Journal of Economich, Technology and Business (JETBIS)
Volume 4, Number 10 Oktober 2025
p-ISSN 2964-903X; e-ISSN 2962-9330

Challenges and Cybersecurity Solutions in the Era of Big Data and Artificial Intelligence

***Abdul Robi Padri[1], Erwin Iskandar[2]**
Politeknik Siber Cerdika Internasional, Indonesia[1, 2]
*Correspondence: abdulrobi@polteksci.ac.id

KEYWORDS:
cybersecurity; big data; artificial intelligence; cyber threats; zero trust; machine learning

ABSTRACT
The development of big data and artificial intelligence (AI) technology has brought about fundamental transformations in data management and utilization across various sectors. However, despite the various benefits offered, this era has also given rise to increasingly complex and diverse cybersecurity challenges. This study aims to identify the main cybersecurity challenges faced by organizations in the era of big data and AI, analyze the effectiveness of available security solutions, and formulate a strategic framework for strengthening AI-based cybersecurity. The study used a systematic literature review (SLR) approach by analyzing 87 scientific articles published between 2019 and 2024 from various databases such as Scopus, IEEE Xplore, and Google Scholar. The results show that the main cyber threats include adversarial attacks against AI models (32.4%), massive data leaks (28.7%), and AI-based ransomware attacks (21.5%). The most effective solutions include implementing a Zero Trust Architecture framework, machine learning-based threat detection, and layered data encryption. This research produces an integrated strategic framework that combines AI technology with a proactive security policy approach. These findings are expected to serve as a reference for organizations, cybersecurity practitioners, and policymakers in designing security systems that are adaptive and responsive to the ever-evolving threats in the digital era.

INTRODUCTION

The rapid digital transformation has positioned big data and artificial intelligence as two key pillars driving innovation across various industrial sectors. Global data volume is projected to reach 175 zettabytes by 2025, making data management and protection an increasingly critical challenge for organizations worldwide (Reinsel et al., 2018). This phenomenon of massive data exploitation has significantly expanded the cyberattack surface, opening new vulnerabilities previously unaddressed by conventional security systems.

The urgency of this research is reinforced by the fact that global losses from cybercrime are estimated to reach US$8 trillion in 2023 and projected to surge to US$10.5 trillion by 2025 (Cybersecurity Ventures, 2023). This figure surpasses losses from natural disasters and illustrates the seriousness of the cyber threat to global economic stability. While the advent of artificial intelligence enhances defense capabilities, threat actors are also exploiting it to launch more sophisticated, stealthy, and destructive attacks.

Several theories and supporting data support this research. According to the IBM Cost of a Data Breach Report (2023), the average global cost of a data breach reached USD 4.45 million per incident, the highest figure in recorded history. Furthermore, the report noted that organizations that integrated AI and security automation were able to reduce the data breach cycle by an average of 108 days compared to those that did not, confirming the relevance of AI technology in modern cyber defense architecture (IBM Security, 2023).

Previous research has explored many partial aspects of this problem. Buczak & Guven (2016) laid a crucial foundation for the application of machine learning in network intrusion detection, while Apruzzese et al. (2022) explored adversarial machine learning threats targeting AI-based security systems themselves. Meanwhile, Sun et al. (2020) addressed privacy challenges in big data environments using a differential privacy approach. While these studies are invaluable, studies that integrate the challenges of the big data ecosystem, the evolution of AI-based threats, and provide comprehensive, integrated solutions are still limited.

The research gap identified in the literature review indicates that most existing studies are partial and sectoral, focusing on only one threat dimension or a specific technical solution without holistically considering the complex interactions between threats, big data infrastructure, and AI capabilities. The absence of an analytical framework that integrates these three dimensions represents a significant gap that needs to be filled, especially in the context of the rapidly evolving cyberthreat landscape (Lim et al., 2019).

The novelty of this research lies in the development of an integrated strategic framework that combines an in-depth analysis of cyberthreat typologies in the era of big data and AI with empirically validated, evidence-based solution mapping. Unlike previous studies, this research adopts a systematic approach that simultaneously encompasses technical, organizational, and regulatory dimensions, resulting in a layered security model adaptable to various types and scales of organizations (Cheng et al., 2022).

Based on the background and gap identification above, this study has three main objectives: (1) identifying and classifying the main cybersecurity challenges facing organizations in the era of big data and AI; (2) analyzing the effectiveness of various available cybersecurity solutions, both from a technical and strategic perspective; and (3) formulating a strategic framework for strengthening artificial intelligence-based cybersecurity that can serve as an implementation guide for organizations. These three objectives are designed to provide meaningful practical and theoretical contributions to the development of the cybersecurity discipline (P. Mishra & Koehler, 2006)

RESEARCH METHOD

Types of Research

This study employed a Systematic Literature Review (SLR) approach, adhering to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) protocol. The SLR approach was chosen because it allows for a comprehensive, objective, and replicable synthesis of knowledge from relevant, high-quality literature. This type of research is qualitative-quantitative (mixed methods), dominated by content and bibliometric analysis of scientific articles discussing cybersecurity in the context of big data and AI (Kitchenham & Charters, 2007).

Population and Sample

The study population included all English and Indonesian-language scientific articles published in journals indexed by Scopus Q1-Q2, IEEE Xplore, ACM Digital Library, and Google Scholar between 2019 and 2024, with the main topics being cybersecurity, big data, and artificial intelligence. An initial search using predetermined keywords yielded 1,247 articles as the initial population. After a multi-step selection process

based on strict inclusion and exclusion criteria, the final sample analyzed consisted of 87 articles. The selection of the 2019-2024 period was based on considerations of relevance to the development of generative AI technology and the current big data ecosystem (Page et al., 2021).

Research Instruments

The research instruments used included: (1) a literature search protocol containing standardized search strings, namely ("cybersecurity" OR "cyber security" OR "information security") AND ("big data" OR "data analytics") AND ("artificial intelligence" OR "machine learning" OR "deep learning"); (2) a data extraction form containing columns for recording title, author, year, methodology, main findings, and challenge/solution categories; and (3) a quality assessment sheet using the Critical Appraisal Skills Programme (CASP) instrument that has been modified for the context of information technology research (Moher et al., 2014).

Data Collection Techniques

Data collection was conducted in several stages. First, a systematic search was conducted in the Scopus, IEEE Xplore, ACM Digital Library, and Google Scholar databases using predefined search strings. Second, the search results were exported to the Mendeley reference management application for deduplication. Third, screening was conducted in two stages: title- and abstract-based screening by two researchers independently, followed by full-text screening considering the inclusion and exclusion criteria . Concordance between researchers was measured using Cohen's Kappa coefficient, which yielded a κ value of 0.82 (very good), indicating a high level of consistency in the article selection process (Liberati et al., 2009).

Research Procedures

The research procedure followed the PRISMA steps, which consist of four phases: (1) Identification — a literature search across all designated databases and identification of additional articles through cross-referencing; (2) Screening — filtering out duplicates and articles that did not meet the initial criteria based on title and abstract; (3) Eligibility — assessing the eligibility of full texts based on comprehensive inclusion/exclusion criteria; and (4) Inclusion — determining the final articles that met all criteria for in-depth analysis. Articles were included if they specifically addressed cybersecurity in the context of big data and/or AI, used a clear methodology, and were published in reputable indexed journals or proceedings (Tranfield et al., 2003)

Data Analysis Techniques

Data analysis employed three complementary approaches. First, qualitative content analysis was conducted by thematic coding 87 articles using NVivo 12 software to identify recurring patterns, themes, and categories of challenges and solutions. Second, bibliometric analysis was conducted using VOSviewer to

map citation networks, co-authorship, and keyword co-occurrence, resulting in a comprehensive visualization of the research landscape. Third, narrative synthesis and evidence mapping were used to integrate findings from multiple studies into a cohesive and practically actionable conceptual framework (Braun & Clarke, 2006).

Table 1. Literature Selection Results Based on the PRISMA Flow

PRISMA Phase	Number of Articles	Information
Identification	1,247	Search results for all databases
Screening (duplicates removed)	893	After deduplication
Screening (title & abstract)	412	Passed the initial screening
Eligibility (full text)	143	Full text available
Included (final)	87	Analyzed in depth

Source: Results of research data processing, 2024

RESULTS AND DISCUSSION

Cyber Threat Landscape in the Era of Big Data and AI

An analysis of 87 selected articles reveals that the cyberthreat landscape in the era of big data and AI has evolved significantly compared to the previous decade. Threats are no longer limited to exploiting simple technical vulnerabilities, but have evolved into adaptive, persistent attacks capable of exploiting the complexity of modern data infrastructure. The identified threat taxonomy includes seven main categories, each with distinct characteristics, targets, and impacts (Shu et al., 2023)

Adversarial attacks against AI models were the most frequently discussed threat in the reviewed literature, appearing in 28 of the 87 articles (32.4%). This type of attack involves the deliberate manipulation of input data to cause the AI model to make incorrect predictions or decisions, undetected by conventional monitoring systems. The implications of adversarial attacks are serious, particularly in security-critical domains such as malware detection, facial recognition systems, and autonomous vehicles, where misclassification can be fatal (Goodfellow et al., 2014).

Massive data breaches ranked second, occurring 28.7% of the total literature analyzed. Big data ecosystems, which aggregate billions of data records from various sources, create highly attractive targets for threat actors. The Verizon Data Breach Investigations Report (2023) noted that 74% of all data breaches involve a human element, whether through human error, misuse of access rights, or social engineering enhanced by generative AI techniques (Singh, 2025).

AI-powered ransomware has emerged as the fastest-growing threat, with incidents growing 130% between 2021 and 2023. The latest generation of ransomware leverages machine learning algorithms to selectively select targets, determine optimal attack timing, and dynamically set ransom amounts based on an assessment of the victim's financial capacity. These capabilities make modern ransomware far more effective and difficult to detect than its predecessors (Brewer, 2022).

AI-enhanced insider threats also received significant attention in 15.7% of the literature reviewed. AI enables malicious actors within organizations to disguise their suspicious activity within normal behavioral patterns, leverage predictive analytics to identify gaps in surveillance systems, and automate the exfiltration of large amounts of data without triggering alarms for

conventional security systems. This challenge is compounded by the difficulty of distinguishing legitimate from malicious user activity in a highly dynamic data environment (Matthew Collins et al., 2018; Theis M. et al., 2022).

Attacks on AI pipeline infrastructure—including training data poisoning, model manipulation, and attacks on the model supply chain—represent a relatively new threat category with far-reaching implications. By injecting contaminated data into the AI model training process, attackers can covertly embed backdoors that allow them to control the model's future behavior. These attacks are particularly dangerous because their effects are not felt until well after the system has been deployed in production (Chen et al., 2019; Wang et al., 2024).

The threat of deepfakes and AI-based disinformation is also increasingly relevant as a cyberattack vector, particularly in the context of advanced social engineering and segmented phishing attacks. Generative AI technology enables the creation of fake audio, video, and text content that is nearly indistinguishable from the real thing, making it highly effective for manipulating specific targets, including senior executives, in Business Email Compromise (BEC) attacks (Chesney & Citron, 2019)

AI-powered Distributed Denial of Service (DDoS) attacks have also shown significant sophistication. AI-powered botnets are now able to adapt their attack patterns in real time based on the response of the target's defense systems, selecting the most effective attack vectors, and coordinating attacks across millions of compromised devices simultaneously. This makes conventional DDoS mitigation increasingly inadequate and requires an equally adaptive defense approach (Bhardwaj et al., 2020).

Table 2. Categorization of Cyber Threats in the Era of Big Data and AI

Threat Category	Main Target	Frequency (%)	Impact
Adversarial AI Attack	ML/AI Model	32.4%	Tall
Massive Data Breach	Database & Data Lake	28.7%	Very high
AI-Powered Ransomware	Critical Systems & Files	21.5%	Very high
Insider Threat + AI	Internal Sensitive Data	15.7%	Tall
AI Pipeline Poisoning	AI Training Model	12.3%	Very high
Deepfakes & Disinformation	Users & Executives	10.8%	Medium-High
AI-Enhanced DDoS	Network Infrastructure	9.1%	Tall

Source: Results of literature analysis, 2024

AI-Based Cybersecurity Technical Solutions

An analysis of 87 articles shows that AI-based cybersecurity solutions have rapidly evolved and encompass multiple layers of information system architecture. These solutions are no longer merely reactive to known threats, but are increasingly oriented toward proactive prediction and prevention through real-time processing of large amounts of data. The effectiveness of AI-based

solutions has consistently been reported to be higher than conventional rule-based approaches in detecting complex and previously unseen threats (Buczak & Guven, 2016)

Machine Learning-Based Intrusion Detection Systems (ML-IDS) are the most widely researched and implemented solutions. Unlike signature-based IDS systems, which can only recognize cataloged threats, ML-IDS can identify behavioral anomalies indicating new attacks based on statistical patterns learned from historical data. Research by Ullah & Mahmoud (2022) shows that an LSTM-based deep learning model achieved an intrusion detection accuracy of 98.7% on the NSL-KDD dataset, with a significantly lower false positive rate than conventional methods (Ullah & Mahmoud, 2022).

Behavioral Analytics and AI-based User and Entity Behavior Analytics (UEBA) play a crucial role in detecting insider threats and attacks that use compromised legitimate credentials. This technology builds a profile of normal behavior for each user and entity on the network and then continuously monitors for deviations from that baseline. Modern UEBA solutions powered by deep learning are capable of detecting subtle behavioral changes that would otherwise go unnoticed by static security rules or human oversight (Matthew Collins et al., 2018; Theis M. et al., 2022).

The AI-based Automated Threat Intelligence Sharing Platform enables real-time, standardized exchange of threat information between organizations. The platform uses Natural Language Processing (NLP) to extract Indicators of Compromise (IoC) from various sources, including dark web forums, incident reports, and commercial threat feeds, then correlates and distributes this information in STIX/TAXII format. This approach significantly shortens response times to attack campaigns targeting multiple organizations simultaneously (Tounsi & Rais, 2022).

AI-Powered Security Orchestration, Automation, and Response (SOAR) integrates various security tools that previously operated in silos and automates incident response based on AI-enhanced playbooks. By automating repetitive and time-consuming incident response tasks, SOAR platforms enable security analysts to focus on more complex case investigations. Enterprise case studies show that SOAR implementation can reduce Mean Time to Response (MTTR) by up to 85% (Patel & Patel, 2022).

Federated Learning for Cybersecurity has emerged as an innovative approach that enables collaborative threat detection model training without the need to directly share sensitive data. In a federated learning architecture, models are trained locally on each node, and then only model parameters (not data) are aggregated on a central server. This approach is particularly relevant in scenarios where direct sharing of sensitive data is not possible due to regulatory, privacy, or competitive constraints (Li et al., 2020).

Explainable AI (XAI) for security has become an increasingly important component as reliance on automated decisions in security operations increases. AI models that cannot explain the reasoning behind their decisions (black-box) pose trust and accountability challenges, particularly in the context of digital forensic investigations and regulatory compliance. XAI techniques such as LIME and SHAP enable security analysts to understand the contribution of specific features to each threat detection decision, increasing model trustworthiness and debugability (Amarasinghe et al., 2019).

Homomorphic encryption and privacy-preserving computation offer a revolutionary solution for processing encrypted data without first decrypting it. While still facing computational performance challenges, advances in fully homomorphic encryption (FHE) have begun to pave the

way for practical implementation in sensitive data analytics scenarios. Integrating this technique with the big data ecosystem has the potential to eliminate the fundamental trade-off between data utility and privacy that has been a major obstacle in managing large-scale sensitive data (Acar et al., 2018).

Table 3. Effectiveness of AI-Based Cybersecurity Solutions

Security Solutions	Threats Addressed	Accuracy/Effectiveness	Implementation Maturity
ML-IDS	Intrusion & Anomaly	94-99%	High (Production-ready)
UEBA	Insider Threat	89-96%	Tall
AI-SOAR	Multi-threat	MTTR -85%	Tall
Federated Learning IDS	Distributed Threat	91-97%	Growing
XAI Security	Manipulation Model	Qualitative	Currently
Homomorphic Encryption	Data Breach	Height (theoretical)	Low-Medium

Source: Literature synthesis, 2024

Zero Trust Architecture Framework in Big Data Ecosystem

Zero Trust Architecture (ZTA) has evolved from a philosophical concept into a standardized implementation framework, especially after the National Institute of Standards and Technology (NIST) released official guidance through Special Publication 800-207 in 2020. The core principle of ZTA is 'never trust, always verify,' which means that no entity—whether a user, device, or application—is trusted by default, even if they are within the same network perimeter. The application of ZTA in the big data ecosystem fundamentally shifts the security paradigm from a perimeter-based approach to an identity-centric one (National Institute of Standards and Technology, 2020; Phiayura & Teerakanok, 2023).

The implementation of ZTA in big data infrastructure involves several key components that work synergistically. Micro-segmentation divides the network into small, cryptographically isolated zones, limiting the lateral movement of attackers even if they successfully compromise one of the segments. AI-enhanced Identity and Access Management (IAM) continuously verifies identities based on multiple factors, including user behavior, device context, and real-time risk analytics. This approach has been shown to significantly reduce the impact of attacks that successfully penetrate the outermost layer of defense (Abdelmagid et al., 2026)

Implementing ZTA on big data platforms such as Hadoop, Spark, and Kafka requires significant architectural adaptations. Unlike traditional enterprise environments, big data platforms are often designed to maximize open data access to support analytical collaboration, which inherently contradicts the principle of least privilege in ZTA. Research by C.Kanmani Pappa (2024) proposes an adaptive ZTA framework for big data that implements dynamic attribute-based access

control (ABAC), where access rights are determined not only by user identity but also by the sensitivity of the accessed data and the context of the request (C.Kanmani Pappa, 2024).

Continuous Monitoring and Validation (CMV) is the third pillar of ZTA, which becomes extremely complex in the context of big data due to the sheer volume of events that must be monitored. Modern big data platforms can generate millions of security events per second, far exceeding the analytical capacity of humans or conventional SIEM systems. The solution is to integrate AI-powered security analytics capable of processing real-time event data streams, correlating anomalies across sources, and prioritizing alerts based on dynamically calibrated risk scores (Kindervag et al., 2010, 2016)

Deep learning-based Network Traffic Analysis (NTA) is an essential component in implementing ZTA for big data. Deep learning models, particularly transformer-based architectures, demonstrate superior capabilities in analyzing network traffic patterns and identifying suspicious communications that may indicate data exfiltration or Command and Control (C2) attacks by malware. A study by Lotfollahi et al. (2020) demonstrated that a deep learning approach can classify encrypted network traffic with up to 98.8% accuracy, surpassing traditional traffic analysis methods (Lotfollahi et al., 2020).

The main challenges in ZTA adoption are implementation complexity and significant costs. A Gartner survey of Fortune 500 companies found that 60% of organizations face difficulties defining appropriate segment boundaries and managing highly granular access policies at scale. Existing IAM maturity, legacy systems not designed for ZTA, and a limited pool of security professionals with in-depth ZTA understanding are the most frequently reported implementation barriers (Edo et al., 2022; Keshav Jena, 2023).

Integrating ZTA with data protection regulations such as the European GDPR and Government Regulation No. 71 of 2019 concerning the Implementation of Electronic Systems and Transactions in Indonesia requires special attention. The data minimization principle in the GDPR and the data security obligations stipulated in domestic regulations naturally align with the ZTA philosophy. However, technical implementation that simultaneously meets the requirements of both frameworks requires a holistic approach that considers legal, technical, and operational aspects in an integrated manner (Sirur et al., 2018)

The ZTA maturity model proposed in this study identifies five levels of adoption, ranging from Traditional (Level 0) to Optimal ZTA (Level 4), with assessment parameters covering identity, device, network, application, and data dimensions. Organizations can use this model as a roadmap to plan a gradual and measured transition to ZTA, aligned with resource capacity and acceptable risk levels. Higher maturity levels are positively correlated with a significant reduction in security incidents, based on data from 42 case studies analyzed (Phiayura & Teerakanok, 2023)

Regulatory, Privacy, and Ethical Challenges of AI in Cybersecurity

Regulatory and ethical dimensions form a critical context that both constrains and shapes the development of cybersecurity solutions in the AI era. The proliferation of global data protection regulations—including GDPR (Europe), CCPA (California), PIPL (China), and sectoral regulations such as HIPAA for the healthcare sector—creates a highly complex compliance landscape for multinational organizations. The incompatibility between these varying regulatory requirements

significantly complicates the implementation of uniform and efficient security solutions (Hoofnagle et al., 2019).

The privacy versus security dilemma is one of the most widely discussed fundamental tensions in the reviewed literature. On the one hand, AI-based security systems require access to large amounts of user behavioral data to establish accurate baselines and detect anomalies with high precision. On the other hand, the collection and processing of such data can violate individual privacy rights and contradict the principle of data minimization. Without a clear legal framework and transparent consent mechanisms, implementing effective AI security solutions can put organizations at serious litigation and reputational risk (Hosein, 2019)

Algorithmic bias in cybersecurity AI systems poses serious ethical and operational implications. Threat detection models trained on biased historical data—for example, data that overrepresents certain user groups or activity patterns—can produce disproportionately high false positive rates in underrepresented groups. In the context of system access, this can have systemic discriminatory effects and potentially violate the principles of fairness and non-discrimination increasingly codified in emerging AI regulations across jurisdictions (Mehrabi et al., 2021).

Rapidly evolving AI regulations, such as the EU AI Act, which establishes a risk-based framework for AI systems, directly impact the development and deployment of AI-based security systems. AI systems used in security-critical contexts are classified as 'high-risk' under the EU AI Act, which requires them to meet stringent standards for transparency, accuracy, robustness, and security. Compliance with these regulations adds a layer of complexity to the development of AI security solutions, but at the same time promotes increased quality and trust in these systems (Veale & Zuiderveen Borgesius, 2021)

Data sovereignty and data localization are increasingly strategic issues in the context of global cybersecurity. Many countries, including Indonesia, through the Personal Data Protection Law (PDP Law) passed in 2022, require citizens' data to be stored and processed within their national borders. This provision directly impacts the architecture of cloud-based security solutions and the ability to share threat intelligence across borders, both of which are essential elements of an effective modern cybersecurity ecosystem (Asogwa, 2020).

Transparency and accountability in automated decision-making by AI security systems are becoming increasingly pressing regulatory demands. When AI systems automatically block access, isolate devices, or flag transactions as suspicious, audit mechanisms must be in place to examine the rationale behind each decision. Article 22 of the General Data Protection Regulation (GDPR) grants individuals the right not to be subjected to automated decisions that significantly impact their data, requiring adequate explainability from AI systems used in security operations (Goodman & Flaxman, 2017)

Ethical hacking and vulnerability disclosure in the AI ecosystem present new challenges that have not been fully addressed by existing legal frameworks. Security researchers identifying vulnerabilities in AI models face significant legal uncertainty: can adversarial testing of AI systems used in critical domains constitute unauthorized computer access? The need for clear safe harbor provisions for responsible AI security research is a consistent policy recommendation in the reviewed literature (Findley et al., 2024)

The AI ethics framework for cybersecurity proposed in this study adopts four key principles: fairness (equality in protection and impact), accountability (clear responsibility for system

decisions), transparency (openness about how the system works), and safety (the reliability and security of the system itself)—known as the FATS framework. Consistent application of this framework throughout the development and operation cycle of AI security systems is expected to reduce ethical risks while increasing technical effectiveness by enhancing stakeholder trust (Jobin et al., 2019)

Integrated Strategic Framework for Strengthening Cybersecurity in the AI Era

The synthesis of the entire literature analysis resulted in an Integrated Cybersecurity Strategic Framework (ISCFR) consisting of four interrelated layers: Governance & Compliance, Architecture & Technology, Operations & Intelligence, and Human Capital & Culture. The ISCFR was designed to address the fundamental weaknesses of a siloed security approach that focuses solely on technical solutions, by holistically integrating risk management, technology, operational processes, and human factors. This framework was validated through comparison with 12 case studies of successful implementations in the reviewed literature (Cheng et al., 2022)

The Governance & Compliance layer within the KSTKS serves as the foundation that defines the direction and boundaries of security operations. Its key components include a risk-based security policy calibrated to an organization's specific threat profile, a compliance framework with applicable regulations, an AI governance mechanism that ensures the responsible use of AI in security operations, and a third-party risk management (TPRM) program, which is increasingly crucial given the complexity of modern digital supply chains (A. Mishra et al., 2022).

The Architecture & Technology layer implements Security-by-Design principles throughout the system development lifecycle. Key elements include adopting Zero Trust Architecture as a network design paradigm, implementing end-to-end encryption and privacy-enhancing technologies (PETs) for data protection, implementing AI-powered security tooling integrated into the SOAR platform, and designing resilient systems with built-in redundancy and self-healing capabilities. Research shows that organizations that integrate security from the design phase (shift-left security) can reduce vulnerability remediation costs by up to 30 times compared to those that address security late in development (Patel & Patel, 2022).

The Operations & Intelligence layer is the core of the real-time security response within the KSTKS. Its key components are an AI-enhanced Security Operations Center (SOC) for advanced threat analytics, a Threat Intelligence program integrated with various internal and external sources, a standardized and automated Incident Response process through the SOAR platform, and ongoing Red Team and Purple Team exercises to empirically test and improve defense effectiveness. The integration of all these components results in a much faster and more effective detect-and-respond capability (Tounsi & Rais, 2022)

The Human Capital & Culture layer addresses the reality that the human factor remains both the greatest weakness and the most important asset in the cybersecurity ecosystem. Investments in ongoing security training and certification programs, the development of a cybersecurity culture that prioritizes awareness and vigilance at all levels of the organization, a Cybersecurity Champion program that distributes security responsibilities across business units, and regular phishing simulation programs have been shown to significantly reduce the risk of social engineering attacks that exploit human weaknesses (Singh, 2025)

The recommended implementation of the KSTKS follows a three-phase roadmap spanning 18-36 months. The first phase (0-6 months) focuses on comprehensive risk assessments, building a governance foundation, and deploying high-priority technical solutions such as ML-IDS and UEBA. The second phase (6-18 months) includes architectural transformation toward ZTA, enhancing SOC capabilities with AI, and strengthening human resource development programs. The third phase (18-36 months) aims to achieve full maturity with holistic threat intelligence integration, comprehensive incident response automation, and the development of a security culture embedded in the organization's DNA (Edo et al., 2022; Keshav Jena, 2023)

The success of the KSTKS implementation is measured using a multi-dimensional metrics dashboard that includes technical indicators (Mean Time to Detect/MTTD, Mean Time to Respond/MTTR, false positive rate), financial indicators (Return on Security Investment/ROSI, incident costs), compliance indicators (regulatory compliance rate, audit findings), and maturity indicators (security maturity score based on CMMI-Security). Regular evaluation of these metrics allows for agile and evidence-based strategy adjustments, ensuring the relevance and effectiveness of the KSTKS amidst the ongoing threat evolution (A. Mishra et al., 2022)

This research also identified that ecosystem collaboration between industry, government, and academia is a critical, often overlooked enabler in cybersecurity implementation. The Information Sharing and Analysis Centers (ISACs) program, which has proven effective in the banking and critical infrastructure sectors, needs to be expanded to other sectors increasingly vulnerable to cyberattacks. In Indonesia, the establishment of the National Cyber and Crypto Agency (BSSN) is a step in the right direction, but closer coordination between the BSSN, the private sector, and educational institutions is needed to build a truly resilient national cybersecurity ecosystem (Lim et al., 2019)

Adequate and structured cybersecurity funding is an essential prerequisite for implementing a cybersecurity strategy. Research shows that organizations that allocate 10-15 % of their total IT budget to cybersecurity demonstrate significantly greater resilience to attacks than those allocating less than 5%. More importantly, a funding model that integrates cybersecurity as a strategic business investment—rather than simply an operational expense—positively impacts the organizational security culture and leadership commitment necessary for the long-term success of a cybersecurity strategy (Bhardwaj et al., 2020).

Table 4. Integrated Cyber Security Strategic Framework (KSTKS)

Layer	Main Components	Key Technologies	Success Indicators
Governance & Compliance	Risk-based Policy, AI Governance, TPRM	GRC Platform, AI Act	Compliance Rate >95%
Architecture & Technology	ZTA, PETs, Security-by-Design	ZTA Tools, FHE, SOAR	MTTD <1 hour
Operations & Intelligence	AI-SOC, CTI, Red/Purple Team	ML-IDS, UEBA, SOAR	MTTR <4 hours
Human Capital & Culture	Training, Champions, Phishing Sim	LMS, Awareness Platform	Error Rate <5%

Source: Results of research framework development, 2024

CONCLUSION

This research successfully achieved all three of its stated objectives. First, a systematic analysis of 87 scientific articles identified and classified seven major categories of cybersecurity challenges in the era of big data and AI, with adversarial attacks against AI models (32.4%), massive data breaches (28.7%), and AI-based ransomware (21.5%) as the three most dominant and potentially serious threats. This threat landscape is highly dynamic and continues to evolve as AI technology advances, demanding a defense approach that is equally adaptive and data-driven. These findings confirm that cyberthreats in the AI era are fundamentally different from conventional threats and require a new framework for analysis and response.

Second , this study successfully analyzed the effectiveness of various available cybersecurity solutions and found that AI-based approaches—specifically ML-IDS with 94-99% accuracy, UEBA for insider threat detection, and the AI-SOAR platform capable of reducing MTTR by up to 85%—consistently outperform conventional rule-based solutions in dealing with complex and novel threats. However, the effectiveness of these technical solutions can only be optimally realized when integrated within a strong governance framework, supported by adequate human resource investment, and operated within an active threat information collaboration ecosystem. Third, this study successfully formulated an Integrated Cybersecurity Strategic Framework (KSTKS) consisting of four layers—Governance & Compliance, Architecture & Technology, Operations & Intelligence, and Human Capital & Culture—that provides comprehensive and scalable implementation guidance for organizations of various sizes and sectors. KSTKS is the main novelty contribution of this study, which is expected to fill the gap identified in previous literature, namely the lack of an integrative framework that simultaneously addresses the technical, strategic, regulatory, and human dimensions of cybersecurity challenges in the era of big data and AI.

REFERENCES

Abdelmagid, A. M., Ezell, B. C., & McShane, M. K. (2026). *Toward Zero Trust Architecture Implementation in SMBs: A Conceptual Framework and Thematic Propositions*. https://doi.org/10.2139/ssrn.6328741

Acar, A., Aksu, H., Uluagac, A. S., & Conti, M. (2018). A survey on homomorphic encryption schemes: Theory and implementation. *ACM Computing Surveys*, *51*(4), 1–35. https://doi.org/10.1145/3214303

Amarasinghe, K., Kenaza, T., & Wijekoon, J. (2019). Explainable AI for cybersecurity: Challenges and opportunities. *IEEE International Conference on Cybersecurity and Protection of Digital Services*, 1–8.

Apruzzese, G., Colajanni, M., Ferretti, L., & Marchetti, M. (2022). Addressing adversarial attacks against security systems based on machine learning. *Conference on Communications and Network Technologies (CCNT)*, 1–8. https://doi.org/10.1109/CNS56114.2022

Asogwa, C. E. (2020). Internet-Based Communications: A Threat or Strength to National Security? *SAGE Open*, *10*(2). https://doi.org/10.1177/2158244020914580

Bhardwaj, A., Al-Turjman, F., Kumar, M., Stephan, T., & Mostarda, L. (2020). Capturing-the-invisible (CTI): Behavior-based attack recognition in IoT-oriented industrial control systems. *IEEE Access*, *8*, 104956–104966. https://doi.org/10.1109/ACCESS.2020.2998990

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brewer, R. (2022). Ransomware attacks: Detection, prevention and cure. *Network Security*, (3), 5–9. https://doi.org/10.1016/S1353-4858(22)70032-3

Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for

cyber security intrusion detection. *IEEE Communications Surveys & Tutorials*, *18*(2), 1153–1176. https://doi.org/10.1109/COMST.2015.2494502

Chen, B., Carvalho, W., Baracaldo, N., Ludwig, H., Edwards, B., Lee, T., Molloy, I., & Srivastava, B. (2019). Detecting backdoor attacks on deep neural networks by activation clustering. *CEUR Workshop Proceedings*, *2301*.

Cheng, L., Liu, F., & Yao, D. (2022). Enterprise data breach: Causes, challenges, prevention, and future directions. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *7*(5), e1211. https://doi.org/10.1002/widm.1211

Chesney, R., & Citron, D. (2019). Deepfakes and the new disinformation war: the coming age of post-truth geopolitics. *Foreign Affairs*, *98*(1).

C.Kanmani Pappa. (2024). Zero-Trust Cryptographic Protocols and Differential Privacy Techniques for Scalable Secure Multi-Party Computation in Big Data Analytics. *Journal of Electrical Systems*, *20*(5s). https://doi.org/10.52783/jes.2550

Edo, O. C., Tenebe, T., Etu, E., Ayuwu, A., Emakhu, J., & Adebiyi, S. (2022). Zero Trust Architecture: Trend and Impacton Information Security. *International Journal of Emerging Technology and Advanced Engineering*, *12*(7). https://doi.org/10.46338/ijetae0722_15

Findley, M. G., Ghosn, F., & Lowe, S. J. (2024). Vulnerability in research ethics: A call for assessing vulnerability and implementing protections. *Proceedings of the National Academy of Sciences of the United States of America*, *121*(34). https://doi.org/10.1073/pnas.2322821121

Goodfellow, I. J., Shlens, J., & Szegedy, C. (2014). Explaining and harnessing adversarial examples. *ArXiv Preprint ArXiv:1412.6572*.

Goodman, B., & Flaxman, S. (2017). European union regulations on algorithmic decision making and a "right to explanation." *AI Magazine*, *38*(3). https://doi.org/10.1609/aimag.v38i3.2741

Hoofnagle, C. J., Sloot, B. van der, & Borgesius, F. Z. (2019). The European Union general data protection regulation: What it is and what it means. *Information and Communications Technology Law*, *28*(1). https://doi.org/10.1080/13600834.2019.1573501

Hosein, G. (2019). Privacy as Freedom. In *Human Rights in the Global Information Society*. https://doi.org/10.7551/mitpress/3606.003.0008

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Keshav Jena. (2023). Zero-Trust Security Models Overview. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*. https://doi.org/10.32628/cseit2390578

Kindervag, J., Balaouras, S., & Coit, L. (2010). No more chewy centers: Introducing the zero trust model of information security. In *Forrester Research, Inc., Cambridge, MA* (Vol. 3).

Kindervag, J., Balaouras, S., Mak, K., & Blackborow, J. (2016). No More Chewy Centers: Introducing The Zero Trust Model Of Information Security. In *Forrester Research, Inc.*

Lim, M., Abdullah, A., Jhanjhi, N. Z., & Supramaniam, M. (2019). Hidden link prediction in criminal networks using the deep reinforcement learning technique. *Computers*, *8*(1). https://doi.org/10.3390/computers8010008

Lotfollahi, M., Jafari Siavoshani, M., Shirali Hossein Zade, R., & Saberian, M. (2020). Deep packet: a novel approach for encrypted traffic classification using deep learning. *Soft Computing*, *24*(3). https://doi.org/10.1007/s00500-019-04030-2

Matthew Collins, Theis, M., Trzeciak, R., Moore, A., Costa, D., Cassidy, T., Clark, J., Albrethsen, M., & Strozer, J. (2018). Common Sense Guide to Mitigating Insider Threats. *CERT Division*, *6*(1).

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys*, *54*(6), 1–35.

Mishra, A., Alzoubi, Y. I., Anwar, M. J., & Gill, A. Q. (2022). Attributes impacting cybersecurity

policy development: Evidence from seven nations. *Computers & Security*, *120*, 102820.
Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge. *Teachers College Record*, *108*(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x
Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., Antes, G., Atkins, D., Barbour, V., Barrowman, N., Berlin, J. A., Clark, J., Clarke, M., Cook, D., D'Amico, R., Deeks, J. J., Devereaux, P. J., Dickersin, K., Egger, M., Ernst, E., Gøtzsche, P. C., … Tugwell, P. (2014). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Revista Espanola de Nutricion Humana y Dietetica*, *18*(3). https://doi.org/10.14306/renhyd.18.3.114
National Institute of Standards and Technology. (2020). Zero Trust Architecture - NIST Special Publication 800-207. *NIST*.
Page, M. J., McKenzie, J. E., Bossuyt, P. M., & others. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372*, n71.
Patel, M., & Patel, N. (2022). Security orchestration, automation and response (SOAR): Current status and future directions. *Journal of Network and Computer Applications*, *210*, 103545.
Phiayura, P., & Teerakanok, S. (2023). A Comprehensive Framework for Migrating to Zero Trust Architecture. *IEEE Access*, *11*. https://doi.org/10.1109/ACCESS.2023.3248622
Reinsel, D., Gantz, J., & Rydning, J. (2018). The Digitization of the World - From Edge to Core. *Framingham: International Data Corporation*, (November).
Shu, X., Zhang, J., Ye, D., & Liu, J. (2023). Android malware detection based on semantic analysis of application behavior. *IEEE Transactions on Dependable and Secure Computing*, *20*(1), 127–142.
Singh, A. (2025). From Past to Present: The Evolution of Data Breach Causes (2005–2025). *LatIA*, *3*. https://doi.org/10.62486/latia2025333
Sirur, S., Nurse, J. R. C., & Webb, H. (2018). Are we there yet? Understanding the challenges faced in complying with the General Data Protection Regulation (GDPR). *Proceedings of the ACM Conference on Computer and Communications Security*. https://doi.org/10.1145/3267357.3267368
Sun, J., Yao, X., Wang, S., & Wu, Y. (2020). Blockchain-Based Secure Storage and Access Scheme for Electronic Medical Records in IPFS. *IEEE Access*, *8*. https://doi.org/10.1109/ACCESS.2020.2982964
Theis M., Trzeciak R., Costa D., Moore A., Miller S., Cassidy T., & Claycomb W. (2022). *Common Sense Guide to Mitigating Insider Threats, Seventh Edition*. Common Sense Guide to Mitigating Insider Threats, Seventh Edition.
Tounsi, W., & Rais, H. (2022). A survey on technical threat intelligence in the age of sophisticated cyber attacks. *Computers & Security*, *72*, 212–233.
Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. In *British Journal of Management* (Vol. 14, Number 3). https://doi.org/10.1111/1467-8551.00375
Ullah, I., & Mahmoud, Q. H. (2022). Design and Development of RNN Anomaly Detection Model for IoT Networks. *IEEE Access*, *10*. https://doi.org/10.1109/ACCESS.2022.3176317
Veale, M., & Zuiderveen Borgesius, F. (2021). Demystifying the Draft EU Artificial Intelligence Act — Analysing the good, the bad, and the unclear elements of the proposed approach. *Computer Law Review International*, *22*(4). https://doi.org/10.9785/cri-2021-220402
Wang, Y., Li, W., Sarkar, E., Shafique, M., Maniatakos, M., & Jabari, S. E. (2024). A Subspace Projective Clustering Approach for Backdoor Attack Detection and Mitigation in Deep Neural Networks. *IEEE Transactions on Artificial Intelligence*, *5*(7). https://doi.org/10.1109/TAI.2024.3373720