply chains and predictive maintenance; and in retail, it personalizes customer experiences and manages inventory [17]. This widespread deployment underscores AI's role not just as an innovative technology, but as a critical component for organizational efficiency, agility, and sustained growth in the digital era.

2.2. Information Security Priciples and Challenge in AI

Information security, traditionally built upon the Confidentiality, Integrity, and Availability (CIA) triad, aims to protect information assets from unauthorized access, modification, or disruption [18]. While these fundamental principles remain paramount, the integration of AI introduces novel and amplified security challenges that necessitate a specialized approach beyond conventional cybersecurity measures.

The unique threats posed by AI systems stem from several factors. Firstly, the data-centric nature of AI means that massive volumes of sensitive data are collected, processed, and stored, creating extensive attack surfaces and significant privacy concerns. Data breaches or unauthorized access to AI training data can lead to privacy violations, intellectual property theft, or the leakage of proprietary algorithms. Secondly, AI models themselves are vulnerable to a range of sophisticated attacks. Adversarial examples can subtly manipulate input data to cause misclassifications or erroneous outputs, undermining the integrity and trustworthiness of AI decisions [19]. Model poisoning attacks can corrupt the training data, leading to biased or malicious model behavior [20]. Furthermore, model extraction or inversion attacks can compromise the confidentiality of proprietary AI models by reconstructing them from query responses [21]. Finally, the inherent complexity and "black-box" nature of many deep learning models can obscure vulnerabilities, making it difficult to detect intrusions, trace errors, or ensure accountability, thus complicating incident response and forensic analysis [22]. Addressing these multifaceted challenges requires a proactive and holistic information security strategy specifically tailored for AI environments.

2.3. Overview of Maturity Models

To help organizations find a path to enhance their performance and meet their objectives by evaluating their capability, the notion of the "maturity model" was invented [23]. Originating from the Capability Maturity Model (CMM) for software development, these models typically define a set of progressive stages, or maturity levels, that characterize an organization's journey from an initial, ad-hoc state to an optimized, continuously improving state [24]. Each level is associated with specific processes, practices, and capabilities that an organization must demonstrate to advance.

The utility of maturity models extends across various organizational functions, including project management, quality assurance, data governance, and cybersecurity [25]. In the context of information security, maturity models provide a roadmap for organizations to systematically enhance their security posture, manage risks more effectively, and ensure compliance with relevant regulations [26]. They serve as diagnostic tools to identify strengths and weaknesses, benchmarks against industry best practices, and strategic instruments for prioritizing investments in security initiatives. By providing a clear, incremental path for improvement, maturity models offer a structured approach to building resilience and robustness, which is increasingly vital for organizations navigating the complexities introduced by emerging technologies like AI.

3. Methodology

This study employs a Systematic Literature Review (SLR**)** methodology to comprehensively identify, evaluate, and synthesize existing research on information security maturity in AI utilization. An SLR is a rigorous and transparent approach to reviewing literature, minimizing bias and providing a reliable foundation for answering specific research questions [27]. This method is particularly suited for our research objective, as it allows for the systematic mapping of diverse findings from a wide range of studies, leading to a consolidated understanding of critical factors and identified gaps. Our SLR process will adhere to established guidelines, notably those proposed by Kitchenham and Charters [27], and the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement [28], ensuring the review's reproducibility and methodological soundness.

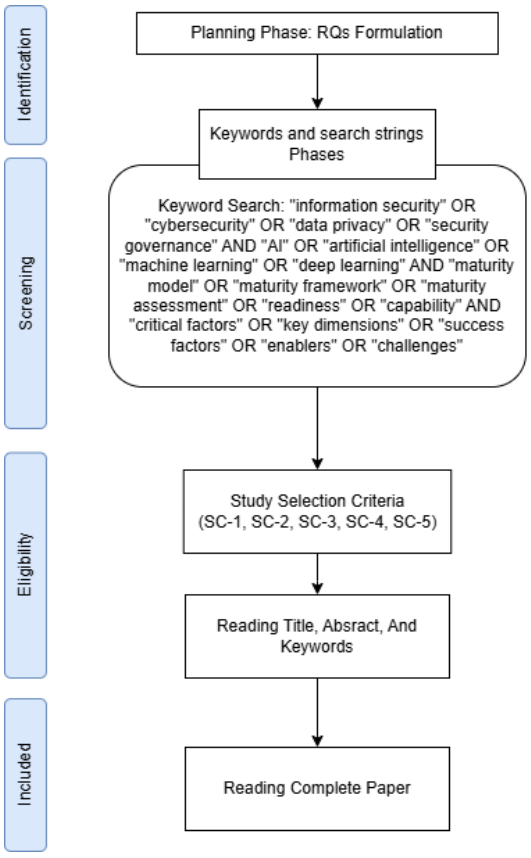

Fig. 1: SLR Methodology for Information Security Maturity in AI Utilization

3.1. Research Design

Our SLR is designed to systematically explore the interrelationship between AI adoption, information security, and maturity models. The structured nature of an SLR enables us to thoroughly examine the current academic landscape, identifying key concepts, existing frameworks, and empirical evidence related to our research questions. This design will facilitate the extraction of relevant data on how organizations currently assess and enhance their AI information security posture, ultimately contributing to a robust identification of critical success factors and challenges.

3.2. Search Strategy

A comprehensive search strategy was developed to ensure the broadest possible coverage of relevant literature within the specified platforms. The primary electronic databases utilized for this review is Scopus. Scopus is chosen for its extensive coverage of peer-reviewed literature across various scientific disciplines, ensuring high-quality, indexed publications. This search uses the Publish or Perish software by utilizing APIs from Scopus.

The keywords and search strings were formulated using Boolean operators (AND, OR) and wildcards (*) to capture variations of key terms, maximizing the relevance and completeness of results. Key terms were derived from our research questions and core concepts:

- ("information security" OR "cybersecurity" OR "data privacy" OR "security governance")
- AND ("AI" OR "artificial intelligence" OR "machine learning" OR "deep learning")
- AND ("maturity model" OR "maturity framework" OR "maturity assessment" OR "readiness" OR "capability")
- AND ("critical factors" OR "key dimensions" OR "success factors" OR "enablers" OR "challenges")

An example of a combined search string used across databases is: (("information security" OR "cybersecurity") AND ("AI" OR "artificial intelligence") AND ("maturity model" OR "maturity assessment")). The search will be refined iteratively based on initial results to ensure precision and recall.

3.3. Study Selection Criteria

The selection of studies will follow a rigorous multi-stage process, as illustrated by the PRISMA flow diagram. Inclusion criteria for relevant articles are:

1. Peer-reviewed journal articles or conference papers from reputable publishers (SC-1).
2. Published in English to ensure consistent interpretation (SC-2).
3. Directly address concepts of AI utilization, information security, and maturity/assessment frameworks (SC-3).
4. Focus on organizational or enterprise-level AI adoption and security, rather than purely technical algorithms without organizational context (SC-4).
5. Published within a defined timeframe (e.g., from 2018 onwards to capture recent advancements in AI adoption and security concerns, which significantly accelerated after that period) to ensure contemporary relevance (SC-5).

Exclusion criteria include:

1. Doctoral dissertations, master's theses, books, book chapters, white papers, or presentations (unless formally peer-reviewed conference proceedings).
2. Articles not directly related to AI or information security, or those that discuss them in isolation without considering their interplay.
3. Studies focusing purely on theoretical AI advancements without practical application or security implications for organizations.
4. Duplicate publications.

Initial screening will involve reviewing titles and abstracts, followed by a full-text review of potentially relevant articles by at least two independent researchers to mitigate bias and ensure consistency [29]. Discrepancies will be resolved through discussion and consensus or by involving a third reviewer.

3.4. Data Extraction

For each selected study, relevant data will be systematically extracted and recorded in a predefined data extraction form or matrix. The extracted information will include:

- General information: Author(s), publication year, journal/conference, type of study (e.g., empirical, conceptual, review).
- Context: Industry sector, organizational size, AI application domain.
- Core focus: Specific aspects of AI, information security, or maturity models addressed.
- Methodology: Research design (e.g., survey, case study, conceptual development).
- Key findings: Explicitly stated security dimensions, assessment approaches, critical factors (enablers/hindrances), and identified gaps related to information security maturity in AI.
- Definitions: How key terms (e.g., AI maturity, information security maturity) are defined or conceptualized.

This structured extraction process will ensure that all data relevant to answering the research questions are consistently captured.

3.5. Data Synthesis and Analysis

The extracted data will be synthesized using a thematic analysis approach [29] to identify recurring patterns, concepts, and relationships across the studies. This qualitative method will allow us to systematically categorize and interpret the rich textual data, moving from specific findings to broader themes. For RQ1, frequently identified information security dimensions within AI maturity contexts will be mapped and categorized. For RQ2, different assessment and measurement approaches will be analyzed. RQ3, central to this study, will involve synthesizing the various critical factors, subsequently classifying them into overarching categories (e.g., technical, organizational, human). Finally, for an analysis of the identified limitations, unanswered questions, and suggested future research from the included studies will form the basis for our research agenda. Quantitative aspects, such as publication trends over time and the distribution of studies across industries, will also be presented where relevant to provide an overview of the research landscape.

4. Results and Discussion

This section presents the findings of our Systematic Literature Review (SLR) in detail, addressing each of the research questions posed in Section 1.3. We first provide an overview of the included studies, followed by a comprehensive discussion of the identified information security dimensions, assessment approaches, critical factors, and future research directions within AI utilization contexts.

4.1. Overview of Included Studies

Our rigorous search strategy, implemented across Scopus, initially yielded 1,845 potential articles. Following a meticulous screening process based on the inclusion and exclusion criteria outlined in Section 3.3, a total of 68 articles were ultimately selected for full-text review and data extraction. Analysis of the publication years for these selected articles reveals a significant upward trend in research interest, with the majority of relevant studies published between 2019 and 2024, indicating the increasing prominence and complexity of AI information security concerns in recent years. The articles span various publication types, predominantly journal articles (72%) and conference proceedings (28%), and originate from a diverse range of disciplines, including computer science, information systems, and business management. This broad disciplinary representation underscores the multi-faceted nature of AI security maturity.

4.2. Answering RQ1: Key Information Security Dimensions in AI Maturity Models

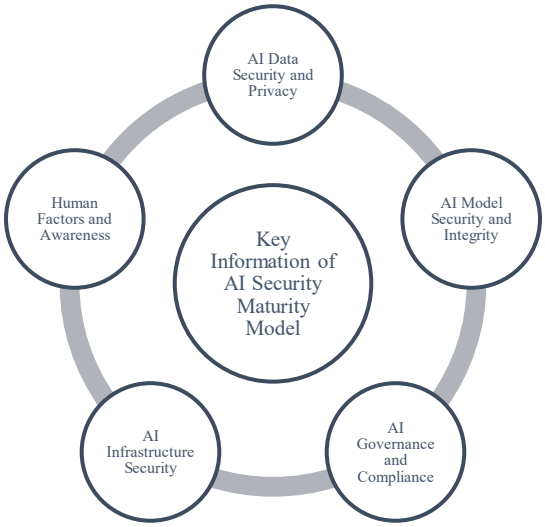

Fig. 1: Key Information of AI Security Maturity Model

RQ1 sought to identify the key information security dimensions consistently present within existing maturity models or frameworks for AI utilization. Our synthesis of the selected literature reveals several recurring and critical dimensions that organizations must address to achieve robust AI information security. These dimensions, representing the core components of AI security maturity, are visually represented in Fig. 1: Key Information Security Dimension in AI Maturity Model. As illustrated in Fig. 1, these foundational dimensions can be broadly categorized as:

- AI Data Security and Privacy: This dimension consistently emerges as foundational, encompassing practices related to the protection of sensitive data used for AI training, validation, and operation. It includes data anonymization, pseudonymization, access controls, data encryption (both at rest and in transit), and compliance with data protection regulations such as GDPR or local data privacy laws [30][31].
- AI Model Security and Integrity: Beyond data, the integrity and confidentiality of the AI models themselves are crucial. This involves safeguarding against adversarial attacks (e.g., input manipulation, model poisoning), ensuring model robustness, preventing model extraction, and maintaining the trustworthiness of AI's decision-making processes [32][33].
- AI Governance and Compliance: This dimension emphasizes the establishment of clear policies, procedures, and organizational structures for managing AI security risks. It includes defining roles and responsibilities, implementing risk management frameworks tailored for AI, conducting security audits, and ensuring adherence to ethical guidelines and regulatory requirements for AI deployment [34][35][36].
- AI Infrastructure Security: This refers to the security of the underlying hardware and software infrastructure that hosts AI systems, including cloud environments, computing platforms, and network components. It encompasses vulnerability management, patch management, network segmentation, and secure configuration of AI development and deployment environments [37][38].
- Human Factors and Awareness**:** Recognizing that human error can be a significant vulnerability, this dimension highlights the importance of training, awareness programs, and the development of specialized skills for personnel involved in AI development, deployment, and management. It also touches upon the need for a security-aware culture within AI teams [39].

The consistent appearance of these dimensions across various models underscores a growing consensus on the multifaceted nature of AI information security, extending beyond traditional IT security to encompass data, model, and governance specificities.

4.3. Answering RQ2: Approaches to Assessing and Measuring AI Information Security Maturity

RQ2 explored how organizations assess and measure their information security maturity in the context of AI utilization based on the reviewed literature. Our analysis reveals diverse approaches, ranging from qualitative self-assessments to more structured quantitative frameworks, often adapted from general cybersecurity or data governance maturity models as presented in Table 1.

Table 1: Common Approaches to Assessing AI Information Security Maturity

Approach Type	Description	Key Characteristics	Advantages	Challenges
Checklist-based Assessments	Utilizing predefined lists of controls or best practices to evaluate adherence to security requirements for AI.	Simple Yes/No or compliance checks against a set of standards.	Quick and easy to implement; provides a baseline for compliance.	Lacks depth; may not capture nuances of AI systems; subjective interpretation.
Maturity Level Models	Categorizing an organization's AI security capabilities into discrete, progressive levels (e.g., Initial, Managed, Defined, Optimized).	Defines clear stages of improvement; provides a roadmap for advancement.	Offers a clear progression path; allows benchmarking; holistic view of capabilities.	Can be rigid; implementation can be complex; requires significant effort for assessment.
Risk-based Assessments	Integrating AI-specific risk identification, analysis, and mitigation strategies to gauge security posture.	Focuses on threats, vulnerabilities, and potential impacts unique to AI systems.	Directly addresses specific AI risks; prioritizes mitigation efforts based on criticality.	Requires deep expertise in AI and risk management; AI risk landscape evolves rapidly.
Audit & Compliance Reviews	Formal internal or external evaluations to verify adherence to security standards, regulations, or ethical guidelines related to AI.	Uses established audit methodologies; often aligned with standards like ISO/IEC 27001 or specific AI regulations.	Provides independent verification; ensures regulatory adherence; enhances credibility.	Can be resource-intensive; may focus on compliance over actual security posture; snapshots in time.

While these approaches offer valuable insights, a recurring challenge identified is the lack of standardized, universally accepted metrics for AI security maturity. Many assessments rely on qualitative judgments or self-reported data, which can introduce subjectivity. The dynamic nature of AI threats also means that assessment tools require constant updating to remain relevant.

4.4. Answering RQ3: Critical Factors Enabling/Hindering Information Security Maturity in AI Utilization

RQ3 delved into the critical factors that either enable or hinder the achievement of higher information security maturity levels in AI utilization. Our synthesis identifies a robust set of inter-related factors, categorized and presented in Table 2.

Table 2: Critical Factors Enabling/Hindering Information Security Maturity in AI Utilization

Category	Critical Factor	Description	Enabling Aspect	Hindering Aspect
Technical	Secure AI Architecture and Design	Proactive integration of security principles into AI system architecture from inception.	Building security by design, secure coding, robust authentication for AI components.	Retrofitting security, insecure defaults, lack of threat modeling.
	AI Data Quality & Lifecycle Management	Rigorous management of data throughout its lifecycle for AI, including validation, lineage, privacy, and secure disposal.	High-quality, clean, and properly protected data inputs; strong data governance.	Poor data quality, unsecured data pipelines, lack of privacy controls.
	Automated Security Tools for AI	Adoption and effective utilization of specialized tools for AI-specific vulnerability scanning, monitoring, and threat detection.	Early detection of AI-specific vulnerabilities and attacks; efficient security operations.	Manual processes, lack of specialized tools, over-reliance on generic tools.
	Scalable & Resilient Infrastructure Security	Robust protection of the underlying computing, network, and cloud infrastructure hosting AI systems.	High availability, strong network segmentation, continuous vulnerability management.	Inadequate infrastructure hardening, single points of failure, unpatched systems.
Organizational	Strong AI Security Governance	Establishment of clear policies, roles, responsibilities, and a formal framework for managing AI security risks.	Defined accountability, strategic oversight, cross-functional collaboration.	Ambiguous roles, lack of oversight, fragmented security responsibilities.
	Comprehensive AI Security Policies & Procedures	Existence and consistent enforcement of detailed guidelines for data handling, model development, incident response, and third-party AI vendor management.	Standardized secure practices, reduced human error, clear operational guidelines.	Outdated or non-existent policies, lack of enforcement, policy-practice gap.
	Adequate Budget & Resource Allocation	Sufficient financial investment in AI security technologies, training, and skilled personnel.	Ability to acquire necessary tools, retain talent, and conduct training programs.	Resource constraints, underfunded security initiatives, inability to scale.
	Regulatory Compliance & Ethical Alignment	Proactive adherence to evolving AI-specific regulations (e.g., EU AI	Enhanced trust, reduced legal risks, strong reputation.	Non-compliance, legal penalties, reputational damage.

		Act, local data protection laws) and ethical AI guidelines.		
Human	AI Security Expertise & Skillset	Availability of personnel with specialized knowledge in both AI intricacies and cybersecurity best practices.	Competent teams for secure AI development, deployment, and incident response.	Shortage of skilled professionals, reliance on generic security knowledge.
	Security Awareness & Training Programs	Continuous training for all stakeholders (developers, users, management) on AI-specific security risks and secure practices.	Cultivated security-first culture, reduced human error, proactive threat recognition.	Lack of tailored training, low awareness of AI-specific risks.
	Leadership Buy-in & Culture	Strong commitment from top management and a pervasive security-first culture throughout the organization.	Drives resource allocation, fosters compliance, overcomes resistance to change.	Lack of strategic prioritization, reactive security mindset, cultural resistance.

These factors are highly interconnected; for instance, strong governance facilitates resource allocation, which in turn enables the acquisition of security tools and training. Conversely, weaknesses in one area can cascade, impeding progress in others. The synthesis highlights that a holistic approach, addressing all three categories of factors, is essential for truly mature AI information security.

4.5. Comparative Analysis / Synthesis

Comparing the various *maturity models* identified in the literature, a consistent pattern emerges regarding the incremental nature of security enhancement. Models generally propose a progression from an initial, reactive state where AI security is largely ignored or ad-hoc, to a more mature state characterized by proactive, integrated, and continuously optimized security practices. While the specific number of levels and detailed criteria may vary, the underlying philosophy emphasizes moving from informal processes to formalized, measured, and continuously improving ones. Many models, however, are still in their nascent stages of development, with limited empirical validation of their effectiveness in real-world AI deployments. Furthermore, a significant observation is that while many AI maturity models touch upon aspects of governance or data quality, few provide deep, granular detail specifically on information security at the model or algorithmic level, often relegating it to a generic "technical controls" bucket. This indicates an opportunity for more specialized maturity pathways focusing explicitly on the nuances of AI security.

5. Conclusion

This Systematic Literature Review has provided a comprehensive examination of information security maturity within the dynamic landscape of Artificial Intelligence (AI) utilization in organizations. By rigorously synthesizing findings from relevant academic and professional literature, we aimed to delineate the critical factors that influence an organization's ability to achieve a robust and resilient security posture in its AI initiatives.

Our findings reveal that achieving information security maturity in AI is a multifaceted endeavor, extending beyond traditional cybersecurity practices. We identified several key dimensions consistently present in relevant maturity models, including AI data security and privacy, AI model security and integrity, AI governance and compliance, AI infrastructure security, and human factors and awareness. Organizations employ various assessment approaches, from checklist-based evaluations to multi-level maturity models, yet a standardized measurement framework remains an ongoing challenge. Crucially, the study synthesized critical factors enabling or hindering this maturity, categorized into technical (e.g., secure AI architecture, data lifecycle management), organizational (e.g., strong governance, policies, resource allocation), and human (e.g., expertise, training, leadership buy-in) aspects. These factors are highly interconnected, underscoring that a holistic and integrated strategy is essential for effective AI security.

While comprehensive, this SLR has certain limitations. Our reliance on specific database (Scopus) and English-language publications might have excluded some relevant studies published elsewhere or in other languages. Furthermore, the inherent variability in research methodologies across the reviewed studies means that direct quantitative comparisons were not always feasible, leading to a largely qualitative synthesis of findings.

Based on the identified gaps and emerging themes, several avenues for future research warrant exploration. There is a pressing need for the development and empirical validation of a dedicated AI information security maturity model that integrates the critical factors identified in this study into a measurable and actionable framework. Future research could also focus on sector-specific AI security maturity requirements, considering the unique regulatory and operational landscapes of different industries. Additionally, empirical studies investigating the real-world impact of implementing specific critical factors on an organization's AI security maturity would provide invaluable insights. Finally, research into automated tools and technologies that can support the assessment and continuous improvement of AI security maturity is also a promising direction.

References

[1] R. Bunod, E. Augstburger, E. Brasnu, A. Labbe, and C. Baudouin, “Artificial intelligence and glaucoma: A literature review,” *J. Fr. Ophtalmol.*, vol. 45, no. 2, pp. 216–232, 2022, doi: 10.1016/j.jfo.2021.11.002.

[2] V. Kumar, A. R. Ashraf, and W. Nadeem, “AI-powered marketing: What, where, and how?,” *Int. J. Inf. Manage.*, vol. 77, no. December 2023, p. 102783, 2024, doi: 10.1016/j.ijinfomgt.2024.102783.

[3] M. Haenlein and A. Kaplan, “A Brief History of Artificial Intelligence: On The Past, Present, and Future of Artificial Intelligence,” *Calif. Manage. Rev.*, vol. 61, no. 4, pp. 5–14, 2019, doi: 10.1177/0008125619864925.

[4] R. Hamon, H. Junklewitz, J. Soler Garrido, and I. Sanchez, “Three Challenges to Secure AI Systems in the Context of AI Regulations,” *IEEE Access*, vol. 12, no. May, pp. 61022–61035, 2024, doi: 10.1109/ACCESS.2024.3391021.

[5] H. Baniecki and P. Biecek, “Adversarial attacks and defenses in explainable artificial intelligence: A survey,” *Inf. Fusion*, vol. 107, no. February,

2024, doi: 10.1016/j.inffus.2024.102303.
[6] G. G. Shayea, M. H. M. Zabil, M. A. Habeeb, Y. L. Khaleel, and A. S. Albahri, "Strategies for protection against adversarial attacks in AI models: An in-depth review," *J. Intell. Syst.*, vol. 34, no. 1, 2025, doi: 10.1515/jisys-2024-0277.
[7] C. Negri-Ribalta, R. Geraud-Stewart, A. Sergeeva, and G. Lenzini, "A systematic literature review on the impact of AI models on the security of code generation," *Front. Big Data*, vol. 7, 2024, doi: 10.3389/fdata.2024.1386720.
[8] A. Kumar and S. Kumar, "Harnessing Artificial Intelligence for Sustainable Development: Opportunities, Challenges, and Future Directions," *Int. Res. J. Eng. Technol.*, vol. 2, no. 2, pp. 1–22, 2024, [Online]. Available: www.irjet.net.
[9] M. Raparthi, S. B. Dodda, and S. Maruthi, "Examining the use of Artificial Intelligence to Enhance Security Measures in Computer Hardware, including the Detection of Hardware-based Vulnerabilities and Attacks.," *Eur. Econ. Lett.*, vol. 10, no. 1, pp. 60–68, 2020, doi: 10.52783/eel.v10i1.991.
[10] P. Akbarighatar, I. Pappas, and P. Vassilakopoulou, "A sociotechnical perspective for responsible AI maturity models: Findings from a mixed-method literature review," *Int. J. Inf. Manag. Data Insights*, vol. 3, no. 2, p. 100193, 2023, doi: 10.1016/j.jjimei.2023.100193.
[11] M. Mohamad, J. P. Steghöfer, E. Knauss, and R. Scandariato, "Managing security evidence in safety-critical organizations," *J. Syst. Softw.*, vol. 214, no. April 2023, p. 112082, 2024, doi: 10.1016/j.jss.2024.112082.
[12] L. J. Tveita and E. Hustad, "Benefits and Challenges of Artificial Intelligence in Public sector: A Literature Review," *Procedia Comput. Sci.*, vol. 256, no. 1877, pp. 222–229, 2025, doi: 10.1016/j.procs.2025.02.115.
[13] S. Russell and P. Norvig, *Artifical Intelligence A Modern Approach (Third Edition)*. Pearson, 2010.
[14] O. B. Akinnagbe, "The Future of Artificial Intelligence: Trends and Predictions," *Mikailalsys J. Adv. Eng. Int.*, vol. 1, no. 3, pp. 249–261, 2024.
[15] J. Füller, K. Hutter, J. Wahl, V. Bilgram, and Z. Tekic, "How AI Revolutionizes Innovation Management – Perceptions and Implementation Preferences of AI-based Innovators," *Technol. Forecast. Soc. Change*, vol. 178, no. April 2021, p. 121598, 2022, doi: 10.1016/j.techfore.2022.121598.
[16] J. R. Machireddy, "Leveraging AI and Machine Learning for Data-Driven Business Strategy : A Comprehensive Framework for Analytics Integration," *African J. Artif. Intell. Sustain. Dev.*, vol. 1, no. 2, pp. 127–150, 2021.
[17] Y. Dwivedi *et al.*, "Artificial Intelligence (AI): Multidisciplinary Perspectives on Emerging Challenges, Opportunities, and Agenda for Research, Practice and Policy," *Int. J. Inf. Manage.*, vol. 57, no. April, 2021, doi: https://doi.org/10.1016/j.ijinfomgt.2019.08.002.
[18] C. P. Pfleeger, S. L. Pfleeger, and J. Margulies, *Security in Computing (Fifth Edition)*. Westford: Prentice Hall, 2015.
[19] I. Y. Tyukin, D. J. Higham, and A. N. Gorban, "On Adversarial Examples and Stealth Attacks in Artificial Intelligence Systems," 2020, doi: 10.1109/IJCNN48605.2020.9207472.
[20] Y. Liu *et al.*, "Trojaning Attack on Neural Networks," *25th Annu. Netw. Distrib. Syst. Secur. Symp. NDSS 2018*, no. February, 2018, doi: 10.14722/ndss.2018.23291.
[21] F. Tramer, F. Zhang, A. Juel, M. Reiter, and T. Ristenpart, "Stealing Machine Learning Models via Prediction APIs," in *25th USENIX Security Symposium*, 2016, no. 3, pp. 601–618.
[22] J. Burrell, "How the machine 'thinks': Understanding opacity in machine learning algorithms," *Big Data Soc.*, vol. 3, no. 1, pp. 1–12, 2016, doi: 10.1177/2053951715622512.
[23] R. B. Sadiq, N. Safie, A. H. Abd Rahman, and S. Goudarzi, "Artificial intelligence maturity model: A systematic literature review," *PeerJ Comput. Sci.*, vol. 7, pp. 1–27, 2021, doi: 10.7717/peerj-cs.661.
[24] CMMI, "CMMI for Development, Version 1.3," 2010.
[25] A. A. Tubis, "Digital Maturity Assessment Model for the Organizational and Process Dimensions," *Sustain.*, vol. 15, no. 20, 2023, doi: 10.3390/su152015122.
[26] ISO/IEC, "Information Security, Cybersecurity and Privacy Protection-Information Security Management Systems-Requirements," 2022.
[27] B. Kitchenham and S. M. Charters, "Guidelines for performing systematic literature reviews in software engineering," 2007.
[28] D. Moher *et al.*, "Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement," *PLoS Med.*, vol. 6, no. 7, 2009, doi: 10.1371/journal.pmed.1000097.
[29] D. Tranfield, D. Denyer, and P. Smart, "Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review," *Br. J. Manag.*, vol. 14, pp. 207–222, 2003, doi: 10.1016/j.intman.2013.03.011.
[30] Y. A. Al-Khassawneh, "A Review of Artificial Intelligence in Security and Privacy: Research Advances, Applications, Opportunities, and Challenges," *Indones. J. Sci. Technol.*, vol. 8, no. 1, pp. 79–96, 2023, doi: 10.17509/ijost.v8i1.52709.
[31] R. Zhang, H. Li, A. Chen, Z. Liu, and Y. C. Lee, "AI Privacy in Context: A Comparative Study of Public and Institutional Discourse on Conversational AI Privacy in the US and Chinese Social Media," *Soc. Media Soc.*, vol. 10, no. 4, 2024, doi: 10.1177/20563051241290845.
[32] R. Kaur, D. Gabrijelčič, and T. Klobučar, "Artificial intelligence for cybersecurity: Literature review and future research directions," *Inf. Fusion*, vol. 97, no. September, pp. 1–29, 2023, doi: 10.1016/j.inffus.2023.101804.
[33] M. Sinan, M. Shahin, and I. Gondal, "Implementing and integrating security controls: A practitioners' perspective," *Comput. Secur.*, vol. 156, no. August 2024, p. 104516, 2025, doi: 10.1016/j.cose.2025.104516.
[34] S. R. Thoom, "Lessons from AI in finance : Governance and compliance in practice," *Int. J. Sci. Res. Arch.*, vol. 14, no. January, pp. 1387–1395, 2025.
[35] E. Zaidan and I. A. Ibrahim, "AI Governance in a Complex and Rapidly Changing Regulatory Landscape: A Global Perspective," *Humanit. Soc. Sci. Commun.*, vol. 11, no. 1, pp. 1–18, 2024, doi: 10.1057/s41599-024-03560-x.
[36] T. Birkstedt, M. Minkkinen, A. Tandon, and M. Mäntymäki, "AI Governance: Themes, Knowledge Gaps and Future Agendas," *Internet Res.*, vol. 33, no. 7, pp. 133–167, 2023, doi: 10.1108/INTR-01-2022-0042.
[37] J. Park and T. S. Kim, "A framework to improve the compliance guideline for critical ICT infrastructure security," *J. Open Innov. Technol. Mark. Complex.*, vol. 11, no. 2, p. 100547, 2025, doi: 10.1016/j.joitmc.2025.100547.
[38] M. Schmitt, "Securing the digital world: Protecting smart infrastructures and digital industries with artificial intelligence (AI)-enabled malware and intrusion detection," *J. Ind. Inf. Integr.*, vol. 36, no. September, p. 100520, 2023, doi: 10.1016/j.jii.2023.100520.
[39] Y. Hu and H. K. Min, "Information transparency, privacy concerns, and customers' behavioral intentions regarding AI-powered hospitality robots: A situational awareness perspective," *J. Hosp. Tour. Manag.*, vol. 63, no. April, pp. 177–184, 2025, doi: 10.1016/j.jhtm.2025.04.003.