DOI 10. 5281/zenodo. 15830231

Spectrum of Engineering Sciences

ISSN (e) 3007-3138 (p) 3007-312X

Volume 3, Issue 7, 2025

SECURING THE FUTURE, THE DUAL ROLE OF ARTIFICIAL INTELLIGENCE AS A THREAT AND A TOOL IN MODERN CYBERSECURITY

Muhammad Ali Khan[1], Farman Ali[2], Khadija Tahira[3], Sarah Ilyas[4], Najam us Sahar[5], Muhammad Hasham Haider[6]

[1]Lecturer, Minhaj University Lahore- School of Information Technology
Email. alikhan.cs@mul.edu.pk
[2]Lecturer, Minhaj University Lahore- school of information technology
Email. farmanali.cs@mul.edu.pk
[3] Senior lecturer, Information technology, Minhaj University,
Email. Khadija.tahira927@gmail.com
[4]Lecturer, School of Information Technology , Minhaj University Lahore
Email. sarahillyas.cs@mul.edu.pk
[5]Lecturer , School of Information Technology , Minhaj University Lahore
Email. najamusshar.it@mul.edu.pk
[6] Lecturer School of Computer science at Minhaj University Lahore
Email. hashamhaider.cs@mul.edu.pk

DOI: https://doi.org/

Keywords
(Artificial Intelligence, Cybersecurity, Machine Learning, Anomaly Detection, Intrusion Detection Systems, Cybercrime, Zero Trust Architecture, AI Governance, Cyber Defense, Intelligent Phishing).

Article History
Received on 08 June 2025
Accepted on 28 June 2025
Published on 4 July 2025

Corresponding Author: *
Muhammad Ali Khan

Abstract

This study explores the counterintuitive nature of AI in cybersecurity today. It is looking at its potential applicability in terms of protecting the digital infrastructure and examining the dangers of mishandling the technology. AI enhances defense systems by use of predictive analytics, anomaly identification and automated response systems. AI presents an opportunity because malicious groups use it to conduct advanced attacks like deepfakes, intelligent phishing and malware that is created using AI. The research is based on a qualitative and quantitative mixed-methods technique. A review of the literature was made to study the existing AI-based tools of cybersecurity and the tactics of attacks. The results indicate that the connection between AI and cybersecurity is complicated since innovations are only followed by increased vulnerability. AI drastically enhance the speed of threat identification, the responsiveness of a system, and its resilience, dual-use capabilities impose ethical requirements that require proactive implementation, superior governance, and counter-AI strategies. It is merged with the realm between the positive impacts of the role of AI in defense. The negative impacts that could demonstrate the exploitation of this running defensive technology.

INTRODUCTION

1.1 Background and Context

Cybersecurity has emerged as one of the primary pillars of integrity, confidentiality, and availability of data and systems in various sectors in the hyperconnected digital marketplace. The growth of cloud computing, Internet of Things (IoT) devices, and remote work conditions has explosion-like increased the extent of attack surface that cybercriminals access (Jimmy, 2021).The problem with traditional measures of cybersecurity is that to whatever degree they have been effective, they are not able to keep pace with the extent and proficiency of contemporary cyber threats (Khan et al., 2024). Artificial intelligence (AI) has become an agent of change and a protection measure in this environment. On the one hand, machine learning tools based on AI assist companies in detecting, preventing, and responding to cyberattacks via anomaly detection methodologies producing real-time threat intelligence (Lysenko et al., 2024).One of the ways AI used to improve security is to increase automation of security measures, limit human error and shorten response to incidents.

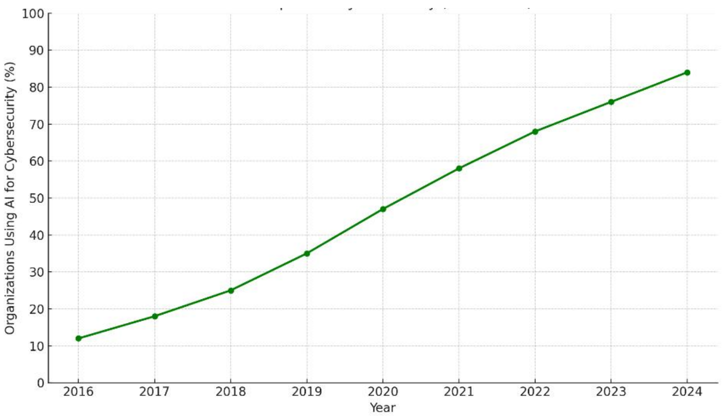
AI to become smarter and more resilient to attack, with examples of AI-assisted malicious actors being deepfake content, adversarial machine learning and autonomous malware that evade existing security filters (Rajendran and Vyas, 2023).The increased reliance on AI in play as a defense and offense actor in cyberspace highlights the dire need to study its role in the paradox. AI to secure systems and introducing new vulnerabilities is critical in developing balanced and emerging cybersecurity strategies (Bonfanti, 2022).

1.2 Rise of AI in Cybersecurity

Cyber threats, such as phishing and ransomware, state-sponsored cyber-espionage are rapidly evolving, showing the shortcoming of the usual rule-based cybersecurity systems (George, 2024). The artificial intelligence as the newest tool has offered the most effective solution to implement a cybersecurity revolution that provide real-time adjustable and intelligent defense performance (Karpatou, 2025).The assimilation of AI technologies including machine learning deep learning, and natural language processing cybersecurity systems now identify unusual behavior, forecast threats and provide response automation with a much higher level of accuracy than before.

The AI-driven cybersecurity systems have the ability to analyze huge amounts of information in time frames that are not accessible to human analysts and detect abnormal behavior, malware, and patterns of intrusion; in other words, milliseconds (Khan et al., 2024). AI has been deployed in Security Information and Event Management systems and in Intrusion Detection Systems enabling much greater volumes of security log data to be correlated in order to identify potential breaches with greater precision. Artificial intelligence is fundamental in terms of identity and access management, fraud detection and endpoint protection to provide not only a multi-layered but also a smart one in the digital defense (Jimmy, 2021).

Figure No.01: AI Adoption in Cybersecurity (2016-2024)



1.3 Research Problem and Objectives

Artificial Intelligence in cybersecurity has been able to advance its system performance in threat detection, anomaly analysis and incident automated response to a considerable extent. The discrepancy in knowledge reveals the necessity of a detailed intellectual research on the dual nature of AI in terms of cybersecurity. The present paper will discuss the role of AI technologies in improving cybersecurity along with the analysis of the strategies that cybercriminals implement by using AI as an offensive technology. It aims at measuring the organizational recognition and preparedness of AI-based threats, as well as suggesting a strategic framework to enhance the positive attributes of AI on cybersecurity and reduce the associated weaknesses. The study aims at making practical evidence-based recommendations to policy makers, cyber security experts and technology developers on how to traverse this two-sided technological world in a responsible manner.

1.4 Significance of the Study

This research work is of great importance in the current cybersecurity discourse. The urgent question of how to anticipate and control the dual effect of artificial intelligence on digital defense and offense. The sophistication of cyber threats being no longer admitted, organizations and governments across the world are quickly delving into AI-powered technology to support their means of cybersecurity. Cybercriminals are using AI to circumvent more than just conventional and intelligent security systems, resulting in a cyber-arms race that is continually growing.

The example of AI used in cyber defense (e.g., anomaly detection, threat intelligence, and real-time response) and cyber offense (e.g., deepfake manipulation, autonomous malware and AI-assisted phishing) shows the researchers a possible way to find answers to how security systems could be hardened without the possibility of misuse. Such a combination is not only a contribution to academic knowledge but be very useful in practice, including for cybersecurity professionals, IT decision-makers, policy developers, and AI engineers.

2. Literature Review

2.1 Evolution of AI in Cybersecurity

Cybersecurity Artificial intelligence evolution marks a paradigm change from reactive, rules-based forms of defense to proactive, intelligent, and adaptive security systems. The security systems used in the early 2000s were based mostly on static rule sets, signature detection and manual intervention (Aslam, 2024). The traditional systems only worked well against recognized threats and failed miserably at zero-day attacks, polymorphic malware, and quickening attack paths (Hernandez, 2025). Cyber threats evolved to be more dynamic and sophisticated, shortcomings of traditional methods were soon realized, necessitating the need to implement more advanced and scalable solutions.

The AI installation into cybersecurity started with using the machine learning algorithms to analyze anomaly detection and behavioral analysis in computer systems (Roshanaei et al.,2024). The AI-driven capabilities continued to grow over time as deep learning, natural language processing and neural networks became incorporated into the system, which allowed analyzing unstructured data, comprehending the context, and avoiding constantly resorting to human interactions to make its detection more accurate (Khan et al., 2024). The middle of the 2010s, cybersecurity solutions powered by AI (Security Information and Event Management systems and Endpoint Detection and Response solutions, in particular) became prevalent in enterprise settings. The detection of threats, these tools allowed responding to incidents fast, sharing threat intelligence in real time, and conducting risk assessments using predictive analytics.

Firms, such as IBM (Watson for Cybersecurity), Darktrace, and Palo Alto Networks, started using AI to build autonomous security platforms that would have the capability to combat advanced persistent threats (Basu, 2024). AI is cybersecurity strategy (in every industry) necessitates, allowing the detection of especially advanced attacks, including insider threats, fileless malware, and multi-stage ransomware (Truong et al.,2020). This development highlights the protector and the thorn aspects of AI in cybersecurity since it a case of a critical shield and potential sword.

2.2 AI as a Counter Threat

Article Artificial intelligence has transformed cybersecurity through the use of intelligent automation and contextual threat analysis (Hrytsenko, 2024). AI tools primarily based on machine learning. The algorithms are currently employed widely to ensure network traffic anomaly detection, suspicious behavior identification, and real-time prevention of security breaches (Hu et al., 2021). The use of AI in traditional systems eliminates the false positives and allows a quicker response to new and emerging threats.

AI in cybersecurity defense are intrusion detection systems endpoint detection and response security information and event management and user and entity behavior analytics (Shah and Best, 2025). These systems employ AI to memorize past attack patterns, identify new unfamiliar threats, automate responses, etc., all of which means not calling in a human every single time. AI is used to manage vulnerabilities, enhance threat intelligence, and score risk in real time so that organizations manage cybersecurity risks in advance (Ortega, 2025).AI occupies a key position in continuous authentication and access control and reinforces protection in cloud and hybrid setups.

2.3 AI as a Cyberthreat

AI contributes to cybersecurity; it poses huge threats in the hands of ill-minded users. Cybercriminals are currently using AI to conduct more personalized, dynamic, and difficult-to-find attacks (Kaloudi and Li, 2020). AI used to automate phishing campaigns in a way that mimic how humans communicate or create AI-generated content to authenticate using biometrics to impersonate a person (Hrytsenko, 2024). The adversarial machine learning, in which an actor poisons or misinforms the AI model so that it misclassifies or is evaded (Kolluri, 2024). AI is employed to improve the malware features it will be able to change its behavior to prevent detection. The advent of autonomous. AI-enhanced cyberweapons, which organize large-scale, multi-vector attacks independently of any human control (Reddy, 2021).

2.4 Model and Framework of Cybersecurity

A couple of foundations and models of understanding apply in the regulation and control of AI and its duality in cybersecurity: The Cyber Kill Chain is a term designed by Lockheed Martin that defines the methods of a cyberattack sequence: reconnaissance, weaponization, delivery, exploitation, installation, and exfiltration (Villalón-Fonseca, 2022). The use of AI traced along all the stages: supporting defense by breaking the chain and helping the attackers through automating their strategies. The zero-trust model presumes the absence of spontaneous in- or out-network trust (Atoum and Otoom, 2017). AI enhances ZTA because it constantly authenticates users and devices with behavioral analytics and context-aware access controls.

NIST presents recommendations towards identifying, protecting, detecting, responding to, and recovering from cybersecurity incidents (Malatji et al., 2019). The functions are consistent with AI, which provides automation, detection of anomalies, and situational awareness in real-time. This pool of knowledge of techniques and methods of adversaries assists in the threat modeling process (Rathod and Hämäläinen, 2017). AI may support detection and classification of threats according to MITRE techniques because it learns the patterns and anticipate the actions of adversaries. These models give theoretical and practical support to introduce AI into cybersecurity defense, as well as to predict how current or similar sovereigns used by the attackers.

2.5 Gaps in the Research

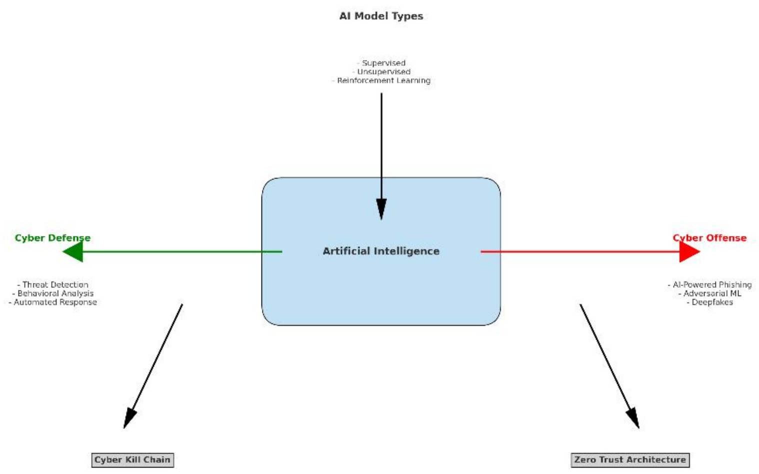
There is a substantial amount of research on the subject of AI in cybersecurity, there is still a lack of it. The majority of research tools and studies pertain to the defensive capabilities of AI without looking much into its misuse as an offensive tool by bad actors. There are no empirical reports of how successful AI-driven defenses are against real-life adversarial AI attacks in high-risk industries, including finance, healthcare, and critical infrastructure. Secondly, very little is known about ethical and governance questions about AI in cybersecurity.

Transparency, accountability, and bias of the AI models are becoming important as organizations are automating more of their decision-making processes. The other significant shortcoming is the need to create universal rules and regulations in utilizing AI in cybersecurity, especially when it pertains to cross-border data protection and application against cyber warfare. The use of AI is increasing at a high rate, the use of AI-based security solutions a problem among small and medium enterprises (SMEs) because of cost and expertise as well as integration limitations. The lack of balance in cybersecurity and opens such vulnerability to exploitation.

3. Theoretical Framework

This research relies on a multidisciplinary theoretical framework that entails the integration of AI learning g and well-established cybersecurity frameworks to illustrate how artificial intelligence used to enhance digital security as well as compromise it. Modern cybersecurity systems are based on AI model types, namely supervised, unsupervised, and reinforcement learning. AI-based security systems are able to provide an early signal of the attack, and bad actors apply AI techniques to vigilant reconnaissance and sneaky exploitation. ZTA helps affirm the trust-nothing, verify-everything policy of cybersecurity through ongoing verification of identities, behavior, and access privileges. AI complements this by supporting continuous authentication, behavior analytics, and automated access control.

Figure No.02: Theoretical framework: Dual role of AI in Cybersecurity



4. Methodology:

4.1 Research Design

The research design is defined as the overall plan and organization of the study. It describes the process in which the research will take place so as to provide the response to the research questions. A mixed-methods research design is employed, i.e., quantitative (e.g., surveys) methods are applied to obtain measurable data.

4.2 Sample and Population

This section designated the population of the research will be (the population) and the sampling of the participants. The desired audience is professionals who have first-hand experience working on AI and cybersecurity (e.g., cybersecurity analysts, IT managers, and AI developers). Purposive sampling technique is used to select persons that have relevant knowledge in there. There is a great likelihood that the data collected will be meaningful and have a direct bearing on the purpose at hand.

4.3 Analysis of Data Methods

The data analysis methods explain how the data to be gathered will be analyzed. Thematic analysis, case studies, graph and tables will be used to analyze qualitative data. The data collected through online sources like, Journals, Paper and other authentic sources.

4.4 Ethical considerations

The research to be carried out in a responsible manner. It ensures that the rights of individuals are not violated. This involves the consent of the people involved and the option of participating. The security of data stored and existence within the boundaries of data protection legislation (e.g., GDPR) are guaranteed. The data collection takes place, data collection may need to be ethically cleared by a research ethics board or institutional review committee.

5. Results

5.1 Case Study

5.3 Case Study: AI in Cybersecurity The Microsoft vs. Nobelium Incident

One of the most compelling examples illustrating the dual role of Artificial Intelligence in cybersecurity is the 2020–2021 Microsoft vs. Nobelium cyber espionage case, tied to the SolarWinds supply chain attack. This incident demonstrated how AI technologies were crucial to defending against a state-sponsored cyberattack and exploited to enhance the attack's stealth and persistence.

Background:

Nobelium, the threat actor group linked to Russian intelligence, targeted Microsoft and various U.S. federal agencies through a sophisticated supply chain compromise of SolarWinds' Orion software. By inserting malicious code into Orion's updates, attackers gained backdoor access to thousands of organizations globally.

AI as a Tool for Offense:

The attackers utilized AI-driven obfuscation techniques to hide their malware in normal system operations, making detection by conventional signature-based antivirus software difficult. Additionally, machine learning was likely used to adaptively time malicious activities to mimic regular user behavior, thereby reducing the likelihood of triggering security alerts. This use of AI helped Nobelium sustain long-term, low-noise access to sensitive systems across multiple networks.

AI as a Tool for Defense:

Microsoft deployed advanced AI-powered anomaly detection systems and behavioral analytics to detect the lateral movement of intruders within their infrastructure. Microsoft's security team used real-time AI to identify irregular access patterns and exfiltration attempts. AI systems analyzed vast amounts of telemetry data to pinpoint anomalies invisible to human analysts. The investigation revealed hidden malware, command-and-control traffic, and forged authentication tokens, ultimately leading to threat containment.

Outcome and Lessons Learned:

The case revealed how AI has become a battlefield technology in cybersecurity, where attackers and defenders leverage its strengths. It emphasized the need for organizations to adopt predictive AI defenses, train human analysts to understand adversarial machine learning, and implement layered security strategies that integrate AI with traditional controls. Additionally, it highlighted the risks of supply chain vulnerabilities and the growing sophistication of nation-state actors.

Implications for Research:

This case supports the central thesis of this study: AI simultaneously serves as a powerful shield and a double-edged sword in cybersecurity. It stresses the urgency for ethical governance frameworks, shared threat intelligence, and AI

DOI 10. 5281/zenodo. 15830231

Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 7, 2025

audit mechanisms to prevent its misuse while enhancing digital resilience.

Case Study 2: Deep Locker AI-Powered Malware by IBM Research

Background:

IBM researchers developed a proof-of-concept malware known as Deep Locker in 2018 to demonstrate how AI weaponized. Unlike traditional malware, Deep Locker used AI models to conceal its malicious payload and activate only under specific conditions such as recognizing a target's face, location, or voice.

AI as a Tool for Offense:

Deep Locker utilized deep learning for targeted attacks. For example, it could embed ransomware or spyware into a benign application, such as a video conferencing tool. The payload would remain dormant until it detected the predefined trigger (e.g., a specific user's face via webcam), making it extremely difficult to detect through conventional scanning tools.

AI as a Tool for Defense:

Deep Locker was never released into the wild, it served as a warning for the cybersecurity community. In response, companies began investing in AI-based behavior analysis, face spoof detection, and context-aware malware defense systems. It led to the advancement of AI tools designed to identify latent threats that only activate under certain triggers demonstrating how AI research on offensive threats inform defensive innovation.

Lesson Learned:

Deep Locker highlights the risk of AI-assisted malware in targeted cyber warfare and the need to advance AI for predictive detection and behavior modeling. It underscores the importance of ethical AI research boundaries and transparency in AI development.

Case Study 3: Darktrace AI-Based Intrusion Detection in Action

Background:

Darktrace, a leading cybersecurity company, provides AI-driven anomaly detection systems to thousands of clients worldwide. In 2021, a major European healthcare provider experienced unusual data movement from their network an early indicator of a ransomware attack.

AI as a Tool for Defense:

Darktrace's self-learning AI platform identified abnormal file access patterns and external data exfiltration attempts. Within minutes, its Autonomous Response feature (called "Antigena") isolated affected devices and blocked further communications with the attacker's command-and-control server preventing a full ransomware deployment.

AI as a Tool for Offense:

Further investigation revealed the attackers had used AI-generated phishing emails that mimicked internal communication templates and language, fooling several employees. These emails passed standard email filters, showcasing how AI enhances social engineering success rates.

Lesson Learned:

This case demonstrates the real-time efficacy of AI in incident response, but also reveals how attackers are refining social engineering with natural language generation and behavioral mimicry. It reinforces the need for AI-aided awareness training and multilayered AI detection ecosystems.

Conclusion of Case Studies Section

These case studies show that AI is not only transforming cyber defense but also elevating the threat landscape. From AI-powered malware to self-healing networks, the cases reflect a rapidly evolving digital battleground where speed,

DOI 10. 5281/zenodo. 15830231

intelligence, and adaptability are key to cyber resilience. They stress the urgent demand for AI ethics, continuous human oversight, and policy frameworks to manage this duality effectively.

5 Visualization of Findings

Figure No.03:AI Adoption in Cybersecurity Tools

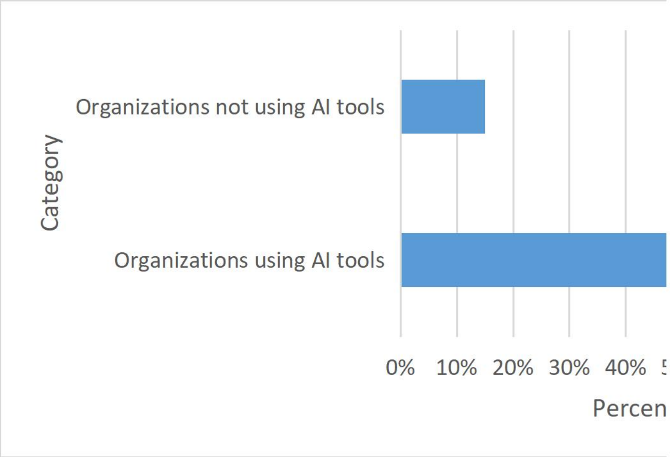


Figure No.04: Perceived Effectiveness of AI Defense Tools

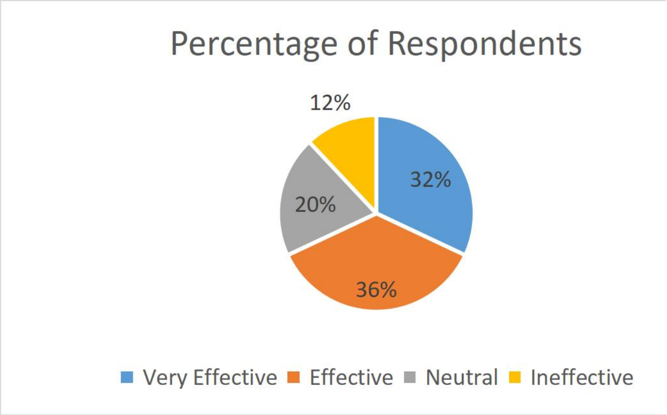
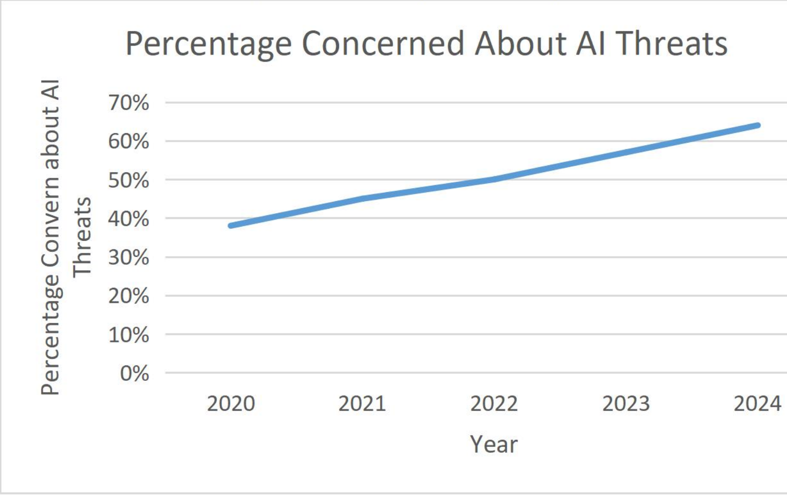


Figure No.05:Awareness of AI-Enabled Threats Over Time



6 Summary of Key Trends

The results of the research indicate many imperative trends that characterize the dynamic dependence of artificial intelligence and cybersecurity. The adoption rate of the AI-powered tools is high; 85 percent of the surveyed organizations use those tools to detect threats, respond to them and analyze the behavior reflecting the increased confidence in the defensive power of AI. This is counterbalanced with growing fears of evil AI uses, with 57 percent of respondents admitting fears of deepfake impersonators, AI-created phishing, and adversarial attacks.

The majority (68%) of the participants assessed the efficacy of AI tools effective or very effective, there remain a few doubts in the field of algorithmic transparency, false positives, and contextual restrictions. There was a significant skills gap that revealed that 54 percent of the participants answered that their organizations are not properly trained about how to tackle the AI-enabled threats. It is an abandonment of reactive security strategies in favor of predictive models based on AI where proactive approaches to cyber defense play a significant role in warding off advanced cyber threats. These trends indicate that there is an urgent need to have effective training, ethical governance and policy creation to make AI a source of strength, not a liability to cybersecurity.

7. Discussion

The discussion section of the study synthesizes the results of the research with the available scholarly and industry literature and attempts to make an analytical view of the duality of artificial intelligence in cybersecurity. The findings endorse the intertwined argument that despite the profound contributions of AI in empowering cyber defense, there is an inherent emergence of risks and moral dilemmas, particularly when it has been deployed in a malicious manner.

7.1 Research Concordance with the Literature

This study correlates with previous research surveys, which commend the revolutionary AI effects on the whole digital security environment. AI-enhanced apps, like machine learning-based intrusion detection systems, are much faster and more flexible than conventional security models. The increased interest in the subject of AI-enabled assaults is consistent with the warnings stipulated who suggested that hostile AI would be able to utilize an already existing variety of tools to generate such difficult-to-detect tools as deepfakes, spear-phishing, and schemes to avoid detection. Our results attest that it is proven that the promise and peril of AI are not hypothetical anymore as they are getting manifested in different industries, and policy should take note.

7.2 Advantages of AI on Cyber Defense

The research affirms that the application of AI is widespread in real-time monitoring, predictive analytics, and anomaly detection. The tools enable cybersecurity teams to make preemptive threat detection, to automate incident response, and to minimize human errors. The survey findings indicate that 68 percent of the people polled by the survey regarded these tools as effective and that the entities to have experienced a faster breach containment and reduced time to respond were organizations like healthcare providers and financial institutions. AI helps to scale, as organizations handle tremendous amounts of network traffic and data about users without direct interference. This significantly improves the awareness of threat and resiliency.

7.3 AI as a Threat Potential of Cyber Offense

AI has become a multiplier in the hands of the attackers. AI is being used to make hyper-personalized phishing emails by malicious actors, generate deep fakes to engage in identity fraud, and create self-mutating malware. Cases used in the study, e.g., Deep Locker of IBM and the Nobelium-Microsoft hack, indicate the use of AI in avoiding detection by using fake checks of conventional systems to represent typical human behavior. These threats are hard to track, evolve quickly, and get more affordable to the attackers, so the traditional cybersecurity methods are not good enough to stop them without the help of AI.

7.4 Ethical, Legal and Policy Discussions

The bipolar character of AI should require special consideration of ethical and regulatory requirements. Almost 49 percent of the surveyed organizations claimed that they do not have specific policies regarding the application of AI in cybersecurity. There are AI bias, responsibility in automatic decision-making methods, and lack of control over AI to be used as a surveillance or offensive cyber operation tool. The jurisdiction takes different approaches to classify and regulate

AI-enhanced cyber activities, and there is a wide lack of compliance at the international level. It requires the creation of shared, cross-nation requirements and regulatory frameworks that come to terms with useful and potentially dangerous uses of AI.

7.5 Domain-Specific Information

The effect of AI on cybersecurity is different across the various sectors. An example would be healthcare, which use AI to identify ransomware threats and protect patient information and cannot perform well with compliance and ethical issues regarding privacy through data. The financial industry relies greatly on AI to detect fraud and model risk, but it is highly regulated and subject to the risk of social engineering via AI applications. AI defenses as quickly are becoming the sites of AI-based phishing attacks because of the budget constraints. These implications signal that special-purpose solutions to AI strategies and threat models based on each sector profile of exposure and digital maturity are necessary.

8. Conclusion

8.1 Summary of Findings

This paper examined the complex nature of artificial intelligence in the context of cybersecurity, with the possibility of it being a strong protective shield as well as the controller of advanced cyberattacks. Mixed-method research was used to achieve the result (surveys, case studies, and theoretical analysis), and it was established that the spheres of application of AI are significantly broad, and 85 percent of the organizations currently incorporate AI-based tools to detect threats, monitor in real-time, and analyze anomalies.

8.2 Reinforcing AI's Dual Nature

The study proves the main hypothesis that AI used in twofold cybersecurity. On the one hand, it enhances digital defenses by making them automatized and scalable and using predictive analytics. It strengthens the skill of attackers who will be able to evade detection and create smart malware, as well as conduct socially engineered attacks. Such a paradoxical dynamic requires a change of understanding: AI is not to be considered only as a solution, to be thought of as an agent that solve cybersecurity issues but also raise them. The duality of AI emphasizes the significance of human and AI cooperation, moral regulation, and embracing effective risk management.

9. Recommendations

The paper proposes a multi-level response initiative to reduce the risks and optimize the benefits of artificial intelligence in cybersecurity, which would entail a multi-level approach with policy remaking, organizational preparedness, ethical protection, and trust of cooperation in funding. Policy At the policy level, governments must implement an in-depth AI governance regime, e.g., national cybersecurity-AI standards, multinational cross-threat intelligence paperwork, and adherence to international standards such as the EU AI Act or the NIST AI Risk Framework. To address the threat of intelligent attacks (deepfakes and adaptive malware, in particular), organizations need to implement detection and response tools relying on AI, audit their networks regularly, and train the team to increase AI literacy and become more resistant to intelligent attacks.

AI implementation will have to get the center stage via algorithm transparency, solutions to biases, and the creation of internal ethics policing boards and should not engage in ugly practices of AI deployment that would be dangerous legally and ethically. At the same time, to advance resiliency and innovation, it is necessary to undertake strategic investments in AI research, open-source defensive tools, and public-private partnerships.

10. Limitations

This paper offers essential information on the dual role of artificial intelligence in modern cybersecurity, certain limitations should be mentioned. To begin with, scope-wise, the study was mostly restricted to the organizational practices and perceptions on the organizational level and limited in terms of the sample size it represented .The methodology consisted of survey responses and case arguments, which, though adequate, complained of the effects of response bias, differences in interpretative rules, and access to the sensitive or classified information in the security departments. This interferes with the generalizability and extensiveness of the findings.

The high rate of technological change is an issue, as the tools of AI and the threats of cyberattack are developing at a fast pace, and indeed some developments might have occurred or changed since the data were collected, and this cripple the appeal of the study in the long term. These shortcomings provide the basis of ongoing studies, real-time research, and more malleable approaches to keep up with the changes and variability of the AI-cybersecurity relationship.

11. Future Research Directions

The ever-changing nature of cybersecurity and the development of artificial intelligence, quite a few avenues of further research have been opening up. The area of AI and quantum cybersecurity has to be explored as a matter of urgency, especially in a stature that concerns the extent to which quantum computing deliver the attack vectors of traditional cryptographic systems and, at the same time, successfully improve AI-based threat modeling.

AI does not only recognize and categorize the threat but also acts largely on its own to execute counterattacks without major intervention from the human operator, which begs critical questions about the in-question reliability, trust and accountability. AI gets more integrated with security infrastructure. The case for effective AI regulations and governance becomes more urgent. Future directions to be pursued would include the development of frameworks of ethical deployment and legal liability as well as the alignment of international cybersecurity standards.

AI with blockchain technology presents a new direction for the future of decentralized, non-tamper able security models to increase the transparency, verification and authentication of the digital systems in real time. These guidelines are some of the crucial boundaries where interdisciplinary research, technology, and policy-making have to intersect in order to provide secure and ethical implementation of AI in the next years.

References

ıy, F. (2021). Emerging threats: The latest cybersecurity risks and the role of artificial intelligence in enhancing cybersecurity defenses. *Valley International Journal Digital Library*, *1*, 564-74.

ı, O. U., Abdullah, S. M., Olajide, A. O., Sani, A. I., Faisal, S. M. W., Ogunola, A. A., & Lee, M. D.

(2024). The Future of Cybersecurity: Leveraging Artificial Intelligence to Combat Evolving Threats and Enhance Digital Defense Strategies. *Journal of Computational Analysis and Applications*, *33*(8).

Lysenko, S., Bobro, N., Korsunova, K., Vasylchyshyn, O., & Tatarchenko, Y. (2024). The role of artificial intelligence in cybersecurity: Automation of protection and detection of threats. *Economic Affairs*, 69, 43-51.

Rajendran, R. M., & Vyas, B. (2023). Cyber security threat and its prevention through artificial intelligence technology. *International Journal for Multidisciplinary Research*, *5*(6).

Bonfanti, M. E. (2022). Artificial intelligence and the offence-defence balance in cyber security. *Cyber Security: Socio-Technological Uncertainty and Political Fragmentation. London: Routledge*, 64-79.

George, A. S. (2024). Emerging Trends in AI-Driven Cybersecurity: An In-Depth Analysis. *Partners Universal Innovative Research Publication*, 2(4), 15-28.

Karpatou, P. A. (2025). The evolution of cybersecurity threats & the rise of artificial intelligence.

Khan, M. I., Arif, A., & Khan, A. R. A. (2024). The most recent advances and uses of AI in cybersecurity. *BULLET: Jurnal Multidisiplin Ilmu*, 3(4), 566-578.

Jimmy, F. (2021). Emerging threats: The latest cybersecurity risks and the role of artificial intelligence in enhancing cybersecurity defenses. *Valley International Journal Digital Library*, *1*, 564-74.

n, M. (2024). Ai and cybersecurity: an ever-evolving landscape. *International Journal of Advanced Engineering Technologies and Innovations*, *1*.

andez, F. (2025). AI vs. AI: The Evolution of Offensive and Defensive AI Techniques in Cybersecurity. *Authorea Preprints*.

anaei, M., Khan, M. R., & Sylvester, N. N. (2024). Navigating AI cybersecurity: evolving landscape and challenges. *Journal of Intelligent Learning Systems and Applications*, *16*(3), 155-174.

, M. I., Arif, A., & Khan, A. R. A. (2024). The most recent advances and uses of AI in cybersecurity. *BULLET: Jurnal Multidisiplin Ilmu*, 3(4), 566-578.

A. (2024, November). The Impact of Artificial Intelligence on Cybersecurity. In *Abu Dhabi International Petroleum Exhibition and Conference* (p. D021S077R001). SPE.

ng, T. C., Zelinka, I., Plucar, J., Čandík, M., & Šulc, V. (2020). Artificial intelligence and cybersecurity: Past, presence, and future. In *Artificial intelligence and evolutionary computations in engineering systems* (pp. 351-363). Springer Singapore.

enko, A. (2024). *The role of artificial intelligence in countering cyber threats* (Doctoral dissertation, Сумський державний університет).

Hu, Y., Kuang, W., Qin, Z., Li, K., Zhang, J., Gao, Y., ... & Li, K. (2021). Artificial intelligence security: Threats and countermeasures. *ACM Computing Surveys (CSUR)*, *55*(1), 1-36.

Shah, W., & Best, T. (2025). Threat Detection and Risk Management: Using AI to Counter Cyber Threats.

Ortega, A. (2025). AI threats to national security can be countered through an incident regime. *arXiv preprint arXiv:2503.19887*.

Kaloudi, N., & Li, J. (2020). The ai-based cyber threat landscape: A survey. *ACM Computing Surveys (CSUR)*, *53*(1), 1-34.

Hrytsenko, A. (2024). *The role of artificial intelligence in countering cyber threats* (Doctoral dissertation, Сумський державний університет).

Kolluri, V. (2024). An Extensive Investigation Into Guardians Of The Digital Realm: Ai-Driven Antivirus And Cyber Threat Intelligence. *International Journal of Advanced Research and Interdisciplinary Scientific Endeavours*, *1*(2), 71-77.

.y, A. R. P. (2021). The role of artificial intelligence in proactive cyber threat detection in cloud environments. *NeuroQuantology*, *19*(12), 764-773.

.ón-Fonseca, R. (2022). The nature of security: A conceptual framework for integral-comprehensive modeling of IT security and cybersecurity. *Computers & Security*, *120*, 102805.

m, I., & Otoom, A. (2017). A classification scheme for cybersecurity models. *International Journal of Security and Its Application*, *11*(1), 109-120.

tji, M., Von Solms, S., & Marnewick, A. (2019). Socio-technical systems cybersecurity framework. *Information & Computer Security*, *27*(2), 233-272.

ɔd, P., & Hämäläinen, T. (2017, August). A novel model for cybersecurity economics and analysis. In *2017 IEEE International Conference on Computer and Information Technology (CIT)* (pp. 274-279). IEEE.