Received: May 6, 2026; Revised: August 4, 2026; Accepted: September 10, 2026; Published: October 30, 2026

The Role of Artificial Intelligence in Predicting and Mitigating Cyber Threats in Philippine Government Institutions

https://doi.org/10.58870/berj.v11i1.96

Roy R. Consulta
College of Arts and Sciences
San Beda University, Manila, Philippines
rconsulta@sanbeda.edu.ph

Peter Paul A. Ocampo
College of Arts and Sciences
San Beda University, Manila, Philippines
pocampo@sanbeda.edu.ph

Abstract

This paper studied the growing complexity of cyber threats facing the Philippine government and the role of artificial intelligence (AI) in combating cybercrime. AI can support cybersecurity by detecting and predicting technology-based threats and implementing automated mitigation strategies. Aligned with United Nations Sustainable Development Goal 16, which promotes peace, justice, and strong institutions, the paper explores AI-driven approaches to protecting public organizations from cyberattacks. It examines AI's role in characterizing threat intelligence and managing response times, as well as its potential to reduce false positives while addressing data privacy concerns. The paper also considers adversarial attacks, vulnerabilities in AI systems, and the need for people-centered controls. By evaluating AI's effectiveness in cybersecurity, the study considers how public organizations can be protected against evolving cyber risks while advancing SDG 16's aim of establishing strong and secure technological structures.

Keywords: cybersecurity, SDG#16, AI, government institutions, Sustainable Development Goals

Introduction

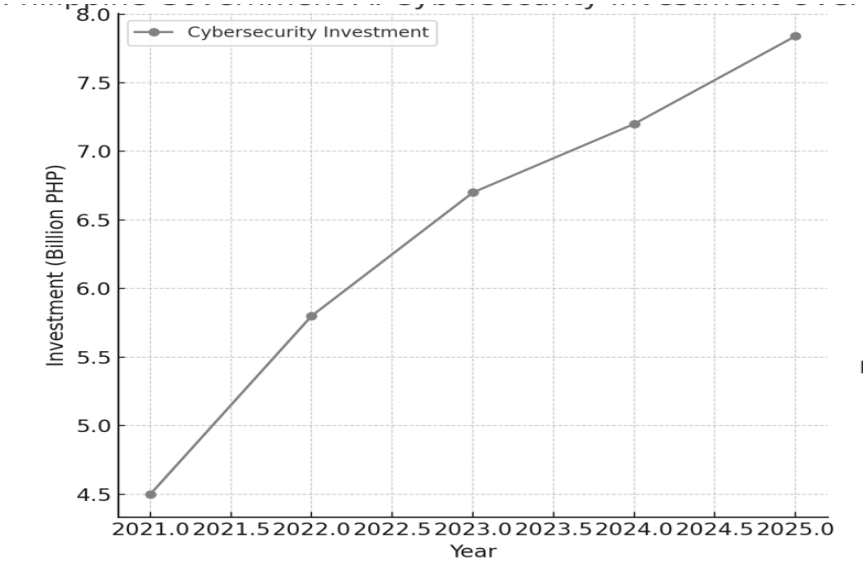
The Philippines present administration eagerly wanted to utilize and employ robust digital transformation initiatives (Flores, 2024). As a result of the creation of artificial intelligence (AI) and digitization, AI is emphasized as the main form of economic development (Bolivar, 2024). Extensive experience and promising developments reveal the state of Artificial Intelligence (AI). The government has analyzed the potential benefits of artificial intelligence for growth and economic diversification, implementing initiatives such as the National AI Roadmap and establishing a National AI Research Center to prioritize AI adoption and research (ITA, 2024). The concept aims to enhance the country's AI potential, boost startups, and integrate AI into various sectors, including healthcare, agriculture, and finance.

According to statistics, Artificial Intelligence (AI) will be the key to improving cybersecurity within Philippine government institutions. In the first quarter of 2024, the Philippines experienced more than four times the amount of malicious activity on the Internet and cyberworld compared to the same period in 2023, with an average of 8,800 attacks daily due to the increase in cyber threats (Cadiang, 2024). Furthermore, regarding budget government initiatives and investment, the 2024-2025 Philippine government budget indicates a significant commitment to cybersecurity, with the Department of Information and Communications Technology (DICT) (Uy & Dy, 2024) receiving a 49.2% increase, totaling PHP 7.84 billion for 2025.

Here are the financial graphs illustrating AI-driven cybersecurity investment in the Philippine government (Uy & Dy, 2024):

Figure 1.

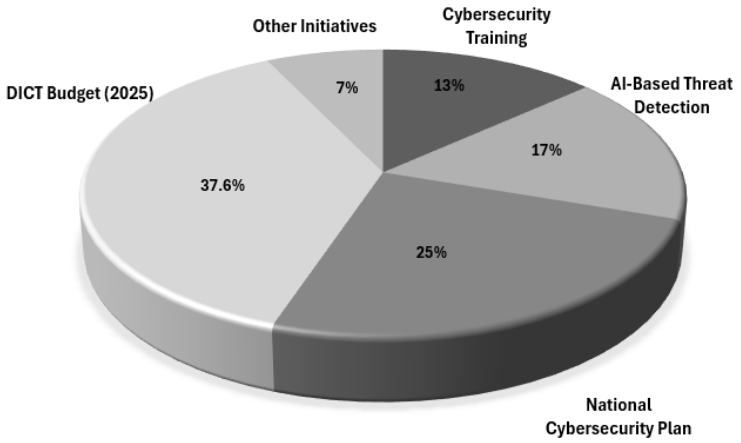
Philippine Government AI Cybersecurity Investment



Shows the increasing trend of government investment in cybersecurity from 2021 to 2025 over time.

Figure 2.

Target Allocation for AI Cybersecurity in the Philippines (2025)



In terms of emerging threats and advancing its artificial intelligence, the Philippines is encountering several challenges and vulnerabilities (DICT Cybersecurity Bureau, 2024); over 84% of Philippine organizations reported being negatively impacted by cybersecurity breaches within their supply chains in 2024, with an average of 3.13 breaches affecting operations (BlueVoyant, The State of Supply Chain Defense: Annual Global Insights Report, 2025). In addition, the rise of AI-powered scams, such as "pig butchering," has been noted, where trafficked individuals are coerced into conducting investment scams. In 2024, Meta removed over 2 million scam-connected accounts from various countries, including the Philippines (Newman & Burgess, 2024). Primary concerns include constrained infrastructure and resources, which can deter the deployment and scaling of AI technologies. Furthermore, there is a shortage of skilled professionals with expertise in AI, creating a talent gap that impacts the ability to develop and implement sophisticated AI solutions. The integration of AI into the cybersecurity infrastructure of most Philippine government institutions is still in its early stages, as they primarily rely on outdated methods for securing their systems, focusing on remediation rather than prevention and quick mitigation (Uy & Dy, 2024). With the current evolution of cyber threats, the focus should be shifted to continuous evaluation of AI-driven cybersecurity implementations, evaluating challenges, and formulation of optimal strategies for cybersecurity adaptability (Luna, 2024).

In this network and social media age, the Philippines' public organizations experience rapid cybersecurity threats, including ransomware, phishing, and advanced persistent threats (APTs). Integrating Artificial Intelligence (AI) in cybersecurity is considered a major upside for organizations that would like to be able to predict and mitigate cybersecurity issues (Kaur, Gabrijelcic, & Klobucar, Artificial intelligence for cybersecurity: Literature review and future research directions, 2023).

By doing research on how AI can be fused into the essential components like network management, encryption, response management, identity access, and regulatory compliance, the study seeks to contribute to the movement geared towards developing more flexible and adaptive cybersecurity frameworks for the Philippine government institutions where there will be special focus on how the government's digital assets may be further protected (Binhammad, Alqaydi, Othman, & Abuljadayel, 2024).

Furthermore, the paper aligns with UN SDG 16 (United Nations, 2023) (Peace, Justice, and Strong Institutions) given the special focus on developing sustainable and resilient institutions through AI-enhanced cybersecurity infrastructure (Dasandi & Mikhaylov, 2019).
Technology can be utilized to achieve objectives that extend beyond traditional business goals. Sustainable development also depends, fundamentally, on upholding human rights and ensuring peace and security (Edralin & Pastrana 2022).

Review of Related Literature

The increasing reliance on digital infrastructure has made government institutions prime targets for cyber threats. Artificial Intelligence (AI) is emerging as a transformative tool in cybersecurity, offering advanced capabilities in threat prediction, detection, and mitigation (Asmar & Tuqan, 2024).

This review provides emphasis in looking at current literature about AI applications in cybersecurity. An example of this is in authentication, where AI can be used to validate user access through comparing access with historical behavioral biometric profiles, which is considered a major upgrade compared to traditional authentication done via password entries (Jongvibool, 2025).

In the modern interconnected world, cyber threats are advancing at an unparalleled speed, aiming at governments, corporations, and individuals equally. Conventional security systems, dependent on rule-based identification, find it challenging to cope with advanced cyberattacks. This area is where artificial intelligence, or AI, acts as a transformative force, changing the way we forecast, identify, and reduce cyber threats before they inflict damage (Ali, Murray, Momin, Dwivedi, & Malik, 2024).

Machine learning (ML) algorithms, such as supervised and unsupervised learning, have demonstrated significant potential in predicting cyber threats. Research by Sommer & Paxson (2010) highlights that anomaly detection techniques in ML can identify suspicious patterns in network traffic, reducing the risk of undetected cyberattacks (Sharma, Sharma, & Jindal, 2021). Deep learning models, such as recurrent neural networks (RNNs) and convolutional neural networks (CNNs), have also been applied to malware detection and behavior analysis (Sharma & Chen,

2021). For years, cybersecurity has primarily been reactive, addressing attacks after they take place. Nonetheless, AI-driven threat forecasting represents a major transition toward a proactive defense approach. By using machine learning, big data, and behavior analytics, AI can foresee the threats before they can infiltrate where risks are minimized and response times significantly improved (Luckin, Cukurova, Kent, & du Boulay, 2022). The current developments in cybersecurity have necessitated the formulation of predictive security models that leverage AI and machine learning, adopting a proactive approach that empowers organizations through advanced detection. AI-powered cybersecurity technologies focus on using big data and analyzing patterns, thereby promoting threat mitigation (Buczak & Guven, 2016). Moreover, Natural Language Processing (NLP) has also been utilized to monitor communications on the dark web to focus on the detection of threats at the early stages (Temara, 2024). Security Orchestration, Automation, and Response (SOAR) that is AI powered help in real-time mitigation of threats (Kinyua & Awuah, 2021) and shows that reactions to threats can be automated. IBM Watson, an independent cybersecurity framework, uses AI for making threat response efficient.

In a recent study, malware detection systems using AI significantly improve antivirus tool performance by being able to scan novel malware agents. Generative adversarial networks (GANs) are utilized to increase the level of detection capabilities using simulations of common attack behavior to improve AI training.

Despite the known benefits of AI in cybersecurity, certain emerging issues arise in this paradigm, particularly adversarial AI threats that manipulate AI models to exploit private data. Assuring data protection and regulation compliance is a cause of concern given that AI may also be used to enhance cyber attacks themselves. For example, deepfake-based phishing uses deepfake tools to mislead cyber users into sharing sensitive information (Brundage et al. 2018).

As a result, compliance to regulations and data protection continue to become critical areas to be prioritized as institutions implement AI-enhanced cybersecurity methods as they continue to align their respective missions and visions with sustainability. This is because sustainable business practices always go with the quest for peace. Its contribution is through their compliance with the local and international laws and other standards to operate their business with transparency, accountability, and

responsibility to create a sound business environment (Edralin and Pastrana 2020).

Adversarial machine learning (AML) is another emerging problem where AI models are manipulated to be able to bypass threat detection systems. In a recent study, small misalignments in datasets can lead to possible AI misclassifications, which can cause reliability issues among AI-based cyber threat solutions (Josyula & Saidireddy, 2024). Adversarial training and more adaptive AI architectures are currently being enhanced to counter these.

Research has shown that AI-driven cybersecurity is considered a major upgrade compared to rule-based systems in the aspects of response timing and flexibility (Kaur, Gabrijelčič, & Klobučar, 2023). However, approaches that mix AI with human insights are often suggested to harmonize automation with inaccuracies (Jiang, Sun, Fu, & Lv, 2024). Further research is needed to improve AI models, upgrade interpretability, and formulate resilient-based techniques to counter the evolving cyber threats.

Statement of Research Problem

Artificial intelligence is an essential tool to go against cyber threats, thereby improving to security and adaptability of institutions, which aligns with United Nations SDG 16 (Herwig, 2015). Through strategic integration of AI, governments and organizations can establish a secure and stable digital environment aligned with sustainability aims (Mondejar et al., 2021).

Cyberattacks in government institutions have become more rampant, particularly more sophisticated ones, i.e., PhilHealth (NPC, 2024) and the Armed Forces of the Philippines (Sigales J. , 2025), and these range from basic ransomware to phishing and advanced persistent threats (APTs). As it is reported on the Insurance Business Magazine website (Libatique, 2024), PhilHealth, 42 million members are exposed to the data breach wherein about 181 million records were copied and exposed to the deep web. It was said that as per the DICT and PhilHealth, the said ransomware only affected the application servers and staff workstations inside the agency and did not infiltrate the central membership database, which contains core sensitive information like contributions, claims, and accreditation.

Although the said incident, which happened on September 22, 2023, did not affect central members database (Libatique, 2024), it reportedly accounted for around 700 GB exposure of personal data of members in the form of IDs, bank cards, and transaction receipts.

In terms of the AFP case, the Philippine Army, which is one of the strong forces within the Armed Forces of the Philippines, affirmed an "illegal access" attempt on its computer network (Ferreras, 2025). It was said that a certain group known as the Philippine Exodus Security (PHEDS) claimed they have infiltrated important systems inside the Department of National Defense and Philippine Army in the form of access to classified data related to the Army's Cybersecurity Battalion, including access to the Philippine Navy's mail system.

However, after a series of technical investigations, the Department of National Defense has declared that its systems were not exposed and remain intact, and the incident was classified as an unsuccessful attempt. (Kabagani, 2025)

In that sense, AI in cybersecurity presents an opportunity to upgrade predictive and proactive capabilities of cybersecurity platforms, but the thorough assessment of how effective AI cybersecurity solutions applied to government institutions' context is not yet fully explored at this point.
The evolving issues and complexities in cybersecurity drive government institutions, particularly in the Philippines have shown that as digital transformation is upgraded, cyberattacks, in the form of ransomware, data breaches, and nation-state cyber espionage, have also increased at alarming levels (Saeed, Altamimi, Alkayyal, Alshehri, & Alabbad, 2023).

The potential benefits of Artificial Intelligence (AI) in cybersecurity are currently well recognized, offering an array of capabilities in aspects like detection of threats, anomaly investigation, and automated mitigation. However, its integration into Philippine government institutions' security frameworks remains low because of concerns related to implementation, false positives, adversarial attacks, data privacy, and human oversight. Evaluating AI-powered threat detection and mitigation platforms in these contexts is still not fully explored.

It also seeks to evaluate the effectiveness of mitigating and predicting AI-driven threat intelligence, assess the challenges associated with AI adoption in government cybersecurity frameworks, and propose

strategies to enhance AI-powered security solutions. The study will also analyze the impact of AI on response times, false positives, and overall security posture while addressing ethical and governance considerations. By examining the factors consistently mentioned in this paper, this research seeks to reveal "*how AI can be mitigated and predicted in improving cyber resilience among Philippine government agencies with a focus on how critical infrastructure and sensitive data are protected and preserved."*

Statement of Specific Objectives

The research seeks to examine the role of Artificial Intelligence (AI) in areas of threat prediction and mitigation among government institutions. The specific objective of the study is to assess how the integration of AI-enhanced threat detection platforms can identify and predict cyber threats targeting government institutions.

Significance of the Study

The integration of artificial intelligence in cybersecurity serves as a paradigm shift into emerging techniques in threat detection, analysis, and mitigation under the context of government security frameworks. Through exploration of this context, this research aims to establish a more localized perspective that also connects global AI developments with Philippine government cybersecurity imperatives. It works toward the improvement of theoretical frameworks by contextualizing AI-driven cybersecurity frameworks and platforms in environments where resources are limited, structures are bureaucratic, and data governance is considered complex.

The study provides useful and up-to-date insights for policymakers, cybersecurity practitioners, and technology developers in the Philippines, given that government institutions are common targets for cyberattacks due to the sensitivity of public data and the nature of digital services. This research also seeks to identify the challenges in terms of data privacy, algorithmic transparency, and the need for human oversight in AI-based cybersecurity systems.

With consideration of both technical and governance concerns, the study aims to identify areas in terms of upgrading national cyber resilience as it supports the development of evidence-based techniques supporting

moves towards investing in AI-capable cybersecurity platforms and developing policies that blend both innovation and ethical factors. As such, it contributes to the larger objective of stabilizing public digital assets, thereby improving public trust in government digital services.

The following recent initiatives and case studies serve as particularly relevant examples to describe and document the known benefits of AI in cybersecurity within the Philippine context:

- Cybersecurity Management System of Department of Information and Communications Technology: The project focuses on the utilization of AI-driven data analytics integratedmonitore CMS to check on unusual network behavior in government agencies. With the use of machine learning models, the system can improve anomaly detection, helping prevent possible breaches via identification of patterns that detect malware and data exfiltration threats (Asia, 2019).

- Anti-Cybercrime Group of the Philippine National Police: The project was built on partnerships with private cybersecurity firms to implement AI-powered tools for cybercrime investigation. The said tools integrated Natural Language Processing (NLP) and behavioral analytics to monitor phishing activities and detect digital fraud on social media platforms and electronic financial systems (Sigales J., 2024).

- Advanced Science and Technology Institute of Department of Science and Technology – The project explored the use of AI in monitoring of network intrusions and incident tracking, where locally trained AI models were used in government networks to automatically react to evolving threats that include ransomware and zero-day exploit attacks (Luci-Atienza, 2021).

- Cybersecurity Framework for Financial Institutions of the Bangko Sentral ng Pilipinas: The project covers guidelines on how local banks can implement AI systems to predict and mitigate fraud and secure online transactions.

These government initiatives demonstrate that AI is beginning to be utilized in modernizing government cybersecurity frameworks and platforms and highlight how public-private joint initiatives with sufficient government support can make projects of this nature successful.

With the analysis of projects like these, the study provides clear evidence on how AI can be factored into cybersecurity across other government institutions, as it supports the development of evidence-based methods and is partnered with policies and investments that promote balance, innovation, security, and improvement of public trust.

With these, the research also contributes not only to the academic body of knowledge but also to the practical upgrading of digital governance in the Philippines. Overall, the study seeks to promote safer digital ecosystems that are in line with ongoing digital transformation and national development goals.

Methodology

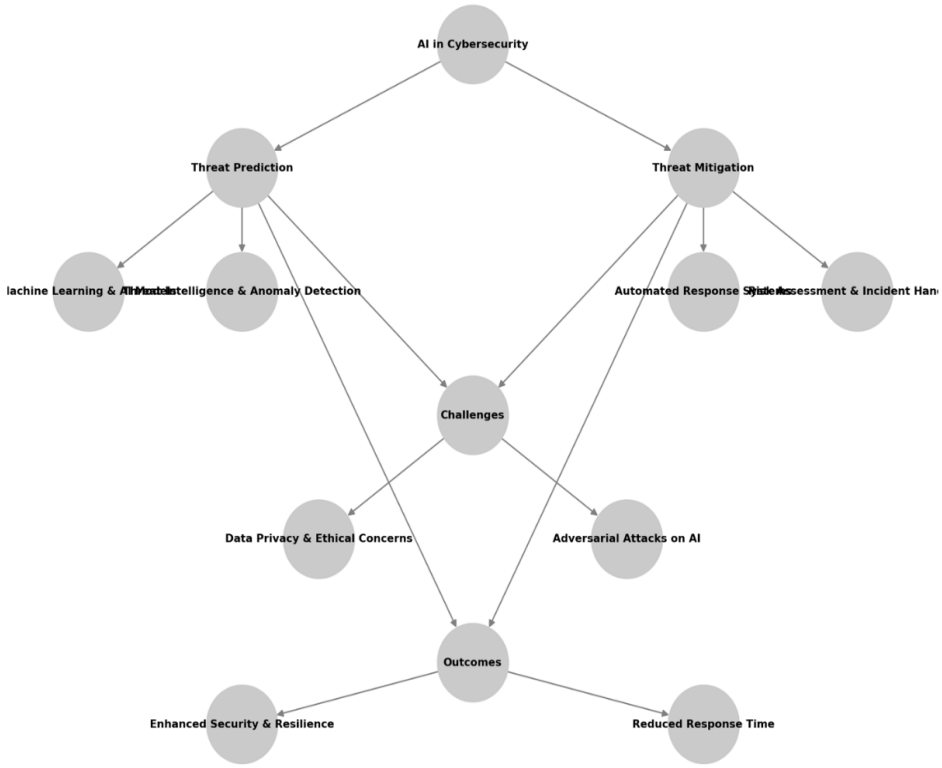
Conceptual Framework

The integration of Artificial Intelligence (AI) in cybersecurity is built upon known theoretical and philosophical points of view that are used in its continued development, implementation, and ethical evolution. This section investigates the relevant theories and philosophical ideas connected to the study of AI-driven cyber threat predictive systems in government institutions.

The graphical model below shows how AI can be used to predict and mitigate cyber threats in the context of government institutions. The said model presents the links of predictive AI, mitigation methods, challenges, and expected outcomes.

Figure 3.

Graphical model illustrating how AI contributes to predicting and mitigating cyber threats in government institutions.

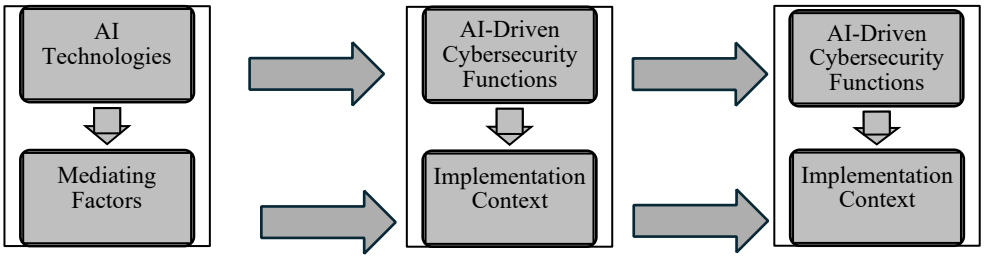


These foundations show that AI-driven threat prediction and mitigation are directly connected to computational-based learning methods, strategies that are adversarial in nature, ethics, and sociotechnical factors. Harmonizing technology efficiency with ethics is also said to be crucial for establishing sound AI cybersecurity platforms and frameworks in government institutions.

This conceptual framework depicts the link between the integration of artificial intelligence and how it affects cyber threat prediction and mitigation within Philippine government institutions. It is built upon systems theory and security models that are adaptive, acknowledging the fact that cybersecurity in public institutions may be classified as a dynamic system directly evolved by modern technology, adaptive policy, and shaped by human factors.

Figure 4.

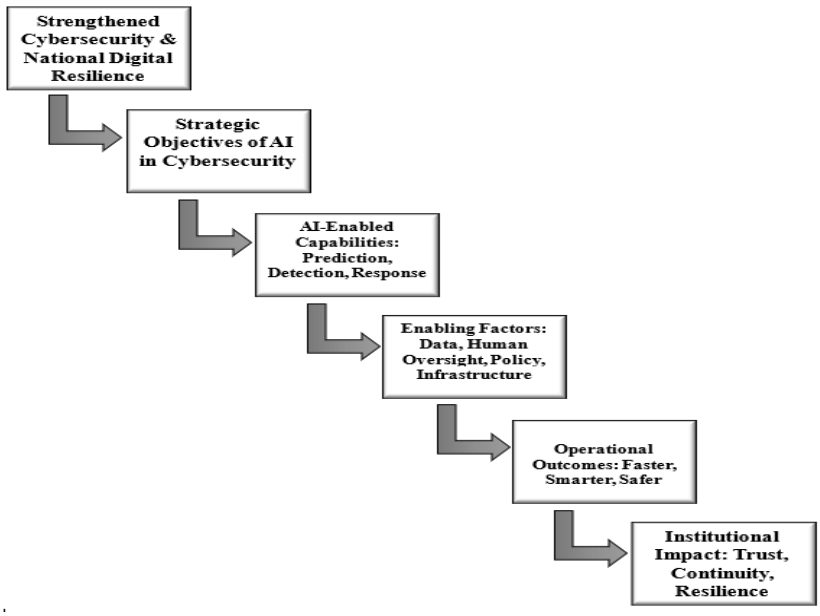
Conceptual Framework Integration of Artificial Intelligence



The overall goal is the protection of government digital assets and public data, including the enhancement of public trust in online services with the strategic implementation of artificial intelligence. With real-time threat detection and quick mitigation, these efforts fortify IT infrastructure against attacks and improve the confidentiality, integrity, and availability of sensitive information. Furthermore, the integration of AI in cybersecurity pushes for transparency and accountability, reinforcing confidence in the use of the government's digital platforms.

Figure 5.

Integration of AI in Cybersecurity



Sources of Data (Secondary)

This study also utilizes secondary data sources to establish a wider contextual and theoretical framework for analyzing the role of artificial intelligence in cybersecurity. The said sources seek to frame a local cybersecurity picture combined with global and regional innovations, offering comparative insights to formulate optimal practices and address implementation challenges.

Government policy documents and strategic frameworks are among the main sources. An example is our *National Cybersecurity Plan 2022* by the Department of Information and Communications Technology that covers strategic priorities for securing and stabilizing important IT infrastructure (DICT Cybersecurity Bureau, 2024). It also provides valuable information and guidelines on how AI can be integrated into the overall national cybersecurity strategy. Similarly, the *National Artificial Intelligence Strategy Roadmap* of the Department of Trade and Industry (DTI) provides foundations for understanding the policy direction on the use of AI and its integration into public sector digital security.

Also, frameworks and legal instruments, like our *Data Privacy Act of 2012* and regulatory compliance by the National Privacy Commission, help frame the ethical and legal aspects behind the use of AI in the processing of sensitive data, paying importance to transparency, data governance, and accountability (NPC, 2024).

This study also explored the existing array of academic literature and peer-reviewed papers related to artificial intelligence, cybersecurity, and data governance. Articles from known repositories: JSTOR, SpringerLink, IEEE, and ScienceDirect is where the theoretical points of view and technical considerations on the use of AI-based threat detection systems were based.

Industry research, product explorations, and white papers published by big cybersecurity companies like FireEye, IBM, Kaspersky, CISCO, Microsoft, and Palo Alto Networks provide modern technical points of view and case studies. Consistently looking at these sources is essential in monitoring new cyber threats, learning about AI-based defense methods, and including global benchmarks for evaluating the cybersecurity posture of organizations, as they contain regional data specific to Southeast Asia

and the Philippines, which is critical in analyzing local context versus global best practices.

We also use publicly available data on cyber incidents to analyze real-world patterns and trends. This includes reports from the Computer Emergency Response Team – Philippines (CERT-PH), advisories from DICT's Cybersecurity Bureau, and publicly reported cases of cyberattacks against government agencies as documented by news outlets and online threat monitoring platforms. These sources provide empirical examples of how AI could potentially improve early warning systems, detect anomalous behaviors, and mitigate cyber threats in real-time.

By synthesizing these various forms of secondary data, the study gains a multi-dimensional view of the existing capabilities, gaps, and opportunities for AI-enhanced cybersecurity in Philippine government institutions. This rich body of evidence supports the interpretation of primary data and ensures that the study is informed by both grounded local realities and global technological developments.

Propositions and assumptions

Propositions (for Secondary Data)

The following propositions are explored by analyzing secondary data, such as case studies, government reports, and white papers:

- **P1:** Secondary case reports suggest that AI improves cybersecurity posture through enhanced detection of previously unknown or sophisticated threats.
- **P2:** National policy documents and roadmaps imply that AI is a strategic enabler in achieving secure digital governance in the Philippines.
- **P3:** Government white papers and audits reveal that successful implementation of AI tools in cybersecurity requires consistent leadership, inter-agency collaboration, and capacity building.
- **P4:** Secondary data sources indicate that while AI shows strong potential, limitations in data quality and infrastructure readiness hinder its effectiveness in some government agencies.

Assumptions (About Secondary Data Sources)

Given the secondary data explored through reliable and effective frameworks, case studies, government reports, and white papers, we make several assumptions:

1. Data Authenticity and Credibility. It is assumed that the secondary data used—such as reports from DICT, DOST-ASTI, PNP-ACG, CERT-PH, and global cybersecurity firms—are accurate, published by reputable institutions, and free from intentional bias or misrepresentation.
2. Timeliness and Relevance. The study assumes that the selected secondary data sources are recent enough (from the past 5–7 years) to reflect current cybersecurity threats, trends, and AI implementations in government systems.
3. Comparability of Contexts. It is assumed that data from both Philippine agencies and international benchmarks can be meaningfully compared, especially where government digital transformation or cybersecurity maturity levels are similar.
4. Consistency in Terminology and Metrics. The research assumes that terms such as "AI implementation," "cyber threat," and "incident mitigation" are used consistently across documents or are at least clearly defined within each source.
5. Ethical Use and Accessibility. It is assumed that all secondary data used are legally accessible and ethically permissible for academic research, including government-published documents, open-access studies, and publicly available white papers.

Method of Data Collection

The study reviews and analyzes publicly available government documents, reports, and policy frameworks, such as DICT's *National Cybersecurity Plan 2022*, the *National AI Strategy Roadmap,* Republic Act No. 10173 (*Data Privacy Act of 2012*), and NPC issuances and cybersecurity regulations (DICT Cybersecurity Bureau, 2024) (NPC, 2024). These documents provide context on the institutional, legal, and policy environment surrounding AI and cybersecurity. The study examines case reports of past cyber incidents involving government agencies where AI was part of the detection or mitigation process. To understand AI's operational impact, we use system audit logs, breach assessments, or incident response evaluations (where available).

Cybersecurity firm white papers (e.g., by IBM, Kaspersky, Cisco, and Microsoft) and regional threat intelligence data offer empirical data and expert analysis relevant to attack trends in the public sector, AI tool performance metrics, and comparisons between traditional and AI-driven defenses.

Method of Data Analysis

Secondary data sources, such as government reports, policy documents, and cybersecurity audits, were analyzed using content analysis. This involves the identification of key references to AI and cybersecurity in policy documents (e.g., DICT's *National Cybersecurity Plan*). In addition, references were categorized into themes such as AI adoption, regulatory compliance, and cyber defense capabilities. Furthermore, analysis of how these documents position AI within broader government cybersecurity strategies and assess the alignment of AI initiatives with actual policy outcomes.

In addition, case studies of AI implementation in Philippine government agencies were examined to understand practical outcomes and lessons learned. The analysis focuses on the identification of the types of AI tools used (e.g., machine learning algorithms for anomaly detection and automated incident response systems). The review focuses on the outcomes of AI usage (e.g., reduction in cyberattacks, improved detection rates) and the context (e.g., agency size, infrastructure readiness, training) that contributed to success or failure.

Trend analysis was performed on historical data from threat intelligence reports and government cybersecurity audits. The analysis involved the identification of changes in cybersecurity outcomes over time (e.g., pre- and post-AI implementation) and the examination of how AI adoption trends correlate with shifts in threat landscapes (e.g., increase in AI-driven cyberattacks or improvements in defense).

SMSR (Systematic Mixed Studies Review) was used to analyze the references that were included in this study. The type of review is ideal given that a narrative review that does not rely on predetermined methods for synthesizing findings (Hong & Pluye, 2019). The methodology used also fits since it utilizes both qualitative and quantitative evidence in analyzing how AI can be used to predict and mitigate cybersecurity threats.

The inclusion and exclusion criteria for the data and sources included in this study are both very important for ensuring the validity, relevance, and reproducibility of the study. These help define which studies are relevant to the research questions and narrow down the volume of references by establishing a proper filtering mechanism.

Presented below are the categories of the references that were included in this study with a well-defined set of criteria.

Table 1.

Categories of the References

Category	Actual / Sample Documents	Inclusion Criteria	Exclusion Criteria
Government policy documents and strategic frameworks	● National Cybersecurity Plan 2022 ● National Artificial Intelligence Strategy Roadmap ● Data Privacy Act of 2012 ● Case studies of AI implementation in Philippine government agencies	Automatically pre-selected since these are the primary documents of the national government related to Cybersecurity and AI	Not applicable
Government case studies	● Department of Information and Communications Technology (DICT) – Cybersecurity Management System (CMS) ● Philippine National Police (PNP) Anti-Cybercrime Group (ACG) ● DOST-ASTI AI applications for network intrusion detection and incident response projects ● BSP's Cybersecurity Framework for Financial Institutions	Automatically pre-selected since these are the primary case studies / projects of the national government related to Cybersecurity and AI	Not applicable

Table 1.

Continued.

Category	Actual / Sample Documents	Inclusion Criteria	Exclusion Criteria
Academic literature and peer-reviewed journals	Examples: IEEE Xplore, ACM, and Springer.	Should contain significant discussion of the following with relevant data: ● Cybersecurity Predictive analysis ● Behavior-based risk detection ● Automated incident response	Unclear methodology and lack of supporting data
Whitepapers	From top technology companies like Amazon, IBM, and Microsoft.	Should contain up-to-date information on technologies being used for: ● AI-enabled cybersecurity defense mechanisms ● Threat intelligence ● Cybersecurity Benchmarks	Information and technology not up to date
Publicly available data on cyber incidents	From national government agencies: ● National Computer Emergency Response Team (CERT) reports ● DICT Advisory reports From industry recognized open threat monitoring platforms: ● Fortinet ● Palo Alto ● IBM	● Automatically pre-selected since these are the primary sources of national government data related to cybersecurity ● Contains data on the number and frequency of different types of cybersecurity attacks within the last 5 years ● Should have clear definition of different threat types	Lack of the ability to integrate the data Lack of similar categories

The reports and policy frameworks in the study related to the national government were directly taken from primary government documents and reports related to AI and cybersecurity. On the other hand, a manual literature search and data collection were conducted on the other sources through:

- **Academic Databases**: Google Scholar, Scopus, PubMed, Web of Science, and IEEE Xplore for academic articles.
- **Industry Reports**: White papers from cybersecurity firms or consultancies working with government agencies.
- **News Articles**: Articles detailing cybersecurity with AI deployments in the Philippine national government.

After conducting a manual literature search using reputable search engines and academic databases, the selected articles underwent full-text screening, which involved reading the entire text to assess their relevance and quality. We excluded studies that do not provide either qualitative or quantitative data or that are not focused on the Philippine government sector.

Mixed methods analysis was done on each of the specific categories, given the extent of variation in the approaches and methodologies among the studies. The table below presents a summary of the number of references for each category, including the actual analyses done for each.

Figure 6.

Number of References and Secondary Data

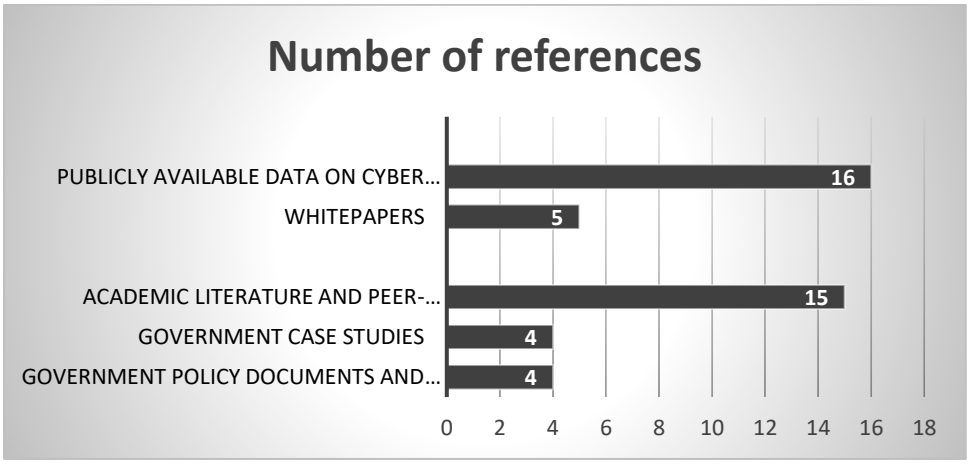


Table 2.

Category and Trend Analysis

Category	Type of analysis	Number of references
Government policy documents and strategic frameworks	Content Analysis	4
Government case studies	Review of context and outcomes	4
Academic literature and peer-reviewed journals	Content Analysis Outcomes Analysis	15
Whitepapers	Content Analysis	5
Publicly available data on cyber incidents	Trend analysis Correlation with threats landscape	16

Research Ethics Approach

Before data collection begins, the study undergoes a thorough ethical review process. The said method ensures that the research follows existing ethical and legal guidelines. The requirements of the review board of the host institution were strictly followed. Also, since the research used government documents related to AI and cybersecurity, approvals were obtained from the respective government agencies to ensure compliance with regulations, privacy rules, and standard protocols for accessing these documents.

Findings

Based on the datasets utilized in this study, which is through the secondary data consists of four (4) government policies, four (4) government case studies, fifteen (15) academic literature sources, five (5) industry white papers, and sixteen (16) cyber incidents reports, the interpretations and the major trend identified in the study are as follows:

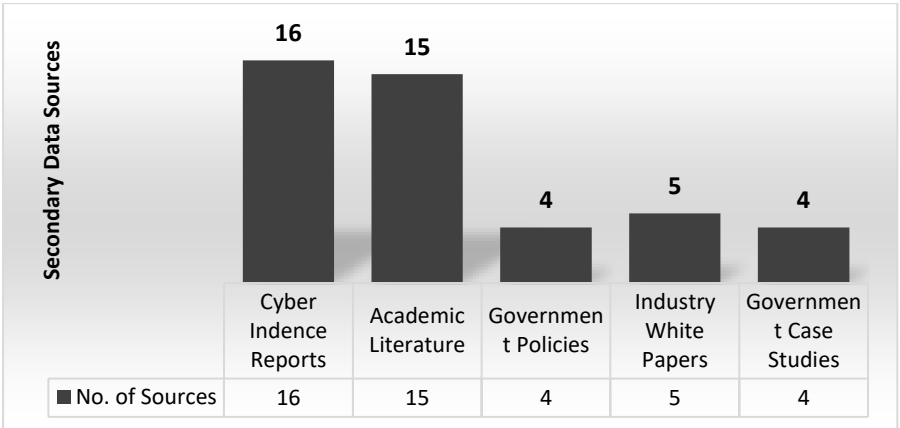
Table 3.

Interpretations and Trend Analysis Identification

Secondary Data Source	Numbers of Sources	Major Trend Identified	Interpretation
Government Policies	4	Increasing national support for AI-driven cybersecurity	Government policies consistently promote AI integration in protecting critical infrastructure and digital governance.
Government Case Studies	4	Growing adoption of AI-related applications	Selected departments and agencies are implementing Artificial Intelligence for intrusion detection, fraud detection, and cyber monitoring.
Academic Literature	15	Strong evidence supporting AI efficiency and effectiveness	Majority of studies report improvements in threat detection, malware analysis, and predictive security analytics.
Industry White Papers	5	Expansion of AI-powered cybersecurity solutions	Industry reports highlight automation, threat intelligence, and Security Operations Center (SOC) modernization and improvements.
Cyber Incidents Reports	16	Increasing sophistication of cyber threats	Reports indicate rising ransomware, phishing, malware, insider threats, and advanced persistent threats requiring AI-based defenses.

Figure 7.

Distribution of Secondary Data Sources



Interpretation

The biggest proportion of the secondary data came from publicly available cyber incident reports with 16 documents, followed by academic literature with 15 studies. These sources provide substantial empirical evidence regarding current cyber threats and AI-based cybersecurity practices. Government policies, government case studies, and industry white papers complement the analysis by providing strategic, operational, and industry perspectives.

Trend Analysis Discussion

The trends analysis shows a clear convergence of evidence across all secondary data sources. Academic literature, case studies, government policies, and industry reports consistently identify artificial intelligence as a transformative technology for strengthening cybersecurity in the Philippine bureaucratic institutions. The most frequently occurring theme, as shown in Table 4 and represented in Figure 8, is that AI enhances cyber threat detection, indicating a broad consensus that AI-based systems outperform traditional rule-based security approaches in identifying sophisticated and emerging attacks.

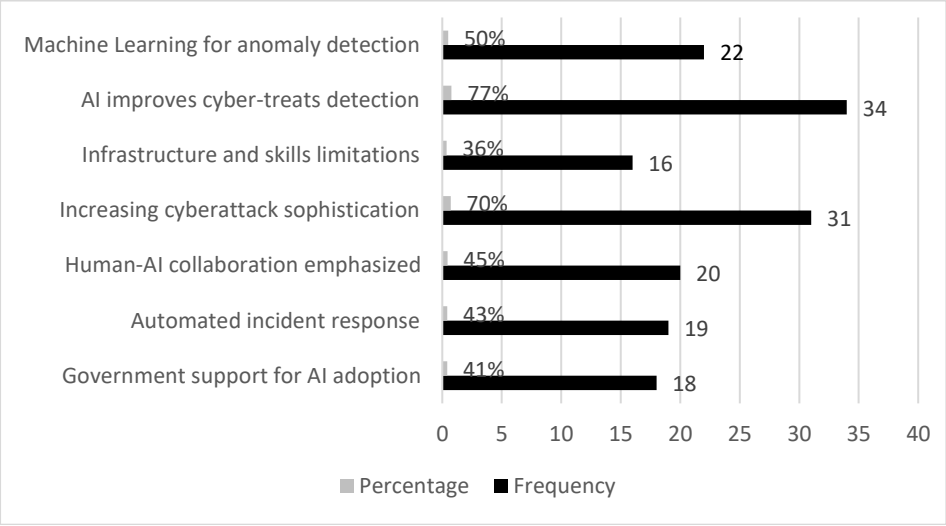
Table 4.

Emerging trends identified from secondary data

Emerging Trends	Frequency	Percentage
Government support for AI adoption	18	41%
Automated incident response	19	43%
Human-AI collaboration emphasized	20	45%
Increasing cyberattack sophistication	31	70%
Infrastructure and skills limitations	16	36%
AI improves cyber-treats detection	34	77%
Machine Learning for anomaly detection	22	50%

Figure 8.

Distribution of Secondary Data Sources



Another observation from the trend is the increasing sophistication of cyber threats. Cyber incident reports reveal a growing number and prevalence of ransomware, malware, phishing, insider threats, and APTs, or advanced persistent threats, which reinforce the need for predictive and adaptive cybersecurity mechanisms. Government policies likewise show increasing institutional commitment to integrating AI into national cybersecurity strategies, while industry white papers emphasize automation, threat intelligence, and security operations center (SOC) modernization.

Despite these positive developments, the secondary data also consistently highlight challenges related to infrastructure readiness, workforce capability, data quality, governance, and ethical artificial intelligence implementation. These findings suggest that the successful adoption of AI in government cyber security depends not only on technological advancement but also on sustained investment in human expertise, institutional capacity, and policy support. So, the trend analysis supports the study's conclusion that AI has become an essential strategic component for building resilient, proactive, and sustainable cybersecurity systems in the Philippines, highlighting the government institutions.

The government faces a huge number of cyber threats daily. An example is phishing, where attackers can clone actual agencies to force employees to expose private information. Given that government offices heavily utilize email for communications, phishing is considered an evolving threat with possible implications related to data breaches and illegal access.

Table 5,

Common Cyber Threats Faced by the Philippine Government

Threat Types	Description	Frequency (2023-2024 Data)	Potential AI Mitigation Strategy
Phishing Attacks	Deceptive emails or messages to steal information	High	NLP-based email filtering and anomaly detection
Malware & Ransomware	Malicious software used for disruption or extortion	Moderate to High	Behavior-based detection using machine learning
DDoS Attacks	Overwhelming servers to cause downtime	Moderate	Network traffic analysis and early alert systems
Insider Threats	Threats originating from within the organization	Low to Moderate	User behavior analytics (UBA)
Zero-Day Exploits	Exploiting unknown vulnerabilities	Low	AI-driven vulnerability scanning and prediction

Ransomware also greatly affects similar agencies, where attackers install malicious scripts that freeze access to systems and require a certain ransom to enable owners to regain control.
Several government institutions have recently been victims of such attacks that caused a halt of operations, resulting in financial losses and damage to reputation.

Another threat comes via DDoS attack that floods online platforms' traffic, causing downtime to online systems and government websites, including online informational platforms used in delivering public services and elections. The study's content and data analysis explain potential AI mitigation strategies for the threat types mentioned above.

Figure 9.

Content Analysis of Cyber Threat Frequency and AI Mitigation Strategy Priority

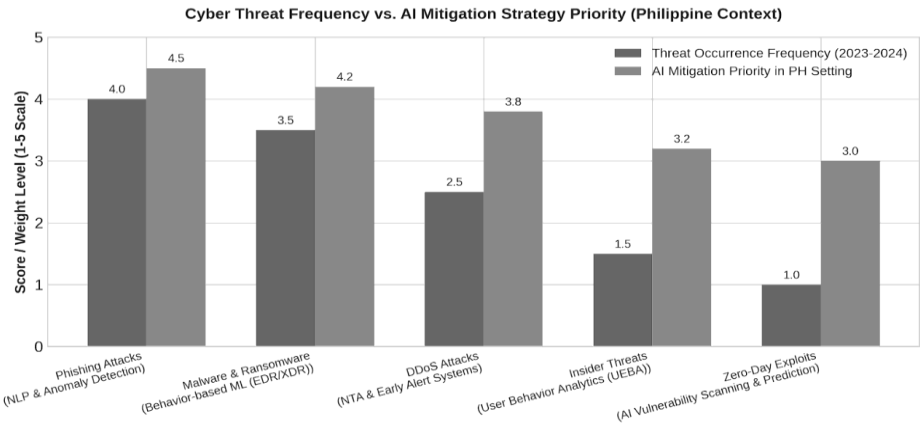


Figure 9 illustrates the comparison of the occurrence frequency of major cyber threats with their possible AI-based mitigation strategies in the Philippine context. The scores were derived from a qualitative-to-quantitative mapping based on the results of content analysis of global cybersecurity threat reports, covering the years 2023–2024, Philippine national cybersecurity frameworks, and the reviewed literature. The blue bars represent the relative frequency per threat, while the orange bars indicate the priority of implementing AI-based mitigation strategies in relation to feasibility, regulatory priorities, effectiveness, and relevance to the Philippine setting.

Phishing attacks got the highest threat occurrence score (4.0) and the highest AI mitigation priority (4.5). This score is an indicator of the use of phishing as the primary attack vector targeting individuals and organizations, particularly through email, SMS, and messaging applications targeting digital financial services. Natural Language Processing (NLP) and AI-based anomaly detection were highlighted with the highest-priority mitigation strategies based on the results of the content analysis, as they provide effective and low-cost solutions for internet use-related threats that are spread via malicious messages and fraudulent websites online.

On the other hand, malware and ransomware obtained threat frequency scores of 3.5 and an AI priority score of 4.2, respectively. Although less frequent than phishing, ransomware should always be

safeguarded against because of the more severe operational and financial consequences it can cause for affected individuals or organizations. Also, the data and literature emphasize behavior-based machine learning, including Endpoint Detection and Response (EDR) and Extended Detection and Response (XDR), as effective approaches for identifying abnormal system behavior, and the high priority aligns with one major direction of the Philippine National Cybersecurity Plan, putting emphasis on protecting critical information infrastructure.

DDoS attacks received a moderate occurrence score (2.5) and a mitigation priority score of 3.8. While these attacks occur less frequently, if they happen, they can significantly disrupt government services and critical infrastructure. Accordingly, Network Traffic Analysis (NTA) and AI-enabled early warning systems are considered strategic investments that any government agency or private organization should allocate for compliance with security and availability requirements.

Insider threats received a low occurrence score of 1.5 but a higher AI mitigation priority score of 3.2. This shows that organizations do not base their security priorities only on how often a threat occurs. Insider attacks can still cause serious damage, especially when sensitive information is involved, making them an important concern for organizations. User and Entity Behavior Analytics (UEBA) helps monitor user activities and identify unusual behavior that may indicate unauthorized access or misuse of data. It also supports compliance with the Data Privacy Act of 2012 and other cybersecurity regulations.

Zero-day exploits had the lowest occurrence score of 1.0 but had a moderate AI priority score of 3.0 based on the content analysis figures. Although such attacks are less common, they are considered very critical because they take advantage of security vulnerabilities that are still unknown to most technical experts. AI technologies can possibly counter this, in the form of tools like vulnerability scanners and those that integrate predictive analytics, which provide early alarm and identify possible weaknesses before they are exploited. However, many organizations in the country still face challenges in adopting these technologies due to limited budgets and the lack of qualified cybersecurity professionals.

Overall, Figure 9 shows that AI mitigation priorities should not be purely based on how on frequencies of cyber threats occur. For example, threats such as phishing and ransomware receive the highest priority

because they occur more often in organizations in the Philippines, and they provide significant disruptions and impact. Also, less frequent threats, including insider attacks and zero-day exploits, should also be closely monitored because of their potential consequences, including violations to cybersecurity regulatory requirements. Given these, prioritizing threat areas and investing in corresponding AI-based solutions should balance the likelihood and potential impact of the cyber threats.

To support the values presented in Figure 1, the qualitative threat levels were converted into a five-point scale through content analysis of global cybersecurity reports published from 2023 to 2024, together with relevant Philippine cybersecurity studies. The assigned scores represent the relative frequencies of each cyber threat based on evidence from both international and local sources included in the review literature of this study.

Table 6.

Threat vector score and supporting literature

Threat Vector	Score	Qualitative Scale	Supporting Literature
Phishing Attacks	**4.0**	High	IBM X-Force (2024); Fortinet (2024); DICT (2024); Caliwan (2024); Sigales (2024)
Malware & Ransomware	**3.5**	Moderate to High	Red Canary (2024); Fortinet (2024); NPC (2024); BlueVoyant (2025)
DDoS Attacks	**2.5**	Moderate	Mishra et al. (2022); CERT-PH (2022); Asia (2019); Sigales (2025)
Insider Threats	**1.5**	Low to Moderate	Kaur et al. (2023); Averia et al. (2022); BSP (2024)
Zero-Day Exploits	**1.0**	Low	Mahboubi et al. (2024); Josyula & Saidireddy (2024); Uy & Dy (2024); Jongvibool (2025)

The assigned scores go with the reviewed literature. Phishing attacks (4.0) received the highest frequency rating because they remain the leading attack vector globally (IBM X-Force, 2024; Fortinet, 2024) and in the Philippines, where phishing, smishing, and online scams account for a significant portion of reported cybercrime incidents (DICT, 2024; Caliwan, 2024; Sigales, 2024).

Malware and ransomware (3.5) were classified as moderate to high due to their continuous occurrence and severe impact. Global reports (Red Canary, 2024; Fortinet, 2024) and local incidents, including major public-

sector breaches and supply chain attacks (NPC, 2024; BlueVoyant, 2025), support this rating.

DDoS attacks (2.5) were rated moderate, reflecting their periodic occurrence, particularly against government systems and critical infrastructure during periods of heightened political or geopolitical activity (Mishra et al., 2022; CERT-PH, 2022; Sigales, 2025).

Insider threats (1.5) received a Low to Moderate score because they occur less frequently than external attacks but remain significant in regulated sectors such as banking and IT-BPO, where compliance requirements emphasize continuous monitoring (Kaur et al., 2023; Averia et al., 2022; BSP, 2024).

Finally, zero-day exploits (1.0) were assigned the lowest frequency score because they are relatively rare and typically associated with advanced persistent threat (APT) groups. Philippine studies indicate that most local cyber incidents continue to exploit known vulnerabilities, weak credentials, or phishing rather than previously unknown software flaws (Mahboubi et al., 2024; Josyula & Saidireddy, 2024; Uy & Dy, 2024; Jongvibool, 2025).

Overall, the literature supports the qualitative-to-quantitative mapping used in Figure 9, demonstrating that the assigned threat frequency scores accurately reflect current global and Philippine cybersecurity trends.

Table 7.

AI-Enhanced Cybersecurity Methods vs Traditional

Criteria	Traditional Methods	AI-Enhanced Methods
Detection Speed	Minutes to Hours	Real-time or near real-time
False Positive Rate	High	Lower due to contextual analysis
Adaptability	Rule-based, static	Self-learning, dynamic
Resource Requirements	High manual oversight	Reduced with automation
Threat Prediction	Limited	Proactive with predictive analytics

Adaptability is also said to be one of the focal areas given the continuously evolving threat environment where government institutions must operate. In Table 2, traditional cybersecurity frameworks are based on rule-based systems under predefined signatures, static protocols, and

manual threat databases. While these systems are the general foundations, their limitations have been exploited in the rapidly evolving cyber threat landscape, as they are inherently reactive. Threats are detected based on already known patterns, and this approach requires depending on constant human intervention to update rules and signatures. In the volatile environment, threats like zero-day may exploit these systems as they normally fail in terms of timely response, where a threat may not be updated yet in the database. Additionally, attackers have become more sophisticated in evading static detection through polymorphic malware, encrypted traffic, and novel attacks vectors. In such scenarios, static systems become obsolete almost as soon as they are updated.

In contrast, self-learning dynamic systems, often powered by artificial intelligence and machine learning, represent a transformative leap in adaptability. These systems do not rely solely on predefined rules. Instead, they continuously analyze data patterns, learn from new inputs, and adjust their detection algorithms based on contextual behavior. For instance, if a user's account suddenly attempts to access classified files at odd hours from an unfamiliar device, an AI-powered system can flag the incident as suspicious (Efe, 2024; Lee, 2024; Hammi, Zeadally, & Nebhen, 2023; Lewke, 2023) even if the activity doesn't match any known threat pattern. Over time, the system becomes more refined in distinguishing between legitimate anomalies and real threats, reducing false positives and enhancing response times.

The dynamic nature of AI also allows for real-time decision-making. Unlike static systems that require scheduled updates, self-learning models update themselves through feedback loops and ongoing training on fresh datasets. In government cybersecurity operations, this capability is particularly valuable. The ability to autonomously adapt to evolving threats means fewer breaches, quicker containment, and a more proactive stance in national cyber defense.

It is also important to note that dynamic systems may also face issues given heavy reliance on datasets, strict AI governance, and mechanisms to counter adversarial changes, but if properly implemented, the advantage outweighs the costs, especially compared to the lack of flexibility of traditional approaches. As a whole, static rule-based systems served as the foundations for cybersecurity frameworks. However, the complexities of evolving threats shifted focus to self-learning dynamic

models. Among Philippine government institutions, investing in adaptable, AI-driven cybersecurity solutions may be considered critical nowadays.

Insider threats also provide potential risk to Philippine government institutions, and employees with access may also cause leaks of information or expose vulnerabilities because of negligence. Also, the risk of insider threats has been increased with the introduction of remote work, where an ideal security environment may not be present. Cyber espionage is also another threat, where the problem can come from external states with the objective of obtaining access to critical political, economic, or military data. Government agencies handling defense, foreign policy, and military-related intelligence are at greater risk against advanced persistent threats, where attacks get undetected for a certain period, heavily exposing critical data.

The government also consistently faces threats in the form of malware, like worms, trojans, and spyware. These can penetrate government networks where backdoors can be created, allowing attackers to maintain access over time. Government agencies maintaining public health data, critical government infrastructure, and law enforcement have been usual targets for malware, and that data breach continues to be a significant issue in the Philippines. Personal information, government data, and private communications are regularly affected, especially when the right security protocols and foundations are not deployed. Such problems often result in the exposure of data leading to identity theft and fraud.

The table below specifies the emerging trends identified in the secondary data. In this study, it has been observed that frequencies represent the number of reviewed documents in which each theme was identified and appeared.

Conclusion

In the ever-evolving landscape of cyber threats, adaptability has become a defining requirement for effective cybersecurity. Static rule-based systems are considered inefficient due to their dependence on manual signatures and dependency on human control. The reactive process makes them insufficient against sophisticated threats that target government systems. Artificial intelligence provides a paradigm shift where the model continuously is being updated through real-time data learning, pattern identification, and autonomous integration to new threat vectors. Detection

of unknown and emerging threats with minimal human dependency enhances the efficiency and resilience of cybersecurity.

For Philippine government institutions, working towards this shift is a strategic move because as digital services are more deployed, the intelligence and adaptability of the platforms must also be enhanced. The shift from static to dynamic cybersecurity models serves as a critical step toward establishing a more secure, proactive, and technologically adaptive government sector.

This study is limited by its reliance on secondary data obtained from published cybersecurity threat intelligence reports, government publications, industry reports, and peer-reviewed literature. While these sources provide credible and up-to-date information, they may contain reporting biases, differences in data collection methodologies, and varying definitions of cybersecurity incidents across organizations and countries. Furthermore, the study did not conduct primary data collection, empirical experiments, surveys, or interviews with cybersecurity practitioners. Consequently, the proposed AI mitigation priorities rely on qualitative content analysis instead of direct implementation or performance evaluation in real-world environments. The findings should therefore be interpreted as a conceptual framework for guiding AI adoption in cybersecurity rather than as empirical evidence of solution effectiveness. Future validation through experimental studies and organizational case studies would further strengthen the proposed framework.

Recommendations

With the evolving landscape of cyber threats and the complexity of attacks targeting government institutions, the following recommendations are geared towards the adaptability and resilience of cybersecurity platforms. Government agencies should start integrating artificial intelligence and machine learning tools that support dynamic models. The government should fast-track the introduction of AI-driven security solutions given that these systems can autonomously detect, analyze, and respond to emerging threats in real-time with reduced dependency on static rule-based methods.

Many government systems still operate under outdated, static, and rule-based frameworks. Therefore, a phased approach to modernization should be implemented to transition these systems toward adaptive

frameworks capable of evolving with emerging threats, and the government should also modernize legacy infrastructure with dynamic security frameworks. Given the results presented in this paper, government institutions should also start creating a specialized national AI cybersecurity task force, which will be managed by the Department of Information and Communications Technology (DICT), which shall oversee monitoring, research, and ethical governance of AI-based cybersecurity solutions among government agencies.

Given this new paradigm and with the introduction of AI systems in cybersecurity, capacity-building initiatives toward AI integrated with cybersecurity should be fortified, as more human specialists will be needed to work alongside maintaining these tools and that these should be supported by the government, which should invest in continuous learning programs focused on AI education, threat monitoring, and dynamic defense techniques.

Additionally, implementing risk management and regular audits with periodic assessments to evaluate the adaptability of current cybersecurity systems is also needed, where consistent evaluation of cybersecurity adaptability measures should factor in mitigation response times, accuracy of detection, system learning capability, and adaptability to novel threats.

Partnerships with private firms specializing in AI and cybersecurity, academic institutions, and research-based organizations should also be established to set parameters for innovative solutions that promote cybersecurity dynamism and resilience. A public-private collaboration for an innovation initiative embracing these recommendations can move beyond the limitations of traditional cybersecurity methodologies and develop a more agile and proactive cybersecurity framework for government institutions that shall continuously evolve to address modern cyber threats and issues.

The findings of this study provide several opportunities for future research. First, the proposed AI-based mitigation framework may be validated through empirical case studies involving government agencies, financial institutions, healthcare organizations, and other critical infrastructure sectors in the Philippines. Second, future studies may evaluate the effectiveness of specific artificial intelligence techniques, such as machine learning, deep learning, natural language processing, and user

behavior analytics, using real cybersecurity datasets and measurable performance metrics, including detection accuracy, false positive rate, and response time. Third, comparative analyses may investigate the applicability of the framework across different organizational sizes, industry sectors, and levels of cybersecurity maturity. Finally, future research may explore emerging AI technologies, including generative AI, large language models (LLMs), autonomous security operations, and explainable artificial intelligence (XAI), to determine their potential in enhancing cyber threat detection, incident response, and cybersecurity decision-making while addressing ethical, privacy, and regulatory considerations.

References

Ali, O., Murray, P. A., Momin, M., Dwivedi, Y. K., & Malik, T. (2024). The effects of artificial intelligence applications in educational settings: Challenges and strategies. *Technological Forecasting and Social Change, 199*, 1–18.

Asia, O. (2019, January 17). DICT launches the Philippines' national cyber intelligence platform. *OpenGov Asia*. https://opengovasia.com/2019/01/17/dict-launches-philippines-national-cyber-intelligence-platform/

Asmar, M., & Tuqan, A. (2024, September 15). Integrating machine learning for sustaining cybersecurity in digital banks. *Heliyon, 10*(17), 1–18.

Averia, A. S., Pascual, G., Yu, W., Mirandilla-Santos, M., Gutierrez, A., Claudio, B., . . . Tajale, F. (2022). *Cybersecurity in the Philippines: Global context and local challenges*. The Asia Foundation–Philippines & Secure Connections.

Benson, M. (2022). *Towards a research guide for cyber threat intelligence* [Master's thesis, Utica University].

Binhammad, M., Alqaydi, S., Othman, A., & Abuljadayel, L. (2024). The role of AI in cybersecurity: Safeguarding digital identity. *Journal of Information Security, 15*, 245–278.

BlueVoyant. (2025, January 27). *The state of supply chain defense: Annual global insights report*. https://www.bluevoyant.com/press-releases/philippines-organisations-negatively-impacted-by-cyber-security-breaches

Bolivar, J. (2024, July 23). Philippine SONA 2024: Ferdinand Marcos Jr. says country embraces AI, digitalization. *CoinGeek*. https://coingeek.com/philippine-sona-2024-ferdinand-marcos-jr-says-country-embraces-ai-digitalization/

Buczak, A., & Guven, E. (2016). A survey of data mining and machine learning methods for cybersecurity intrusion detection. *IEEE Communications Surveys & Tutorials, 18*, 1153–1176.

Cadiang, G. (2024, August 21). 2024–2025 government budget: Strategic investment in cybersecurity aligned with the Philippine Skills Framework. *Lumify Work*. https://www.lumifywork.com/en-ph/blog/2024-2025-government-budget-strategic-investment-in-cyber-security-aligned/

Caliwan, C. L. (2024, December). New PNP-ACG chief vows to build 'cutting-edge cyber defenses'. *Philippine News Agency*. https://www.pna.gov.ph/articles/1240517

Dasandi, N., & Mikhaylov, S. (2019, July). *AI for SDG-16 on peace, justice, and strong institutions: Tracking progress and assessing impact*. Knowledge 4 All Foundation. https://www.k4all.org/wp-content/uploads/2019/07/AI-for-SDG-16-on-Peace-Justice-and-Strong-Institutions-Tracking-Progress-and-Assessing-Impact.pdf

Department of Information and Communications Technology Cybersecurity Bureau. (2024, February). *National cybersecurity plan 2023–2028*. e-Gov PH. https://cms-cdn.e.gov.ph/DICT/pdf/NCSP-2023-2028-FINAL-DICT.pdf

Edralin, D., & Pastrana, R. (2020). The nexus between sustainable business practices and the quest for peace. *Bedan Research Journal, 5*(1), 1–58.

Edralin, D., & Pastrana, R. (2022). Advancing the pursuit of the United Nations Sustainable Development Goals: Initiatives of selected publicly listed companies in the Philippines. *Bedan Research Journal, 7*(1), 1–47.

Efe, A. (2024). Risk modelling of cyber threats against MIS and ERP applications. *Pamukkale Üniversitesi İşletme Araştırmaları Dergisi, 11*(2), 502–530.

Flores, H. M. (2024, July 22). President Marcos pushes for digitalization in first 2 SONAs. *Philstar.com*. https://www.philstar.com/headlines/2024/07/22/2371895/president-marcos-pushes-digitalization-first-2-sonas

Fortinet. (2024). *2H 2023 global threat landscape report*.

Georgescu, T. M. (2021). A study on how the pandemic changed the cybersecurity landscape. *Informatica Economică, 25*(1).

Hammi, B., Zeadally, S., & Nebhen, J. (2023). Security threats, countermeasures, and challenges of digital supply chains. *ACM Computing Surveys, 55*(14), 1–40.

Herwig, C. (2015). *Peace, justice, and strong institutions: Why they matter*. United Nations. https://www.un.org/sustainabledevelopment/wp-content/uploads/2019/07/16_Why-It-Matters-2020.pdf

Hong, Q. N., & Pluye, P. (2019). A conceptual framework for critical appraisal in systematic mixed studies reviews. *Journal of Mixed Methods Research, 13*(4), 446–460.

IBM. (2024). *X-Force threat intelligence index 2024*.

International Trade Administration. (2024, September 20). *Philippines—Country commercial guide: Digital economy*. U.S. Department of Commerce. https://www.trade.gov/country-commercial-guides/philippines-digital-economy

Jiang, T., Sun, Z., Fu, S., & Lv, Y. (2024). Human-AI interaction research agenda: A user-centered perspective. *Data and Information Management*, 1–10.

Jongvibool, P. (2025, March 13). AI-driven cybersecurity: A game changer for organizations in the Philippines. *BusinessWorld*. https://www.bworldonline.com/technology/2025/03/13/658915/ai-driven-cybersecurity-a-game-changer-for-organizations-in-the-philippines

Josyula, M., & Saidireddy, M. (2024). Mitigating adversarial and AI-evasion attacks in cybersecurity: Challenges and solutions. *Journal of Propulsion Technology*, 3152–3162.

Kaur, R., Gabrijelcic, D., & Klobucar, T. (2023). Artificial intelligence for cybersecurity: Literature review and future research directions. *Information Fusion, 97*, 1–29.

Kinyua, J., & Awuah, L. (2021). AI/ML in security orchestration, automation and response: Future research directions. *Intelligent Automation & Soft Computing*, 527–545.

Lee, N. (2024). Cyberattacks, prevention, and countermeasures. In *Counterterrorism and cybersecurity: Total information awareness* (pp. 295–342). Springer International Publishing.

Lewke, D. (2023). *The MIT-IBM CloudSec 16: A cloud cybersecurity benchmarking framework.*

Luci-Atienza, C. (2021, April 8). DOST launches 9 new AI R&D projects. *Manila Bulletin*. https://mb.com.ph/2021/4/8/dost-launches-9-new-ai-rd-projects

Luckin, R., Cukurova, M., Kent, C., & du Boulay, B. (2022). Empowering educators to be AI-ready. *Computers and Education: Artificial Intelligence, 3*, 1–11.

Luna, J. (2024, August 28). The Philippines' National Cyber Security Plan 2023–2028: Roadmap to cyberspace resilience. *Lumify Work*. https://www.lumifywork.com/en-ph/blog/the-philippines-national-cyber-security-plan-2023-2028-roadmap-to-cyberspace/

Mahboubi, A., Luong, K., Aboutorab, H., Bui, H., Jarrad, G., Bahutair, M., . . . Gately, H. (2024). Evolving techniques in cyber threat hunting: A systematic review. *Journal of Network and Computer Applications, 232*, 1–34.

Mishra, S., Albarakati, A., & Sharma, S. (2022). Cyber threat intelligence for IoT using machine learning. *Processes*, 1–22.

Mondejar, M. E., Avtar, R., Diaz, H. B., Dubey, R., Esteban, J., Gómez-Morales, A., . . . Garcia-Segura, S. (2021). Digitalization to achieve sustainable development goals: Steps towards a smart green planet. *Science of the Total Environment*, 1–20.

National Computer Emergency Response Team. (2022). *CERT-PH annual report 2022*.

National Privacy Commission. (2024, April 16). *NPC PhilHealth portal*. https://philhealthleak.privacy.gov.ph/

NCC Group. (2025). *Cyber threat monitor report 2024*.

Newman, L., & Burgess, M. (2024, November 21). Meta finally breaks its silence on pig butchering. *WIRED*. https://www.wired.com/story/meta-pig-butchering-report-2024/

Palo Alto Networks. (2025). *Threat monitor report*.

Red Canary. (2024). *2024 threat detection report*.

Saeed, S., Altamimi, S. A., Alkayyal, N. A., Alshehri, E., & Alabbad, D. A. (2023). Digital transformation and cybersecurity challenges for business resilience: Issues and recommendations. *Sensors, 23*(6666), 1–20.

Sauerwein, C., & Staiger, T. (2021). *Cyber threat intelligence sharing platforms: A comprehensive analysis of software vendors and research perspectives*.

Sharma, N., Sharma, R., & Jindal, N. (2021). Machine learning and deep learning applications—A vision. *Global Transitions Proceedings, 2*, 24–28.

Sigales, J. (2024, December 19). PNP partners with TikTok to boost drive vs. online scams, exploitation. *Inquirer.net*. https://newsinfo.inquirer.net/2016964/fwd-pnp-partners-with-tiktok-to-boost-fight-vs-online-scams-exploitation

Sigales, J. (2025, January 7). AFP assures ‘systems in place’ amid report of Chinese hacking on PH. *Inquirer.net*.