

Generative AI: Emerging Threats and Opportunities in Cyber Security

Amarnath Awasthi¹, Akshat Singh², Akshad³, Ansh Vimal⁴

¹ Assistant Professor, Axis Institute of Higher Education¹

^{2,3,4} Member, Axis Institute of Higher Education

Abstract—Generative AI is a significant part of the modern cybersecurity paradigm, but this does not always reflect well on the state of network security. Tools that have been fine-tuned using large-scale language models and conditional generative adversarial networks are proving to be a double-edged sword. This paper discusses both sides of their impact, as indicated by the quantitative analysis of GenAI's potential in network security research. On the plus side, the study affirms the efficacy of various security measures while postulating new vulnerabilities stemming from the technology's implementation.

Overall, the positive trends involve the application of transformer-based algorithms in network traffic analysis. Specifically, RoBERTa and Security BERT+, when combined with Fully Homomorphic Encryption, demonstrate accuracy in threat classification (0.91 precision, 0.89 recall) when operating on raw packet capture data. In turn, GAN-based data augmentation allowed addressing the issue of imbalanced datasets in network security research.

Meanwhile, the negative implications of using generative AI in cybersecurity were found in the field of attack obfuscation. The study revealed that CGANs could learn to generate background network traffic to be used as camouflage for attack traffic. This includes IoT botnet attacks (Mirai), SQL injection, and man-in-the-middle attacks, among others. This presents a major weakness in network intrusion detection systems, as shown by the results of the experiment. On three different CICIoT2023, LUFlow, and SQL NetFlow datasets, the addition of generative traffic decreased the performance of neural network intrusion detection systems to random chance.

Index Terms—Generative AI, Large language Models (LLMs), Conditional GANs (CGANs), Cybersecurity, Intrusion Detection System (IDS), Adversarial Machine Learning, Explainable AI(XAI), Federated Learning.

I. INTRODUCTION

The current state of the art of cybersecurity is undergoing a tremendous transformation under the influence of the newly emerged paradigm of Generative Artificial Intelligence and its utilization in various forms, including (but not limited to) Large Language, Generative Adversarial, and Conditional GANs. The crucial difference between traditional AI and GenAI is that the latter can be utilized not only to learn the patterns characteristic of a particular category of objects but also to generate new instances of data, engage in adversarial games, hypothesize about the data's meaning, and perform various other tasks previously exclusively performed by humans or conventional artificial intelligence.

Importantly, GenAI is a double-edged sword in cybersecurity, meaning that it can be utilized both to mount attacks and to defend against them. First, it is essential to underline that GenAI and its manifold applications in cybersecurity, including but not limited to threat detection and response, data synthesis, vulnerability analysis, phishing detection, and others, are ultimately beneficial for the defensive aspect of the cyber security landscape. GenAI can help red teams to attack systems more efficiently, which ultimately helps those systems to improve their defenses in an iterative fashion. Moreover, federated learning, encryption, and synthetic data technologies are generally more privacy-friendly, scalable, and secure alternatives to their traditional counterparts.

However, the same technological advancements that benefit cybersecurity defenders can be utilized by malicious actors to mount more sophisticated cyber-attacks. GenAI can be leveraged to generate

personalized phishing attacks, deepfake-based scams, more advanced malware capable of evading conventional protection software, utilizing myriad other AI-driven methods and technologies to overcome conventional cybersecurity measures, and much more. Finally, the utilization of GenAI adds new security risks and threats to the cybersecurity landscape, which, in turn, should be addressed by defenders. Among the most prominent threats posed by GenAI are prompt injection, data poisoning, adversarial examples, data leakage, and myriad other risks unique to the new paradigm of artificial intelligence.

Given the benefits provided by GenAI to both sides of the cybersecurity conflict, it is only reasonable to assume that the paradigm of cybersecurity should shift from traditional, perimeter-centric defense mechanisms to a new paradigm centered around self-sustaining structures capable of thriving in the dynamic, multi-faceted environment of ongoing operations. To that end, adversarial testing and training or XAI utilization to make human judgment a pivotal element of AI-driven cybersecurity solutions is a viable option. The implementation of model governance, privacy-centric AI, provenance technologies, and collaborative approaches will make it possible to accomplish that goal. Ultimately, this research will explore the ways in which GenAI can be utilized to mount attacks and defend against them. It will highlight the most significant threats presented by GenAI that can impact both sides of the cybersecurity conflict while also emphasizing the benefits unique to GenAI-driven cybersecurity solutions. Finally, the research will also outline the ways in which organizations can prepare themselves to resist AI-driven cyber-attacks while still being capable of mounting a formidable attack in the event of a conflict.

II. METHODOLOGY

The study reviews academic research papers, industry reports, cybersecurity advisories, and government publications related to Generative AI and cybersecurity. The literature mainly covers the period from 2019 to 2025. Sources were selected based on their relevance to GenAI-enabled cyber threats, defensive applications, real-world incidents, security frameworks, and emerging opportunities. Particular attention was given to Large Language Models (LLMs), AI-generated phishing, social engineering, malware assistance, deepfakes, prompt injection, data leakage, threat detection, vulnerability analysis, and automated incident response. The selected studies were compared to identify the major threats, opportunities, challenges, and defensive measures associated with GenAI.

III. GENERATIVE AI AND CYBERSECURITY

3.1 Generative AI and Overview of Large Language Models:

Generative Artificial Intelligence refers to a type of AI that creates new data, such as text, code, images, audio, and video. The development of the attention mechanism made it possible to implement generative AI successfully in practice using promising architectures, such as transformers. Modern Large Language Models (LLMs) are the product of developing the self-attention mechanism; therefore, such well-known models as GPT, Gemini, LLaMA, Claude, and others can perform various tasks, from answering questions to writing computer code given the input prompt. Since the release of GPT-3, there has been an exponential increase in the use of such models, which has led, on the one hand, to a significant acceleration of progress in many areas and, on the other hand, to critical security challenges. Generative AI can be used both to protect information and to attack information infrastructure, which emphasizes the importance of studying the defensive capabilities of GenAI and the threats it can pose.

3.2 Generative AI Techniques Used in Cybersecurity:

Table 1- GenAI Techniques used in Cybersecurity

Technique	How it works	Cybersecurity Application
Generative Adversarial Networks (GANs)	Two networks (generator and discriminator) compete, improving realism through adversarial training	Synthetic attack data generation, adversarial testing, deepfake/phishing detection
Large Language Models(LLMs)	Transformer-based models predict and generate text based on massive training corpora	Threat report summarization, phishing email detection, security chatbots, code review
Autoencoders	Neural networks compress and reconstruct input data, flagging poor reconstructions as anomalies	Network anomaly detection, unsupervised intrusion detection
Transformers	Attention-based architecture that models relationships across long sequences of data	Log analysis, malware sequence classification, natural language threat intelligence
Diffusion Model	Learn to reverse a gradual noising process to generate realistic synthetic samples	Synthetic dataset generation for training detection models, red-team simulation
Reinforcement Learning (RL) with Generative Agents	Agents learn optimal actions through reward-based trial and error, often paired with generative components	Automated penetration testing, adaptive defense strategy simulation

Evolution of AI based Cybersecurity:

The development and application of artificial intelligence in cybersecurity has gone through several stages. In the 1980s-1990s, expert systems were used, which were based on rulebased algorithms that could

identify attacks based on their characteristics. However, such systems proved to be somewhat limited since they could not react to intrusive activities that differed from known ones. In the 2000s,

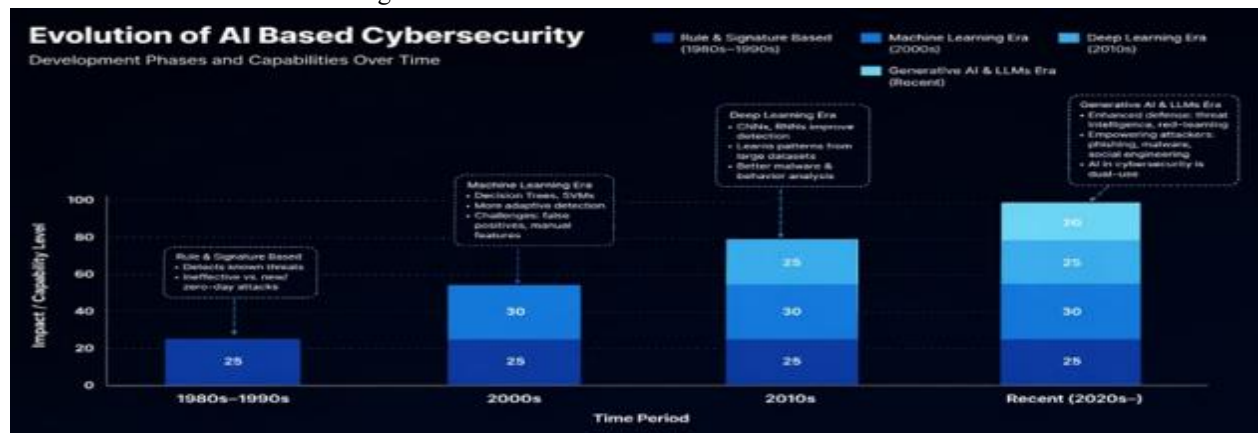


Figure 1-Evolu.on of AI Based Cybersecurity

machine learning algorithms, such as decision trees or support vector machines, were used to classify attacks based on their features extracted from data. However, the use of such algorithms was also limited due to the relatively small amount of data and difficulties in generalizing attacks. The next stage of development began in the 2010s and was marked by the advent of deep learning algorithms. The use of CNNs and RNNs made it possible to detect known types of attacks more reliably and determine anomalous data associated with attacks. Finally, the current stage of AI development in cybersecurity is associated with the use of Large Language Models and Generative AI, which made it possible to take cybersecurity to a new level. While traditional AI is limited to detecting attacks, generative AI can not only detect attacks but also automatically adjust the characteristics of existing ones or create completely new ones.

3.3 Role of Generative AI in Modern Security Systems

Generative AI plays a crucial role in modern security systems since it can be used both to defend information and to attack information infrastructure. On the one hand, GenAI can be used to automatically analyze vast amounts of threat data, extract useful information from it, and detect threats in real-time. For example, SOAR systems can use GenAI to prioritize and triage incoming alerts, analyze and manage incidents, and provide extensive situational awareness by generating situational awareness reports. Another example of the use of GenAI in cybersecurity is automatically analyzing code to find and fix weaknesses, which is especially relevant for critical software that requires high security. On the other hand, generative AI can be used to implement attacks by automatically generating new malware or modifying existing ones to evade detection. At the same time, various techniques from the Generative AI domain can be used to enhance existing security systems, such as GANs or VAEs, which can automatically test and adapt systems. Thus, generative AI can be used both to protect information and to attack information infrastructure, which requires examining the capabilities of such technologies in detail.

IV. EMERGING THREATS OF GENERATIVE AI IN CYBERSECURITY FIELD

4.1. AI Assisted Malware Development:

Generative AI threatens to significantly lower the difficulty and increase the scale and complexity of malware creation, distribution, and detection evasion. Large language and generative models could help would-be attackers rapidly produce polymorphic malware and other evolving, multi-hop attack chains that are challenging to detect and analyze using traditional security technologies. Defenses must include behavioral analysis, threat intelligence, and attack surface monitoring, as well as AI-specific detection and response techniques.

4.2. Data Poisoning and Adversarial Attacks:

Data poisoning and adversarial attacks pose significant security and safety risks by taking advantage of the fact that many Generative AI systems learn from data. By contaminating the data set or using carefully crafted inputs to trick the model into producing output that is not desired, attackers could undermine the accuracy and integrity of many cybersecurity functions. Adversarial examples may also cause AI-driven security systems to fail, allowing attacks to bypass detection. Defenses must be designed to detect and respond to poisoned data sets and include robust adversarial training and model validation, and monitoring.

4.3. Prompt Injection and Model Manipulation:

Prompt injection and model manipulation techniques allow attackers to circumvent safety and security measures in various ways, such as eliciting forbidden responses from a model or subverting its intended function. These attacks could compromise the confidentiality or integrity of data handled by AI systems or undermine their ability to perform critical security-related tasks. Defenses should seek to detect and prevent such attacks by implementing prompt validation, filtering, and monitoring, and other methods.

Table 2-Emerging Threats information

Threat	Description	Potential impact	Example
4.1 AI-assisted malware development	Generative models used to write, obfuscate, or mutate malicious code faster than traditional methods	LLM-assisted polymorphic malware that rewrites itself to evade signature detection	Faster malware creation, harder detection, lower skill barrier for attackers
4.2 Data poisoning and adversarial attacks	Attackers manipulate training data or inputs to corrupt model behavior or bypass detection	Injecting mislabeled samples into a malware classifier's training set to create blind spots	Degraded model accuracy, backdoored models, false negatives in detection systems
4.3 Prompt injection and model manipulation	Malicious inputs crafted to override an LLM's instructions or extract sensitive	A crafted prompt causing a security chatbot to leak internal configuration data	Data leakage, unauthorized actions, loss of trust in AI-driven security tools

V. OPPORTUNITIES IN GEN AI IN CYBERSECURITY

5.1 Threat and Anomaly Detection:

The utilization of generative AI, including such models as GAN, autoencoders, and LLMs, enable the description of the general behavior of systems to detect zero-day attacks, insider threats, and anomalous activities. Generative models can be applied to produce synthetic data, analyze unstructured security telemetry, and reduce alert fatigue by providing contextual incident summaries. Nonetheless, there are several challenges related to adversarial attacks, data poisoning, false positives, high computational costs, and limited model explainability.

5.2 Intrusion Detection Systems:

GANs, autoencoders, and hybrid IDS models based on LLMs can be utilized in intrusion detection systems for detecting zero-day intrusions, addressing an imbalanced dataset problem, and implementing unsupervised anomaly detection to separate normal and intrusive traffic. Additionally, hybrid intrusion detection and prevention systems based on LLMs can be applied for detecting complex multi-stage attacks by analyzing security logs with natural language understanding. The main challenges include intrusive traffic detection, adversarial attacks, high computational costs, and low interpretability.

5.3 Malware Detection and Analysis:

The application of generative models, including GANs and autoencoders, enables the implementation

of malware detection systems, which can identify previously unseen malicious software by learning the behavior of known malware. Moreover, generative models can be applied to analyze and reverse-engineer malware binaries to detect their unique patterns and behaviors. Finally, LLMs can be utilized for accelerating the process of code decomposition, analysis, and understanding of sandboxed malware. Nonetheless, there are several challenges, including adversarial malware creation, concept drift, high computational costs, and limited forensic transparency.

5.4 Threat Intelligence and Incident Response:

Generative AI can be applied to unify threat intelligence feeds, vulnerability databases, system logs, and other resources to provide fast and efficient detection of potential security incidents and support threat hunting activities. Additionally, LLMs can be utilized for creating structured summaries of identified security incidents and establishing efficient response playbooks. The use of LLMs for unifying heterogeneous information about cyber-threats can facilitate the daily operation of SOC teams by decreasing the mean time to detect (MTTD) and respond (MTTR) to security incidents. Nonetheless, the main challenges include hallucination, lack of technical domain knowledge, analyst reliance, and adversarial intelligence poisoning.

VI. COMPARING THE OPPORTUNITIES AND THREATS OF GENERATIVE AI IN CYBERSECURITY

6.1 Defensive and Offensive Use of Generative AI: The same advantages of generative AI that make it applicable in defensive cybersecurity can be weaponized by attackers. In terms of defensive tasks, generative AI was utilized for “accelerating” processes in Security Operation Centers by assisting in “incident classification” and “response recommendation,” creating “GANbased intrusion detection systems (e.g., TadGAN, TAnoGAN),” improving “password hardening (e.g., PassGAN)” and “generating synthetic data to address imbalances in training sets.” At the same time, the same capabilities can be used by malicious actors to generate phishing emails, malware, and deepfake videos. Microsoft’s report on cyber AI signals for 2025 notes a 46% increase in AI-generated phishing content and predicts that LLM-powered malware will comprise “50% of detected incidents in 2025 (up from 2% in 2021)” with nation-state actors using LLMs for reconnaissance and attack planning.

6.2 Benefits and Security Risks:

The benefits of using generative AI in cybersecurity are clear, as it “helps identify threats faster, automates vulnerability scanning, reduces the amount of work required to respond to incidents, and improves predictive security assessments.” However, the adoption of generative AI “comes at the expense of increased attack surfaces for organizations.” For example, “enterprise traffic scanning AI-powered content revealed a 90% traffic increase year-over-year (in one case) and required blocking a significant portion of traffic to AI platforms to reduce exposure.” The same report revealed that “testing adversarial inputs against different AI models uncovered weaknesses in all models tested, and red teamers could exploit these weaknesses to compromise a model-targeted system in under 16 minutes on average.” Another risk is the possibility of “lowering the barrier to entry for cybercriminals” as generative AI grants many advantages that “would have required far more resources to achieve without such technology.”

6.3 How Generative AI Is Changing Traditional Cybersecurity:

Traditional cybersecurity focuses on detecting and responding to incidents after they have occurred, with detection relying on pattern recognition and rule-based analysis of known threats. Using generative AI allows addressing threats that traditional methods would fail to detect, as well as automating responses before human intervention. The paradigm is “predictive” in that it attempts to preemptively detect and neutralize threats before they cause damage. However, such a paradigm is “not sustainable” due to the rapid evolution of malware, with the latter “already changing its code with each iteration to evade detection and attribution by traditional rule-based models.” Generative AI can also perform some tasks that rule-based systems could not do, including “semantic analysis of natural language text to detect sensitive information such as credentials or identities in a range of formats including text, images, video, audio, and log data.” Cybersecurity is shifting towards the use of generative AI because it is “enabling a fundamental change in how we anticipate, detect, respond to, and remediate cyberthreats.” However, it can also be used by attackers, making it a double-edged sword.

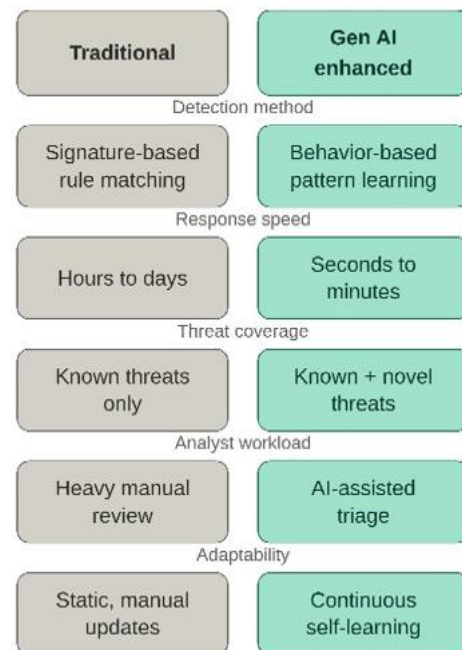


Figure 2 – Traditional vs GenAI Threat Detection

VII. CHALLENGES AND SECURITY REQUIREMENTS

7.1 Model reliability and robustness:

The most significant challenge posed by generative AI in cybersecurity is that these models can be manipulated. The two main ways through which this can occur are adversarial attacks, where the input data is modified slightly to mislead the model, and data poisoning, where the data used to train a model is tweaked to introduce undesirable traits in its functioning. Reports indicate that adversarial examples will constitute 30 percent of cyberattacks on AI by 2025. Thus it becomes imperative to develop robust models that can withstand such attacks. Researchers make adversarial attacks using Clever Hans and Fool box and use adversarial training to develop robust models. Generative adversarial networks have been successfully used to bolster the cybersecurity defenses. For example, in one study, GANs were used to reduce the success rate of evasion attacks by 15 percent and improve the accuracy of intrusion detection systems by 12 percent.

However, a more robust model tends to sacrifice accuracy for robustness, and there is a need to balance the two with regard to cybersecurity applications. Overall, there is a significant security challenge in terms of reliability and robustness.

7.2 Transparency, privacy, and Data Protection:

Another challenge in using generative AI in cybersecurity applications is that these systems can leak data. This risk is particularly acute in AI models that require massive amounts of data for training, fine-tuning, and deployment since these models are likely to incorporate sensitive data in their training samples. Moreover, if the training data is sourced from outside the country, there is an additional risk of cross-border data use and privacy violations. According to a report by Gartner, 40 percent of AI-related data breaches will be caused by the use of generative AI across borders by 2027. Thus, there is a need for increased transparency in AI operations as a means of protecting data.

In fact, transparency is a useful security mechanism because it allows organizations to ascertain whether adversarial attacks are being conducted by adversaries. In other words, AI should be able to explain what it does, and its inner workings should be visible in order

to ensure cybersecurity. For instance, the EU AI Act makes it mandatory to label AI-generated information and require certain measures from organizations developing high-risk AI models concerning robustness, accuracy, and cybersecurity. With regard to privacy and data protection, organizations are required to comply with the principles of privacy by design while developing AI systems. This includes using synthetic data to minimize the risks of exposing data subjects and employing cryptography to protect the integrity of data being processed by AI throughout the model's lifecycle. These considerations illustrate the extent to which transparency and data privacy need to be prioritized in the case of generative AI in cybersecurity applications.

7.3 Governance, ethics, and evaluation:

Governance frameworks and processes are increasingly viewed as critical pillars of AI deployment, particularly in relation to cybersecurity. In fact, without appropriate ethical decision-making and AI governance policies, organizations cannot effectively participate in the responsible and ethical use of AI, let alone realize value from their investments in the space. As such, there is a growing need for effective AI governance in accordance with emerging regulatory requirements. However, reports indicate that most organizations have yet to establish a robust AI governance framework; indeed, only 25 percent of organizations have established AI governance programs, and only 7 percent are at an advanced level, or AI future-ready, in terms of governance. Thus, the majority of organizations have a long way to go in terms of developing AI governance in keeping with the latest requirements.

Regulatory developments are playing a significant role in accelerating the development of AI governance frameworks fit for purpose. For example, the EU AI Act stipulates that general-purpose AI providers must meet certain system-level requirements, including adversarial testing, incident reporting, and cybersecurity requirements. At the same time, NIST's AI Risk Management Framework emphasizes four functions – govern, map, measure, and manage – that should guide AI governance processes. Overall, these developments signify the importance of governance, ethics, and evaluation in the context of generative AI in cybersecurity, as well as the associated costs. Organizations that fail to develop such frameworks

will be unable to meet the requirements and incur financial losses. After all, shadow AI costs organizations six hundred thousand dollars on average in case of a data breach; in addition, twenty percent of organizations have already experienced a data breach due to unauthorized AI usage.

VIII. FUTURE SCOPE

The future of generative AI in cybersecurity will be defined by an arms race between defenders and attackers. Gartner predicts that by 2026, 80% of enterprises will have embedded generative AI application programming interfaces (APIs) or models in their security operations, while attackers will leverage such models for fully automated end-to-end attacks with limited human intervention. The future scope of generative AI in cybersecurity will include researching more robust models with improved resistance to adversarial attacks, developing highly interpretable and trustworthy security solutions, and applying relevant homomorphic encryption and other methods to ensure confidence in the results and protect data. In addition, governance practices will need to evolve from adapting existing policies to addressing unique risks posed by AI systems. The future of generative AI in cybersecurity is dependent on the ability to balance the innovation and control to ensure that the development of such models does not outpace the efforts to secure them.

IX. CONCLUSION

The research in this paper was designed to demonstrate the double-edged nature of Generative AI in application-specific cybersecurity use cases, and indeed, the findings reported by the scholars have confirmed the initial assumptions. The first edge of the sword is represented by the ways in which the technology can be used to improve the performance of protective mechanisms within the targeted area of cybersecurity. The number of such methods is indeed considerable, ranging from the use of fine-tuned transformer-based neural networks for intrusion detection to the utilization of GANs for synthetic data generation, which allows for increased efficiency in malware analysis and improved incident response. Meanwhile, the second edge is represented by the ability of attackers to utilize similar methods and tools

to achieve the opposite goal of defeating cybersecurity measures and causing harm or disruption. In particular, the report reveals that the attack surface is now expanding due to the increased adoption of AI-driven tools among bad actors, including polymorphic malware and phishing campaigns utilizing computer vision, CGAN-driven obfuscation methods, and other emerging threats. Overall, the findings highlight the need to acknowledge the dual impact of Generative AI on organizational security as well as the associated risks and opportunities.

While the paper attempts to identify the ways in which Generative AI tools and algorithms can benefit various domains, including cybersecurity, it also recognizes various challenges and limitations associated with the use of such technologies. For starters, the report highlights the risk of model reliability issues, privacy concerns, and the complexity of establishing appropriate regulation and governance mechanisms in global and multijurisdictional contexts. In addition, the findings emphasize the financial and time costs of implementing new systems, adopting AI-specific practices, and addressing AI-related incidents, including the breach of privacy and data protection regulations.

Ultimately, the findings presented in the paper confirm the authors' thesis statement regarding the need to recognize the double-edged nature of Generative AI. Specifically, the report reveals that the benefits of utilizing such technology in cybersecurity heavily depend on the extent to which it is regulated and designed to operate in a secure and trustworthy environment. Moving forward, it will be interesting to observe how different industries embrace this innovation while balancing opportunities and challenges on both sides. In particular, more attention should be paid to ensuring security, transparency, and proper governance throughout the lifecycle of AI products and services.

REFERENCES

- [1] Ahuja, R. (2025). The risks and opportunities of generative AI in cyber security: Investigating the dual-use nature of generative models for phishing, malware development, and defence through synthetic data and automated threat detection. *International Journal of Innovative Research in*

- Computer and Communication Engineering, 13(2), 1758–1767.
- [2] Alghamdi, N. S., Alqahtani, H. J., & Alghamdi, A. (2024). Phishing attacks in the age of generative artificial intelligence: A systematic review of human factors. *Future Internet*, 16(8), 174.
 - [3] European Union. (n.d.). Accuracy, robustness and cybersecurity. EU Artificial Intelligence Act.
 - [4] Microsoft Corporation. (2025). Cyber Signals: AI-powered deception — emerging fraud threats and countermeasures / Microsoft Digital Defense Report 2025.
 - [5] IBM Security, & Ponemon Institute. (2025). Cost of a data breach report 2025.
 - [6] National Institute of Standards and Technology. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0) (NIST AI 100-1).
 - [7] Hitaj, B., Gasti, P., Ateniese, G., & Perez-Cruz, F. (2019). PassGAN: A deep learning approach for password guessing. In *Applied Cryptography and Network Security (ACNS 2019)*.
 - [8] Geiger, A., Liu, D., Alnegheimish, S., Cuesta-Infante, A., & Veeramachaneni, K. (2020). TadGAN: Time series anomaly detection using generative adversarial networks. *arXiv*.
 - [9] Bashar, M. A., & Nayak, R. (2020). TAnoGAN: Time series anomaly detection with generative adversarial networks. *arXiv*.
 - [10] Neto, E. C. P., Dadkhah, S., Ferreira, R., Zohourian, A., Lu, R., & Ghorbani, A. A. (2023). CICIOT2023: A real-time dataset and benchmark for large-scale attacks in IoT environment. *Sensors*, 23(13), 5941.
 - [11] OWASP Foundation. (2025). OWASP Top 10 for Large Language Model Applications.
 - [12] Microsoft Corporation. (2025). Microsoft Digital Defense Report 2025.
 - [13] IBM Security, & Ponemon Institute. (2025). Cost of a data breach report 2025.
 - [14] National Institute of Standards and Technology. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0) (NIST AI 100-1).
 - [15] Liu, J., Tayeen, A. S. M., Kumar, P., Gong, Q., Jiang, W., Cao, H., Misra, S., & Harikumar, J. (2026). Generative AI and federated learning for intrusion detection systems: A survey. *arXiv*.
 - [16] Rahman, S. A., Tout, H., Talhi, C., & Mourad, A. (2024). Survey on federated learning for intrusion detection system: Concept, architectures, aggregation strategies, challenges, and future directions. *ACM Computing Surveys*.
 - [17] Rjoub, G., Bentahar, J., Abdel Wahab, O., Mizouni, R., Song, A., Cohen, R., Otrouk, H., & Mourad, A. (2023). A survey on explainable artificial intelligence for network cybersecurity. *arXiv*.
 - [18] Nadeem, A., Vos, D., Cao, C., Pajola, L., Dieck, S., Baumgartner, R., & Verwer, S. (2023). SoK: Explainable machine learning for computer security applications. *IEEE European Symposium on Security and Privacy (EuroS&P)*.
 - [19] Alauthman, M., Aslam, N., Al-Qerem, A., Aldweesh, A., & Sureephong, P. (2026). Generative adversarial networks for intrusion detection systems: A comprehensive survey of applications, challenges, and research directions. *Arabian Journal for Science and Engineering*.
 - [20] Tolosana, R., Vera-Rodriguez, R., Fierrez, J., Morales, A., & Ortega-Garcia, J. (2020). Deepfakes and beyond: A survey of face manipulation and fake detection. *Information Fusion*, 64, 131–148.
 - [21] Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71.