

American Journal of Economics and Business Management

Vol. 9 Issue 11 | **pp.** 479-490| **ISSN:** 2576-5973

Available online @ https://www.globalresearchnetwork.us/index.php/ajebm

Article

AI-Driven Cybersecurity Risk Management for AI-Powered Business Systems

Fahmida Akter[1], Mahbub Hasan[2], Mohd Salman Hossain Bhuiyan[3], Muhammad Farhan Nafees[4]

1. Master of Science in Information Technology, Washington university of science and technology, USA
2. Master of Science in Information Studies, Trine University, USA
3. Master of Science in Computer Science, San Francisco Bay University, USA
4. Masters in AI for Business, National college of Ireland, Ireland

* Correspondence: a.fahmida42@gmail.com[1], mahbub0016@gmail.com[2], mohd.salmanhossain@gmail.com[3], farannafees9@gmail.com[4]

Abstract: An integral part of current business systems, Artificial Intelligence is opening up new opportunities for automation, decision support, and intelligent service delivery but also offers new cybersecurity concerns. This research aims to explore the nature of cyber threats targeting AI and LLM environments by analyzing a secondary AI Sec Watch Dataset. The research explores threat severity, threat exploitability, sophistication of attack, targeted AI functions, and threat prioritization to determine meaningful cybersecurity threat patterns. A qualitative research design is used to analyze 3020 threat records and the related security attributes, backed up with quantitative exploratory data analysis. Key threat variables are explored in terms of their distribution and relationships by descriptive statistics and visual analysis. This analysis has shown that medium and high-severity threats are the most common recorded events, which shows that there is a high level of exposure in AI-powered environments. The analysis also illustrates that vulnerabilities of the same severity can have different probabilities of being exploited, thus depending on the number of risk factors assessed. AI risks are not evenly distributed among AI components like large language models, agents, APIs or vector databases, as there are differences in sophistication of attack and concentration of threat. Furthermore, there is a strong correlation between threat prioritization patterns and threat severity, exploitability and impact. The study helps to enhance the understanding of AI cybersecurity threats and offers practical implications that may assist in the risk assessment, security planning, and prioritization process for AI-powered business systems.

Keywords: Artificial Intelligence, Business System, AI, Automation, Cyber Threats, Dataset, Exploratory.

Citation: Akter, F., Hasan, M., Bhuiyan, M. S. H., & Nafees, M. F. AI-Driven Cybersecurity Risk Management for AI-Powered Business Systems. American Journal of Economics and Business Management 2026, 9(11), 479-490.

Received: 09th Aug 2026
Revised: 04th Sep 2026
Accepted: 08th Oct 2026
Published: 11th Nov 2026

Copyright: © 2026 by the authors. Submitted for open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/)

1. Introduction

1.1 Background of Study

Artificial Intelligence (AI) is a key component of modern business systems, supporting intelligent digital operations, customer service, data analysis, and automated decision-making. The rising adoption of generative AI and large language models (LLMs) by organizations has made securing them against targeted cyber threats increasingly critical. Prompt injection, disclosure of sensitive information, vulnerabilities in the supply chain, and data and model poisoning are the most significant risks for LLM applications, as described by the OWASP 2025 guidance [1]. These risks can affect the information's

confidentiality, systems' integrity, operational reliability, and the reliability of AI-generated outputs. Also, AI applications can tap into external tools, databases, plugins, and application programming interfaces, which can create new attack vectors to exploit weaknesses in the tools and databases connected to the AI application. As a result, traditional security measures might not be sufficient to protect against unique vulnerabilities of AI-powered technologies. To effectively identify the threats that pose the greatest risk to businesses, it is important to understand the severity of a threat, how easy it is to exploit, the sophistication of the attack, and the AI components that are vulnerable to attacks. By understanding these traits, companies can identify potentially severe threats from less important ones and make better decisions about cybersecurity. Hence, it is more important than ever to consider AI security intelligence and LLM security intelligence systematically to enhance risk awareness and inform the Cyber Security Risk Management of AI-based business systems [1].

1.2 Problem Statement

As AI becomes increasingly embedded in business systems, it presents new cybersecurity challenges that traditional information security strategies and protocols might not be able to solve. The limitations of existing risk assessment techniques in capturing AI-specific risks, such as adversarial attacks, evasion attacks, and data poisoning, were identified in a 2026 peer-reviewed study [2]. These constraints can pose a challenge for organizations when assessing vulnerabilities in AI models, data pipelines, the process of inference, and the integrated application. While risks need to be identified and prioritized, various levels of risk severity, exploitability, sophistication of an attack, and component exposure can result in varying risk profiles. The 2026 study emphasizes the need to refine and enhance current risk assessment frameworks to account for the unique threat categories and appropriate mitigation strategies associated with AI [2]. Yet, grasping these challenges also demands looking at the available threat intelligence data to see if there are patterns in the reported threats and the impacted AI component. For this reason, this research investigates the threat patterns and the threat prioritization by looking at the severity of the cyber threat, the possibility of the exploit, the level of sophistication of the attack, and the targeted AI components. This study is not a development or validation of the predictive model, but secondary data and exploratory analysis.

1.3 Aim of Research

The goal of this research is to examine the level of severity of a cybersecurity incident, the exploitability, how sophisticated the attack is, and the targeted AI components to figure out the threat patterns and to inform risk prioritization in an AI-powered business system.

1.4 Research Objectives

- To examine the severity, exploitability, and sophistication of AI and LLM security threats.
- To study the correlation between the influences of affected AI elements, cybersecurity effects, and threat prioritization in AI-driven systems.

1.5 Significance of Study

This study is an effort to understand the patterns of AI and LLM security threat intelligence, which helps in understanding the cybersecurity risks associated with AI-powered business systems. Take severity, exploitability, attack sophistication, targeted AI components, and reported cybersecurity impacts into account to give a structured overview of the threat landscape. The results can aid cybersecurity experts in identifying common risk profiles and threat categories to investigate. The study also provides a basis for informed risk prioritization by investigating the relationship between threat characteristics, affected components, and priorities assigned. Academically, it serves as a preliminary study for future research on AI security risk assessment and management.

While the analysis does not prove causation or the accuracy of a predictive model, it can help inform future research and further systematic analysis of cybersecurity threats involving AI.

1.6 Literature Review

1.6.1 Artificial Intelligence in Business Systems and Emerging Cybersecurity Risks

The use of AI in business systems is a sign of the change in the way businesses are making decisions, managing knowledge, and providing services. Gupta et al. [3] establish the theoretical underpinnings of understanding generative AI adoption and implications and suggest directions for future research, including a focus on the impact of generative AI on organizational outcomes. It's crucial as it shifts the way data is accessed, content created, and decisions made in a business, and brings new reliance on model outputs and the quality of the data. Reznikov [4], on the other hand, investigates the patterns of strategic adoption in enterprises, pointing to organisational considerations that can affect the incorporation of generative AI into current operations. Such views are supportive but not conflicting, as there are subtle differences in the emphasis. While Gupta et al. [3] offer more comprehensive theoretical arguments for study of adoption and impact, Reznikov [4] is more closely concerned with enterprise adoption strategy. Moreover, the adoption frameworks do not provide information on whether an organisation has sufficient controls to address AI-related cyber threats. The difference between successful and secure implementation then needs to be explored in greater detail when it comes to the technical architecture and risk exposure of the AI-powered business system.

Jeong [5] moves beyond organisational adoption to the technical architecture needed to make generative AI work in enterprise settings. The study explores an enterprise data-based LLM architecture, along with fine-tuning and direct document integration as solutions to overcome data constraints, and retrieval-augmented generation (RAG) as a method for supporting information retrieval and content generation. In the field of cybersecurity, this architectural viewpoint is especially important since enterprise AI applications could integrate language models with internal documents, databases, and other information resources. These are security issues to consider with access rights, exposure of sensitive information, data integrity, and the reliability of retrieved content. While Gupta et al. [3] and Reznikov [4] present their viewpoints more adoption-centric, Jeong [5] gives an implementation-centric perspective on the structuring of enterprise AI services. The study focuses, however, largely on architecture and information retrieval and does not comprehensively evaluate the risk and vulnerability of cybersecurity threats. As such, its technical contribution contributes to the understanding of potential security issues, but does not necessarily provide a framework for prioritizing or addressing them across various AI elements.

The discussion on adoption is continued by Gupta and Yang [6], who propose a Generative AI technology adoption model for an entrepreneur that captures technological familiarity, system quality, training, support, interaction convenience, and perceived value at the different stages of adoption. Their model extends the enterprise strategy analysis, as it takes into account the users' perceptions and the ability to evaluate and integrate new technologies. Significant to cybersecurity, this can mean that employees and entrepreneurs could utilize AI tools based on their perceived utility without having a robust understanding of potential security concerns. However, it shouldn't be interpreted as an indication that these users actually neglect cybersecurity controls; to be certain, that would necessitate a direct investigation. While Jeong [5] focuses on the architectural view of the problem, Gupta and Yang [6] emphasized the human and organizational factors influencing adoption, and Jeong focuses on the technical configuration of enterprise AI services. All of these perspectives highlight the importance of technical design, as well as organisational practices, in AI security management, but not in isolation.

The common problem in these studies is that they focus on the adoption of AI, integrating AI into the enterprise, and how AI will be of value to the organisation, but not

on the systematic categorization of cyber threats unique to AI. Although Gupta et al. [3] offer approaches to the theoretical analysis of adoption and its implications, and Reznikov [4] describes patterns of strategic enterprise adoption, neither source, as presented here, explicitly quantifies which elements of AI present greater or lesser risk. Jeong [5] provides architectural guidance for enterprise LLM applications, while Gupta and Yang [6] provide some factors to consider for an adoption decision. While these contributions provide complementary perspectives, there remains an important analytical question: how to study the severity of cybersecurity, the exploitability, the sophistication of the attack and the targeted AI components together. Additionally, while organisational advantages like productivity and better information access may imply secure AI deployments, they do not necessarily do so. However, the analysis of structured AI and LLM threat intelligence can help augment this literature, as it can explore the characteristics of reported threats and their prioritization, whilst acknowledging that a causal link between AI adoption decisions and security outcomes cannot be definitively proved by examining the data for these threats.

1.6.2 AI and LLM Security Threats

As large language models (LLMs) are increasingly embedded in intelligent applications, the cybersecurity threat surface has grown to encompass more than software vulnerabilities. Gan et al. [7] discuss security, privacy, and ethical risks of LLM-based agents, focusing on threats from agent autonomy, external tools, and access to sensitive data. Their wider view is relevant as agentic systems may carry out connected functions as opposed to producing text, meaning that any initial weakness may impact other functions and information flows. But a wide threat survey provides a different analysis than research focused on particular attack methods. Singh et al. [8] explore the potential of fooling LLMs with deceptive methods and persuading them using cognitive tricks, underscoring the importance of how interactions can shape the output of LLMs. This view further adds to the agent security debate and addresses the human and linguistic nature of exploitation. However, the susceptibility to misleading prompts and broader dangers of free agents should not be considered the same issue. The former is about the behaviour of the model being manipulated; the latter also relates to the impact of the permissions and architecture of the model and the interactions between the model components. Understanding these threat mechanisms is vital to evaluating exposure to cyber threats in AI-enabled business systems.

Another facet of LLM security is the ability for prompt obfuscation; system prompts may include proprietary instructions that impact the behavior of an application. Pape et al. [9] explore prompt obfuscation to minimize the threat posed by the extraction of system prompts while maintaining the functionality of the prompt. They measure the similarity between the original and obfuscated prompts with several similarity measures, and discuss cyber-attacks under various attacker knowledge settings. While similar to Singh et al. [8] in the focus on protecting system instructions, this work tackles the protection of system instructions rather than primarily the manipulation of the model response by deceptive interactions. The distinction is relevant when deployed in an enterprise environment where a timelier disclosure could provide information about application logic, enterprise procedures, or information about safeguards. Immediate obfuscation should not be seen as an all-in-one cybersecurity solution; however, as system instructions are protected, that does not necessarily mean that access to data can't be used, that malicious tools can't be employed, or that vulnerabilities within the application surrounding the system can't exist. The role it plays in cybersecurity, therefore, is relatively specific, as it provides a protective measure for a single attack surface and not a complete solution to assessing the risks of an AI system. This is particularly helpful in distinguishing between prompt confidentiality and operational security when combined with agent-based threat analysis.

In addition to prompt manipulation, vulnerabilities can also stem from the software frameworks used for building, transforming, and distributing AI models. Deng et al. [10] propose TENSORSCOPE, a differential testing technique for locating inconsistencies between the same application programming interface (API) among deep learning frameworks. Their assessment for 1,658 APIs distributed throughout 6 frameworks identifies 257 bugs of which 230 are new, showing that implementation inconsistencies can pose security and reliability issues that go beyond the underlying model. The discovery expands the scope of the debate on LLM-specific attacks beyond that of the LLM to uncover that the underlying machine learning architecture can pose vulnerabilities even if the intended model operations are not altered. Deng et al. [10] look at system implementation vulnerabilities and framework interoperability, as opposed to Pape et al. [9], who concentrate on protecting system prompts. Gan et al. [7] offer a broader classification of agent threats and Singh et al. [8] show the threats that arise from deceptive interactions. However, the different threats surfaced and the different analytical methods applied mean that it is difficult to make direct comparisons as to the relative severity of threats. For cybersecurity risk management, this generates a requirement for identifying the nature and severity of cyber threats, to clearly understand how cyber threats can be exploited, the sophistication of the attacks, and the AI component that is impacted. These differences are especially important when using structured threat intelligence and evaluating the relation of threat priorities to a security characteristic.

1.6.3 Emerging Approaches to Cybersecurity Risk Assessment in AI-Powered Systems

AI is increasingly used in the cybersecurity risk assessment process to aid in continuous monitoring, threat detection, and risk-informed decision-making. Chowdhury [11] reviews the deep learning methods used for real-time cybersecurity risk assessment in enterprise IT systems, including CNNs, RNNs, transformers, and GNNs. The review points to their ability to analyses and analyse security telemetry, as well as their ability to support continuous risk scoring. There are, however, significant caveats in its results, such as the use of simulated data, variability in evaluation methods, and the lack of long-term production deployments. These restrictions can be important because the detection accuracy in a controlled experiment does not guarantee the assessment of risk in a complex operational environment. Chirra [12] on the other hand, is more concerned with predictive AI for cyber risk assessment, which aims to identify threats before they happen, instead of focusing solely on responding to detected threats. Earlier threat recognition can be achieved with both methods; predictive capability does not mean full risk assessment. Interpretation of the potential implications of threats, systems affected, and relative urgency for mitigation is also a key component of effective assessment.

Predictive risk assessment is more useful if it is linked to structured prioritization and continuous security management. Chirra [12] expounds the role of AI in threat assessment and intelligence gathering, and how predictive analytics can enhance proactive counterattacks. Unlike Chowdhury [11], who reviews more on deep learning architectures and their deployment in real-time applications. Although deep learning can be used to better recognise complex patterns, a risk score generated by deep learning should also be able to be understood and acted on by the security professional. Udeze et al. [13] explore automated cybersecurity risk assessment using machine learning, threat intelligence, anomaly detection, and decision support. They propose a framework that includes data preprocessing, feature engineering, risk scoring, and ongoing monitoring, and acknowledge the challenges of data quality, model bias, privacy, scalability, and adversarial attacks. This is useful, as it is a more general way of relating analytical skills to organisational security decisions. However, the automation of prioritization does not ensure that it is done in an appropriate manner, especially in the absence of threat data or when it is not possible to explain threat scores to decision makers.

Another view is offered by Raza [14], who focuses on proactive cyber defense using AI for vulnerability detection, attack prediction, anomaly detection, and response. The

method focuses on the operational benefits of identifying emerging threats early in their development and avoiding them becoming serious incidents. Raza [14] offers a wider description of proactive defence and the role that machine learning could play in building organisational resilience than Udeze et al. [13] do. The information available about Raza's publications, however, fails to reveal a journal where he published, and his formal publication status and review process are not readily ascertainable. It should therefore be viewed with some skepticism and not be accepted as the same level of evidence as a well-documented peer-reviewed study. More generally, predictive detection and automated assessment are two different, but related, functions. A system can identify an unusual activity but not the business implications, and a risk assessment process must take into account threat likelihood, potential impact, and priorities for mitigating the threat. It is especially relevant for AI systems as the vulnerability can lie in the AI model, application interface, plugins, data sources, or agents that are connected to the system.

One key challenge in all these methods is the difficulty of converting general cybersecurity indicators into more specific risk insights for AI and LLM applications. Chowdhury [11] points to gaps in operational evaluation, and Udeze et al. [13] highlight governance and implementation issues that could impact automated assessment. The importance of proactive security is emphasized by Chirra [12] and Raza [14], but predictive detection does not necessarily mean that the most exposed AI components are the ones that have been prioritized, nor do prioritized components always align with ease of exploitability and severity of impact. These disparities indicate that a more complex approach to threat intelligence needs to be taken, and that a single model output or aggregate risk score should not be used. Exploratory analyses of the severity, CVSS scores, EPSS scores, sophistication of the attack, affected AI components, and recorded threat priorities can be used to identify patterns in available data without asserting causalities or the validity of a predictive model.

1.6.4 Literature Gap

Previous studies have focused on AI adoption, potential threats to security with the use of LLMs, predictive analytical tools, and automated cybersecurity risk assessment. However, a lack of integration of threat severity, exploitability, the sophistication of the attack, targeted AI components, and threat prioritization restricts the understanding of AI security risks. Structured threat intelligence analysis can be explored without the need for predictive models to fill this void.

2. Materials and Methods

2.1 Research Design

This research is exploratory, qualitative in nature, and utilizes a quantitative approach to examine cybersecurity threat patterns in AI-based systems through the analysis of secondary threat intelligence. The research targets the cybersecurity severity, exploitability, sophistication of attacks, targeted AI components, and threat prioritization distribution. Descriptive statistics and visual analysis are used to see if there are patterns, differences, or relationships between the variables selected. This study does not require the development, training, or validation of a new machine learning model. Instead, it uses existing structured data to understand the characteristics of AI and large language model security threats. This design aligns with the research goals and offers an evidence-based picture of the cybersecurity threats and their implications for risk assessment in AI-powered business systems.

2.2 Data Source and Description

The study is based on the AI Sec Watch Dataset (AI/LLM Security Threat Intelligence) published at Zenodo in 2026. The dataset includes 3,020 entries and 44 attributes about AI- and LLM-related cybersecurity threats [15]. The factors include severity, cvss_score, cvss_severity, epss_score, exploit_maturity, attack_sophistication, attack_type, impact_type, ai_component_targeted, llm_specific, and priority [15]. The

threat severity, threat exploitation potential, attack characteristics, affected AI components, reported impacts, and threat prioritization information are provided by these variables. The data does not have to be collected but can be used from existing sources, which is called secondary data. It does not directly measure financial losses at the organisation, operational downtime, or business resilience, but it is a structured format that can provide quantitative analysis of cybersecurity risk patterns.

2.3 Data Preprocessing

Data preprocessing consists of examining the structure of the data, checking for variable names, data types, missing data, and duplicate or inconsistent data before analysis. Selected variables include numerical measures, such as CVSS and EPSS scores, which are checked for proper handling in quantitative analysis, while categorical measures are used as categories, such as severity, sophistication of the attack, and targeted components of AI. The zeros in the dataset are examined to see if they could affect the descriptive data and plots. Records are also duplicated (replicated), and inconsistent category labels are evaluated to minimize the risk of misleading results. Category labels may be standardized (without altering the meaning of the observations) where necessary. The preprocessing stage keeps track of relevant records and documents any transformations and exclusions made for analytical transparency and reproducibility.

2.4 Exploratory Data Analysis

Exploratory data analysis is used to analyze the distribution and relationships of the selected variables of cybersecurity. Descriptive summaries are used to gain an understanding of the frequency of severity categories, the distribution of CVSS and EPSS scores, and the level of sophistication of the attacks. Comparative analysis is a comparison between the severity or attack characteristics of different targeted AI components. The linkages among the different indicators of exploitability, threat impacts, and priority are also examined for the purpose of uncovering potentially significant patterns of risk. Categorical comparisons and numerical associations are understood based on the nature of the measures. The analysis is centered on the identification of concentrations, differences, unusual observations, and relationships that could aid in cybersecurity risk assessment. The associations observed in the study do not suggest causation or future attack behaviour and are explored in an exploratory study.

2.5 Visualization Strategy

The central patterns of the cybersecurity risks are communicated using four complementary figures of data visualization. The distribution of cybersecurity severity categories is shown in a bar chart, with emphasis on the representation of various threat levels. A scatter plot is used to look at the relationship between CVSS and EPSS scores, and to differentiate threat profiles by attack sophistication. A grouped bar chart or a heatmap compares the severity patterns of the targeted AI elements. A threat prioritization is then followed by an integrated risk visualization, which takes threat severity, threat exploitability, threat impacts and threat components into consideration. The use of suitable graphical packages to develop visualizations is appropriate, and labels, formatting and legends are clear and readable. Each figure serves a particular analytical purpose and is part of the interpretation of the research objectives. The visualizations are not predictive models trained or validated, but rather are descriptive and exploratory. They are descriptive and exploratory rather than outputs from a trained or validated predictive model.

2.6 Ethical Considerations

The data set used in this research is a secondary cybersecurity threat intelligence data set and there are no human subjects and no collection of personal information from human subjects involved in the research. Ethical practice thus focuses on the responsibility of managing data and data sources, the capacity to report on the source data accurately, the transparency of analytical methods followed, and the ability to acknowledge the origin

of the data set. Analysis does not change observations to the extent that it alters the reported threat characteristics. Cautious interpretation of findings is provided, especially when there is a lack of information or imbalance in representation of threat categories, which could impact comparisons. This study does not assert that the observed patterns indicate causes, the likelihood of threat occurrence in the future, or direct loss quantification in business. Any possible security concerns are discussed at an appropriate level and not as "exploratory associations" that are presented as "operational predictions". These actions promote research integrity and responsible awareness of AI risks to cybersecurity.

2.7 Design of Conceptual Framework

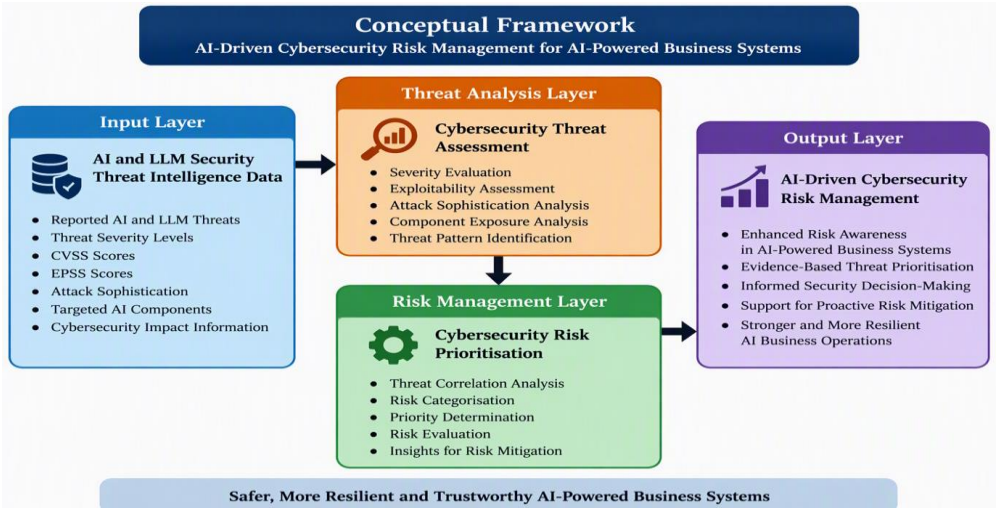


Figure 1. Conceptual Framework Diagram

(Source: Created by learner)

3. Results

3.1 Distribution of Cybersecurity Risk Severity in AI and LLM Threats

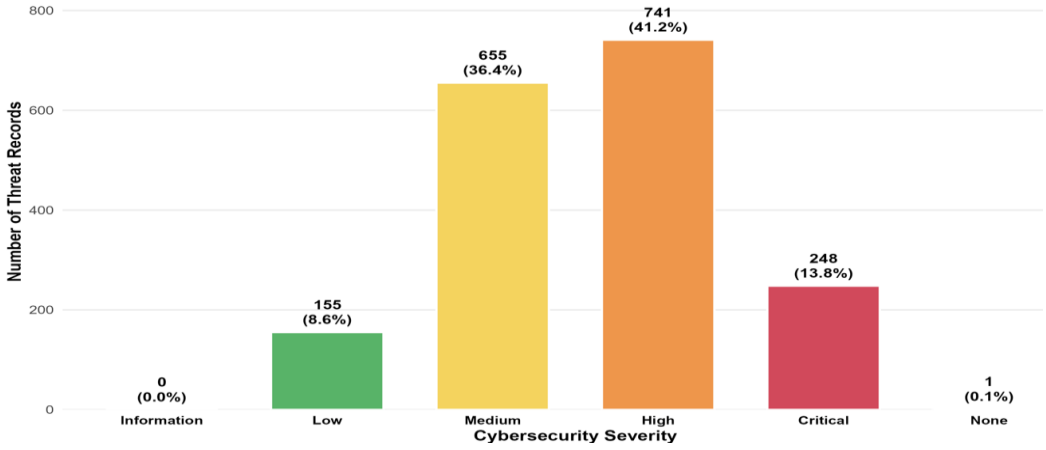


Figure 2. Distribution of Threat Records Across Six Severity Categories

(Source: Created by learner)

The breakdown of the distribution indicates that most incidents recorded are at the High and Medium severity levels, while there are 13.8% at the Critical level. Low-severity events constitute a small percentage – almost negligible in the case of Information, None. This pattern suggests that for the most part, the AI/LLM environment is highly susceptible to threats that can impact its operations and security in meaningful ways. The high percentages documented under each severity level emphasize the importance of systematically recognizing and prioritizing major cybersecurity risks.

3.2 Exploitability and Attack Sophistication across Cybersecurity Threats

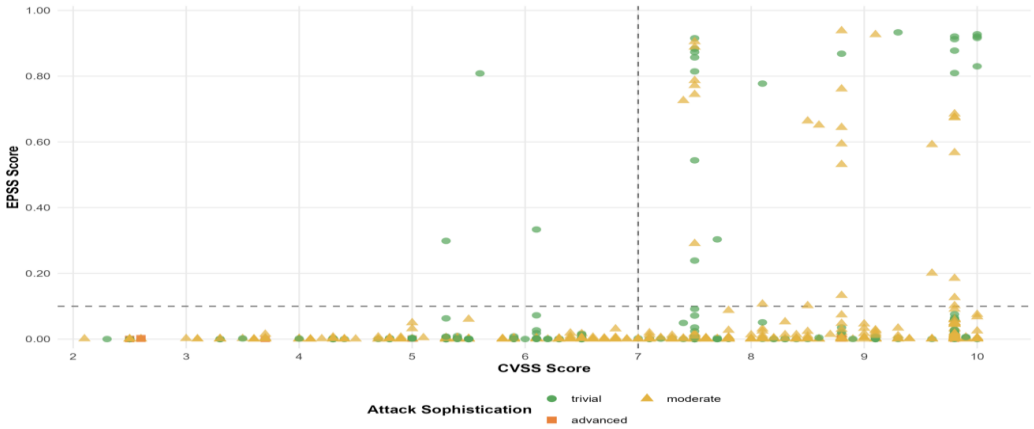


Figure 3. CVSS Severity Score versus EPSS Exploitation Probability

(Source: Created by learner)

The scatter plot shows that vulnerabilities with high CVSS scores often also have a wide range of EPSS scores, suggesting that there are vulnerabilities with high CVSS scores that do not have the same EPSS scores. The majority of threat records lie in the upper half of the CVSS, indicating that there is significant impact potential across the AI-related environment. Also, threats that have moderate/advanced attack sophistication seem to be more prevalent at higher exploitability and severity levels. The results offer a more granular perspective on the varying degrees of risk exposure, based on the severity of vulnerability and the capability of the attacker, helping to inform threat assessment and prioritization decisions.

3.3 Cybersecurity Risk Patterns across Targeted AI Components

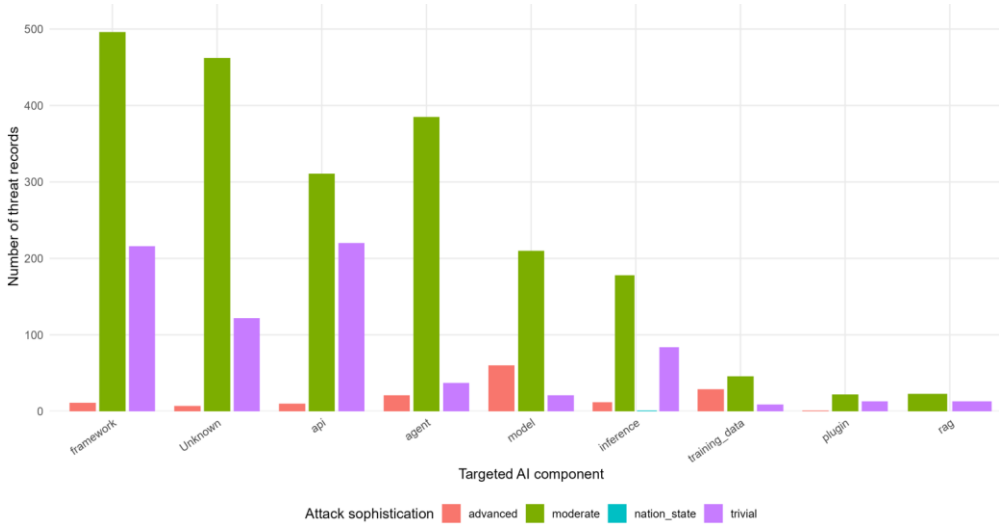


Figure 4. Record Counts Show the Distribution of Sophistication Categories

(Source: Created by learner)

Based on the analysis, there is a significant difference in the distribution of threats among targeted AI components. The highest occurrence of cyber incidents is found in LLMs, agents, vector databases, and APIs, especially in categories of moderate and trivial sophistication. In the meantime, model-level attacks show a higher percentage of higher-level sophistication events. The differences suggest that the threat of risk is not equal across all AI ecosystems and that the elements of particular AI systems are exposed to specific threat characteristics, necessitating specific security focus and protection strategies.

3.4 Integrated Cybersecurity Risk and Threat Prioritisation Patterns

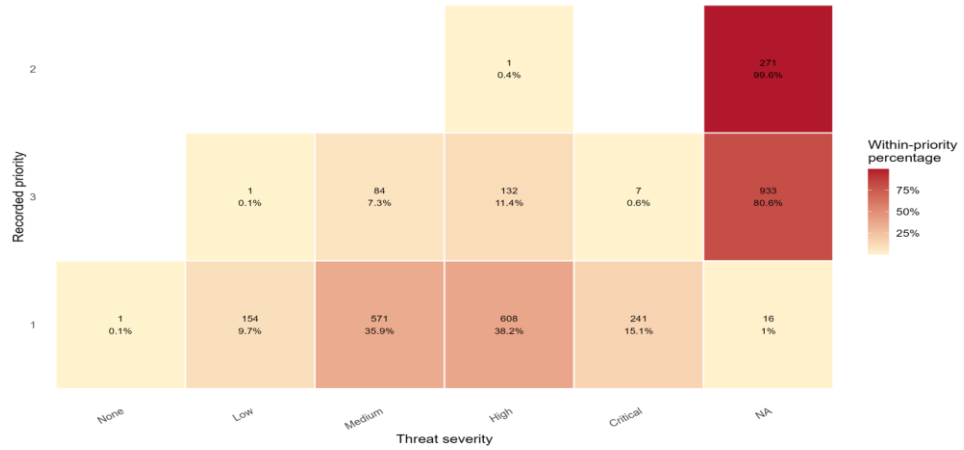


Figure 5. Cell Labels Show Record Counts and Percentages within Each Priority Category

(Source: Created by learner)

The heat map shows a high density of medium and high severity threats across Priority 1 and Priority 3, and high severity threats across Priority 2. High-priority classifications are also prominently represented by critical threats. An interaction between severity, exploitability, and impact seems to influence the prioritization decision, rather than a single factor. The patterns observed reinforce cybersecurity risk assessment activities that are structured and based on evidence.

4. Discussion

It is found that the threats to the security of AI-based business systems are mostly medium severity and high severity, along with a significant number of critical threats. This distribution suggests that the security problems that arise in an AI environment are not merely operational issues but can impact not only the security of the system but also its confidentiality and integrity. The results also demonstrate that severe threats are not just a singularity but a significant part of the threat picture. This focus reflects the need for a structured approach to understanding threat characteristics and underscores the need for careful consideration and prioritization of risks within AI and LLM ecosystems to ensure effective security management.

Further analysis of the exploitability and attack sophistication of these threats can give added understanding of their nature. The correlation between CVSS scores and EPSS values indicates that some technically severe vulnerabilities may have varying levels of likelihood of exploitation. That indicates that severity by itself is not enough to grasp the overall risk status of a threat. Moderate and advanced attack sophistication is seen with higher-scoring vulnerabilities, showing that the capability of the attackers has a significant influence on overall attack exposure. Moreover, the distribution of threats by the targeted AI components shows an unequal distribution of risk. Some components have a higher proportion of incidents, while others have advanced attack techniques targeted at them more. The insights gained from these observations help to better understand the impact of cybersecurity events on various aspects of AI and the nature of these events in the AI landscape.

This is reinforced by the integrated risk and threat prioritization analysis, which indicates that there is a relationship between high-priority classifications and the medium-high and critical severity threats. The high number of severe incidents in priority categories suggests that prioritization is a result of a combination of impact, exploitability and threat characteristics, not just any one of these. The results give a holistic picture of cybersecurity risk patterns of AI-based systems. The outcomes reveal significant connections between severity, exploitation potential, attack sophistication, targeted AI components, and

prioritization results, aiding in making more informed and evidence-based cybersecurity decisions in AI-empowered business environments. .

4.1 Future Work

This research can be extended to future research, with the inclusion of more diverse and larger threat intelligence sources based on various industries, organizational environments, and geographical areas. An analysis of the datasets over a number of years could be used to identify patterns of attacks and changes in the behavior of attackers towards the AI systems. More in-depth analysis can be carried out to explore individual components of AI in more detail and identify component-specific vulnerabilities and security weaknesses. Further research can also investigate the link between threat characteristics and other indicators of organisational impact including service disruption and the effectiveness of organisational responses in terms of resilience and recovery. The fusion of sophisticated statistical techniques and machine learning-powered analytical tools may unlock more insights into the underlying patterns in cybersecurity threat data. Future studies could look into the impact of new technologies, such as autonomous agents, multimodal AI systems and connected business platforms on cybersecurity exposure. Practical relevance and understanding of risk management strategies can be further enhanced by validation of findings through real-world organisational case studies. The initiatives can help build more holistic cybersecurity frameworks for assessing, prioritizing, and managing cyber threats in increasingly complex business environments powered by AI.

5. Conclusion

In this study, cybersecurity threat patterns impacting AI-driven business systems were explored and analyzed by examining crafted AI and LLM threat intelligence data. The results reveal that medium-high and critical severity threats make up a significant number of reported incidents, and exploitability and attack sophistication are different across threat categories. The analysis also reveals that the risks associated with cybersecurity vary across the components of AI, with some being more vulnerable to significant threats. Also, the threat prioritization seems to be affected by the extent of the threat, the likelihood of exploiting it, and the severity of the impact. The study integrates these interdependent elements, gaining insight into the current state of AI cybersecurity. The findings are useful for cybersecurity experts to assess threats, determine susceptible items, and aid in systematic risk assessment methods. The research helps to broaden the understanding of cybersecurity issues related to AI-based business environments and identifies the areas that need further security focus.

REFERENCES

[1] OWASP GenAI Security Project, “OWASP Top 10 for LLM Applications 2025,” OWASP Foundation, 2025. [Online]. Available: https://genai.owasp.org/llm-top-10/

[2] G. Agathangelos-Ioannou and A. Tsohou, “Methodologies for information security risk assessment: shortcomings for analysing artificial intelligence systems,” International Journal of Information Security, vol. 25, Art. no. 98, May 2026, doi: 10.1007/s10207-026-01271-9.

[3] R. Gupta, K. Nair, M. Mishra, B. Ibrahim, and S. Bhardwaj, “Adoption and impacts of generative artificial intelligence: Theoretical underpinnings and research agenda,” International Journal of Information Management Data Insights, vol. 4, no. 1, Art. no. 100232, 2024, doi: 10.1016/j.jjimei.2024.100232.

[4] R. Reznikov, “Leveraging generative AI: Strategic adoption patterns for enterprises,” SSRN Electronic Journal, 2024. [Online]. Available: https://doi.org/10.2139/ssrn.4851632.

[5] C. Jeong, “A study on the implementation of generative AI services using an enterprise data-based LLM application architecture,” arXiv preprint arXiv:2309.01105, 2023. [Online]. Available: https://arxiv.org/abs/2309.01105.

[6] V. Gupta and H. Yang, “Generative artificial intelligence (AI) technology adoption model for entrepreneurs: Case of ChatGPT,” Internet Reference Services Quarterly, vol. 28, no. 2, pp. 223–242, 2024, doi: 10.1080/10875301.2023.2300114

[7] Y. Gan et al., "Navigating the risks: A survey of security, privacy, and ethics threats in LLM-based agents," arXiv preprint arXiv:2411.09523, Nov. 2024. [Online]. Available: https://arxiv.org/abs/2411.09523.

[8] S. Singh, F. Abri, and A. S. Namin, "Exploiting large language models (LLMs) through deception techniques and persuasion principles," in Proc. 2023 IEEE Int. Conf. Big Data (BigData), 2023, pp. 2508–2517, doi: 10.1109/BigData59044.2023.10386814.

[9] D. Pape, S. Mavali, T. Eisenhofer, and L. Schönherr, "Prompt obfuscation for large language models," in Proc. 34th USENIX Security Symp. (USENIX Security 25), 2025, pp. 2323–2342. [Online]. Available: https://www.usenix.org/conference/usenixsecurity25/presentation/pape.

[10] Z. Deng, G. Meng, K. Chen, T. Liu, L. Xiang, and C. Chen, "Differential testing of cross deep learning framework APIs: Revealing inconsistencies and vulnerabilities," in Proc. 32nd USENIX Security Symp. (USENIX Security 23), 2023, pp. 7393–7410. [Online]. Available: https://www.usenix.org/conference/usenixsecurity23/presentation/deng-zizhuang

[11] T. K. Chowdhury, "AI-powered deep learning models for real-time cybersecurity risk assessment in enterprise IT systems," ASRC Procedia: Global Perspectives in Science and Scholarship, vol. 1, no. 1, pp. 675–704, 2025, doi: 10.63125/137k6y79.

[12] B. R. Chirra, "Predictive AI for cyber risk assessment: Enhancing proactive security measures," International Journal of Advanced Engineering Technologies and Innovations, vol. 1, no. 4, pp. 505–527, 2024.

[13] C. T. Udeze, O. Falebita, and S. T. Fasasi, "Artificial intelligence for automated cybersecurity risk assessment," SAMRIDDHI: A Journal of Physical Sciences, Engineering and Technology, vol. 17, no. 4, pp. 21–32, 2025.

[14] H. Raza, "Proactive cyber defense with AI: Enhancing risk assessment and threat detection in cybersecurity ecosystems," 2021. [Online]. Available: https://www.researchgate.net/publication/384323201_Proactive_Cyber_Defense_with_AI_Enhancing_Risk_Assessment_and_Threat_Detection_in_Cybersecurity_Ecosystems

[15] T. J. Luu, *AI Sec Watch Dataset: AI/LLM Security Threat Intelligence*, Version 3.0. Zenodo, Apr. 11, 2026. Available: https://doi.org/10.5281/zenodo.19519470