Topic: Artificial Intelligence in cybersecurity

Sri Sai Sathvik Pidikiti

New Jersey Institute of Technology, College of Humanities and Social Sciences

English 102, Section 074

Professor Gabrielle Rossi

March 29th, 2026

Abstract

The usage of artificial intelligence is influencing the evolution of cybersecurity as AI has made advancements in detection, analysis, and response to Cyber Security Threats. The ability for AI systems to process large datasets, find patterns in the data, and respond to attacks in real time gives them an advantage over traditional security methods. AI can predict threats to an organization and identify them before they become attacks. Another thing is that organizations can use AI as a tool for predictive analysis, using it to identify potential threats before they occur. The effectiveness of AI-enabled security capabilities depends on how well the technology is being utilized and overseen by organizations.

However, AI comes with new risks such as cybercriminals using these systems to automate their cyber attacks, create more complex phishing schemes, and find weaknesses through adversarial attacks. The dependency on large amounts of data to train and build AI creates questions about privacy and ethics, and excessively using AI could lead to reduced human participation, thereby increasing the likelihood of human error and creating potential oversight of undetected vulnerabilities. A combined approach of utilizing both AI and human involvement will be the most effective way to create strong, reliable cybersecurity platforms.

1. Introduction

With significant technological advancement in the modern world has asked for cybersecurity personnel to step up massively. The majority of all companies depend on electronic systems, and their dependence leads to an increase in risk of a cyber attack (Alowaidi et al., 2023). As the attack methods are evolving every now and then, traditional approaches to secure these systems will not be effective (Karki et al., 2024). Moreover, Generatives AI models have

been developed and trained based on millions upon millions of known cyber attacks, which have helped to produce more accurate representations of a potential cyber attack (Ahmed et al., 2024). However, with great power also poses great responsibility. As cyber attacks have a direct proportional relationship with AI advancements, a constant rush between both organizations protecting their systems and individuals attempting to breach them has been ignited. With this information in mind, companies must understand the relationship between these two entities to create effective cybersecurity plans.

2. Background Information

2.1 Evolution AI in Cybersecurity

In the past, cybersecurity has developed so much. Even though old systems did well against known threats they weren't able to detect the new and unknown threats because the traditional system had fixed rules and were monitored by humans. When Machine Learning (ML) was introduced, systems were designed to learn from a particular data set, rather than use traditional rules to detect cyber attacks. These ML models learn from the data and are able to detect unusual behavior or patterns that indicate a cyber attack.

Artificial intelligence (AI) has developed in many areas of cybersecurity, such as malware analysis and threat intelligence. Generative AI can continue to improve itself by learning from new data; this makes AI more effective against evolving cyber threats, generates predictive analytics, and even automated responses to potential cyber threats. Because of this counterattack, updated systems are now more versatile and respond to threats more quickly than before.

3. Risks of AI in Cybersecurity

3.1 Adversarial Attacks

Adversarial attacks pose a serious risk to AI-based cybersecurity systems because they affect how AI learning algorithms develop by changing the training dataset by machine learning algorithms, for example, adding false/misleading data because of this AI's accuracy to identify legitimate threats in trust, some adversarial attackers can simply bypass an entire AI security system (Karki et al., 2024; Kasakliev et al., 2024). This creates a major challenge for organizations that rely on AI models needs more research and development.

3.2 Automated Hacking and Increased Attack Efficiency

Thanks to improved AI technology, cyberattacks have improved both speed and effectiveness. Activities associated with managing cyber attacks along with network reconnaissance and vulnerability identification, are now automated using AI-based automation technologies, reducing the amount of time and effort required to execute these processes. This further causes a state of panic and less time is given for cybersecurity individuals to address the incoming attacks. According to Gupta et al. (2023), generative AI can also generate scripts and strategies used in executing cyber attacks, allowing for further reductions in time and effort associated with executing an attack.

Not only do cyber attacks require less-planning, they can target multiple systems simultaneously. As more and more separate attacks can occur simultaneously, the number of

potential attack targets becomes significantly increased thus increasing vulnerability for the targeted organization and increasing the risk to their information assets (Michael et al., 2023).

3.3 Advanced Phishing and Social Engineering

Phishing attacks are gradually more robust in terms of AI technology. In the past, it was easy to identify phishing attacks via the poor spelling and grammar typically found in phishing emails and other types of communication. Universities and many institutions receive many spam emails from hackers and thieves who try to trap students with reward emails, and try to generate emails similar to those of companies, which are harder to differentiate from the official emails. (Michael et al., 2023).

Therefore, there is a greater likelihood that users will "click" on a bad link or provide sensitive and/or important information to the phishing attacker. As phishing attacks continue to move toward a more advanced stage, it will continue to become increasingly difficult for users to determine whether emails/messages they receive/will receive are "legitimate" or "fake" (Kasakliev et al., 2024).

3.4 Data Privacy and Ethical Concerns

Before creating AI systems with the use of data, it will be important to make sure there are sufficient steps in place to create privacy protections (Taddeo et al., 2019). While training the AI models in real-world , they require huge amounts of data, which might be sensitive data, including personal information about individuals and financial information about businesses. Failure to properly secure sensitive data will lead to privacy violations.

According to Taddeo et al. (2019), trust is a major issue for AI-powered cybersecurity systems. Users need to feel confident that when their data is utilized by organizations, they are doing so in a responsible manner. Another area that will require careful consideration is the legal concerns regarding how AI is being utilized. In my opinion, organizations must make sure their AI systems are fair and secure, and they must implement sufficient rules and regulations to help reduce these issues.

4. Benefits of AI in Cybersecurity

4.1 Real-Time Threat Detection and Response

AI's most significant benefits are how it identifies and responds to threats as they occur. AI systems are always analyzing network traffic and looking for suspicious behavior or activities. After detecting, the system will alert appropriate personnel, which allows for quick response to minimize the overall impact of the attack by the security teams such as blocking or isolating the threat to prevent any further damage(Ahmed et al., 2024). AI is essential when it comes to threat detection in cyber because it can process large amounts of data much faster than humans.

4.2 Predictive Analysis and Proactive Security

With AI being used within organisations it can transition from a reactive security posture to a proactive one (Zacharis et al., 2024). AI systems have the ability to forecast the likelihood of future attacks due to its capability of predicting a possible threat by looking back at past data and recognizing patterns that can be helpful to detect a future attack. Such predictive analysis creates an opportunity for the organisation to prepare accordingly to strengthen their overall security

(Zacharis et al., 2024). This ability is also crucial in increasing security measures along with diminishing the impact of cybercrime.

4.3 Role of Generative AI in Cyber Defence

Generative AI tools in cybersecurity aid with threat analysis, report generation, and recommendations for action. The use of these tools allows organizations to make decisions using large amounts of data to choose the best option, which has increased the efficiency of cybersecurity teams (Gupta et al., 2023). Furthermore, generative AI will help organizations train their employees by simulating real-life cyber attacks so that their cyber security practitioners are prepared for actual events (Kasakliev et al., 2024).

4.4 Machine Learning for Malware and Intrusion Detection

Cybersecurity has recently begun to utilize machine learning for the purpose of detecting malware and intrusions (Martínez Torres et al., 2019). Machine learning models are developed by training on both data from known attacks as well as benign data (Martínez Torres et al., 2019). Once the model has been trained, it is capable of classifying new data, allowing it to identify potential threats. Machine Learning has been an effective approach to detecting attacks that have not been seen before (Handa et al., 2019).

4.5 Email Security and Spam Detection

Additionally, AI is used in securing emails. Email spam is identified through the use of machine learning to evaluate email content (Handa et al., 2019). Bag of words models assist in

the identification of the many different "keywords" associated with spam emails. Any suspicious email that matches one of those keywords will be filtered out and moved to the spam folder (Handa et al., 2019). Filtering these kinds of emails is one of the most common and prevalent uses for AI technology in relation to cybersecurity (Handa et al., 2019). By filtering emails, it helps to reduce the amount of interaction between users and potentially malicious content.

5. Challenges Organizations Face When Using AI

While AI has benefits, there are several problems faced by organisations implementing AI into cybersecurity systems (Michael et al., 2023);

1. Cost is a major challenge since building and maintaining AI systems will require much more resources than small organisations will typically have available to them and may have no chance of adopting said technologies (Alowaidi et al., 2023)
2. Skilled professionals that can implement and maintain AI-based cybersecurity systems are in short supply; they must have expertise in cybersecurity, as well as machine learning data science (Karki et al., 2024)
3. Over-reliance on AI presents yet another challenge; while machine learning data is very reasonable for predicting outcomes, it cannot be counted upon to produce perfect predictions, so human oversight will still be required to confirm that any predictions made by a machine learning model are correct and dependable(Michael et al., 2023).

6. Future Directions of AI in Cybersecurity

Cybersecurity's future will be affected a lot by AI technology (Hoza, 2025). Advances and developments are being accomplished by researchers worldwide to create more advanced and secure AI technologies (Zacharis et al., 2024).

A significant branch of AI development will be improving the robustness of machine learning models against attacks, i.e., providing robust defences against various forms of adversarial attack (Kasakliev et al., 2024). Another major branch of AI development is the introduction of ethical guidelines and regulations on AI usage to make sure that AI is deployed in both a secure and responsible manner (Hoza, 2025). Furthermore, technology will continue to improve in the areas of automation, predictive analysis, and threat intelligence; as this happens, the effectiveness of cybersecurity systems will increase (Ahmed et al., 2024; Zacharis et al., 2024).

7. Conclusion

The ability to detect threats has improved dramatically, and the time it takes to respond has also decreased while the overall performance of the systems being secured as increased. However, as Artificial Intelligence has evolved, it has also produced new types of threats i.e. Adversarial attacks, Automated Cyber Attacks and the violations of Privacy. A critical factor in the development of any Cybersecurity solution is the combination of Artificial Intelligence with Human Intelligence to ensure accurate decision making and responsible use. As Artificial Intelligence continues to evolve within the Cybersecurity arena to create a safe Digital World, there must be adequate oversight, responsible usage, and the implementation of appropriate security strategies that guide the developmental evolution of Artificial Intelligence.

References

Ahmed, M., Okba, K., & Harous, S. (2024). Generative Artificial Intelligence for Cyber Security: Literature Review and Challenges. *Scientific Journal University of Saba Region*, *7*(2), 425–470. https://doi.org/10.54582/TSJ.2.2.109

Alowaidi, M., Sharma, S. K., AlEnizi, A., & Bhardwaj, S. (2023). Integrating artificial intelligence in cybersecurity for cyber-physical systems. *Electronic Research Archive*, *31*(4), 1–21. https://doi.org/10.3934/era.2023097

Gupta, M., Akiri, C., Aryal, K., Parker, E., & Praharaj, L. (2023). From ChatGPT to ThreatGPT: Impact of Generative AI in Cybersecurity and Privacy. *IEEE Access*, *11*, 1–1. https://doi.org/10.1109/ACCESS.2023.3300381

Handa, A., Sharma, A., & Shukla, S. K. (2019). Machine learning in cybersecurity: A review. *Wiley Interdisciplinary Reviews. Data Mining and Knowledge Discovery*, *9*(4), e1306-n/a. https://doi.org/10.1002/widm.1306

Hoza, K. (2025). AI in cybersecurity. *Zeszyty Naukowe Wyższej Szkoły Finansów i Prawa w Bielsku-Białej, 29*(3). https://doi.org/10.19192/wsfip.sj3.2025.4

Karki, S., Hasan, A. B. M. M., & Sanin, C. (2024). Use of ML and AI in Cybersecurity-A Survey. *Procedia Computer Science*, *246*, 1260–1270. https://doi.org/10.1016/j.procs.2024.09.552

Kasakliev, N., Somova, E., & Gocheva, M. (2024). Artificial Intelligence for Good and Bad in Cyber and Information Security. *Mathematics & Informatics*, *67*(1), 82–94. https://doi.org/10.53656/math2024-1-6-art

Martínez Torres, J., Iglesias Comesaña, C., & García-Nieto, P. J. (2019). Review: machine learning techniques applied to cybersecurity. *International Journal of Machine Learning and Cybernetics*, *10*(10), 2823–2836. https://doi.org/10.1007/s13042-018-00906-1

Michael, K., Abbas, R., & Roussos, G. (2023). AI in Cybersecurity: The Paradox. *IEEE Transactions on Technology and Society*, *4*(2), 104–109. https://doi.org/10.1109/TTS.2023.3280109

Taddeo, M., McCutcheon, T., & Floridi, L. (2019). Trusting artificial intelligence in cybersecurity is a double-edged sword. *Nature Machine Intelligence*, *1*(12), 557–560. https://doi.org/10.1038/s42256-019-0109-1

Zacharis, A., Katos, V., & Patsakis, C. (2024). Integrating AI-driven threat intelligence and forecasting in the cyber security exercise content generation lifecycle. *International Journal of Information Security*, *23*(4), 2691–2710. https://doi.org/10.1007/s10207-024-00860-w