The Uses of Generative Artificial Intelligence for Cybersecurity in Organisations

Chipo Chidakwa[1] and Zainab Ruhwanya[2][0000−0003−2339−7154]

[1,2] Cybersecurity and Privacy (CSPR) Research Group, Department of Information Systems, University of Cape Town, Cape Town, South Africa
[1]CHDCHI007@myuct.ac.za, [2]zainab.ruhwanya@uct.ac.za
www.cspr.uct.ac.za

Abstract. Generative Artificial Intelligence (Gen AI) is an emerging technology that has the potential to influence cybersecurity within organisations. The transformative capabilities of Gen AI enable it to adapt to the rapidly evolving cyberspace in which organisations operate. This study represents primary efforts to understand Gen AI's capabilities for cybersecurity and its uses within organisations. The study was conducted through a literature review to investigate the discourse on the use of Gen AI for cybersecurity, based on published academic papers. The findings point to a discourse focused on the capabilities of Gen AI for cybersecurity within organisations, such as threat detection and automated testing, among others. The study also highlights the benefits associated with the usage of Gen AI for cybersecurity, as well as the concerns that arise within organisations. Further research will be necessary to determine employee perceptions of the use of Gen AI for cybersecurity within organisations.

Keywords: Generative Artificial Intelligence, Artificial Intelligence, Cybersecurity, Organisations.

1 Introduction

Organisations are increasingly integrating Generative Artificial Intelligence (Gen AI) into their cybersecurity workflows as the threat landscape and defensive tools evolve in tandem with the rapid development of Artificial Intelligence (AI). Deshpande & Gupta [1] note that there is a dynamic link between Gen AI and cybersecurity due to the protective and flexible capabilities of Gen AI that allow it to evolve and adjust according to the changing landscape of cyber threats. While it is widely acknowledged that eliminating all cyber threats is not feasible, Gen AI has exhibited capabilities that can augment detection, response, and user awareness, effectively minimising their subsequent risks [2, 3]. GenAI possesses the capability to efficiently process heterogeneous data and generate multimodal outputs [4]. When utilised in security contexts, these functionalities facilitate the translation of complex threat intelligence into stakeholder-specific summaries, enable realistic phishing and social engineering simulations for training purposes, and assist in identifying, explaining, or flagging manipulated media

and other malicious content across multiple languages and channels. Consequently, these applications serve to enhance an organisation's security posture.

However, Gen AI has exposed a new realm of risks within the cybersecurity landscape [5]. These risks include the creation of highly convincing phishing at scale, deepfakes, automated social engineering, tool-assisted intrusion, and model-driven errors, such as hallucinations. [5–7]. The cyber risks that arise from the use of Gen AI for cybersecurity have an impact on an organisation's operations. Subsequently, efforts have been made to address these cyber risks, with some organisations implementing mitigation strategies [7]; however, the practice remains uneven.

Existing research has outlined the increased move towards incorporating Gen AI within organisational cybersecurity practices to leverage the opportunities that it presents. The implications of Gen AI usage for cybersecurity within organisations have exposed a gap in the understanding of the current uses of Gen AI for cybersecurity. This research aims to build on the existing research in this field and investigate the current uses of Gen AI for cybersecurity within an organisational context.

2 Background

The increase of AI advancements in the business world has led to the adoption of Gen AI within organisations for cybersecurity. Within organisations, cybersecurity is centred on safeguarding the hardware and software within cyberspace. [8]. Cybersecurity is seen to encompass the provision of capable resources and infrastructure to protect cyberspace by defending against and preventing cyber-attacks. AI enables machines to perform social and cognitive tasks, facilitating communication with other entities and processing high-level information [9]. Due to the advanced capabilities of AI, many organisations have begun incorporating emerging AI tools, such as Gen AI, into their organisational practices. Gen AI is an emerging technology capable of generating new, original content based on information provided to it and made available in various databases [10]. For example, the release of Gen AI tools such as ChatGPT and DALL–E by OpenAI has sparked growing interest in Gen AI within the AI community [11]. Interest has been drawn to the potential uses of Gen AI within organisations for cybersecurity and the subsequent implications.

The rapid rise of Gen AI has presented potential opportunities and threats in the cybersecurity landscape for organisations [5]. Existing research has highlighted a limited focus on research surrounding the uses and capabilities of Gen AI for cybersecurity within organisations, to leverage the potential opportunities and mitigate the possible threats that it presents. This gap implies the need to understand the uses of Gen AI for cybersecurity within organisations. The following research question has been derived to address this research topic and the main objective of this research:

Research Question: What are the uses of Gen AI for cybersecurity in organisations? With the primary objective of this research being to explore the applications of Gen AI for cybersecurity in organisations, this study assessed papers found through a literature search that investigate the intersection of Gen AI and cybersecurity within organisations. The capabilities of Gen AI for cybersecurity, its benefits, and drawbacks for

usage within organisations were investigated. The section below outlines the research methodology used to arrive at the findings.

3 Research Design

This research was conducted as a systematic literature review (SLR). The research protocol was guided by Okoli [12], with the process beginning by deriving keywords to be used for searching relevant literature in academic databases and other online platforms. Inclusion and exclusion criteria were established and used to screen the identified papers and exclude irrelevant ones. The remaining papers were then analysed, and key themes were identified and categorised for the findings.

3.1 Literature Search and Selection Strategy

As the focus of the study is Gen AI usage for cybersecurity in organisations, the key search terms used to source the relevant publications were: ("Generative Artificial Intelligence" OR "Gen AI") AND ("CYBERSECURITY") AND ("ORGANISATIONS.") The selected search terms were inputted into the Scopus and Web of Science databases as the two databases typically cover a wide range of publications relevant to any information systems or technology related discipline [12]. Additional publications were identified through a comprehensive literature search and snowball sampling. This "hybrid" technique was utilised for this research as the area of focus is new and continuously being updated, with added information being uncovered and made available at a rapid rate. Wohlin et al. [13] highlight that this technique is common in SLRs to improve the quantity and enhance the quality of the publications that are analysed. The publications found were then scrutinised against the inclusion and exclusion criteria outlined in Table 1.

Table 1. Inclusion and Exclusion Criteria

Criteria	Include	Exclude
Year	Published between 2021 and 2025. Due to the rapidly evolving nature of Gen AI, it is important for recent sources to be screened and analysed	Research published before 2021
Paper Focus	Exploring the uses of Gen AI for Cybersecurity in organisations	Focus is solely on the technical elements of Cybersecurity
Language	Written in English	Not written in English
Area of Focus	• Capabilities of Gen AI • Benefits of the use of Gen AI in an organisational context • Concerns surrounding the uses of Gen AI in an organisational context • Cybersecurity practices within an organisation that are impacted by AI	• No focus on the capabilities of Gen AI in the context of organisational cybersecurity • Focus on the technical elements of Gen AI and the different Gen AI models

3.2 Screening Process

This study followed the screening process outlined by Okoli [12] as a guide for selecting the appropriate papers. The study employed a two-phase screening approach, which began with searching for literature relevant to the topic using the defined search terms for this study, and then assessed the titles, abstracts, and keywords from the identified studies. Following this, the next phase screened the papers that had been narrowed down in Phase 1 and examined the entire text of the literature to identify studies that contained relevant information to the study focus. By following this screening process, the study gathered a relevant dataset for analysis.

A Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart diagram is used to outline the different phases of the systematic review of the literature [14]. Figure 1 outlines the PRISMA flowchart diagram, illustrating the data collection and screening process undertaken for this literature search.

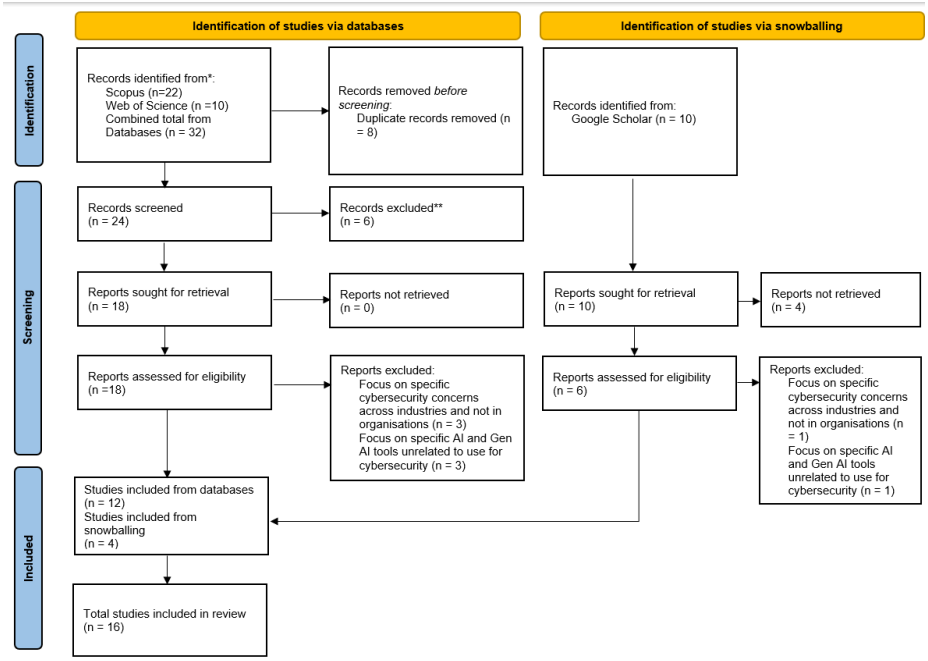

Fig. 1. PRISMA Diagram

The combined sourced papers were then analysed to derive the relevant insights for this research. These papers are listed in table 1 below.

Table 2. Systematic Literature Review Papers

	Paper Title		Paper Title
[2]	Cyber Attack Prediction: From Traditional Machine Learning to Generative Artificial Intelligence	[15]	Role of Artificial Intelligence based Chat Generative Pre-trained

			Transformer (ChatGPT) in Cyber Security
[16]	Lateral Phishing With Large Language Models: A Large Organization Comparative Study	[3]	Examine the Role of Generative AI in Enhancing Threat Intelligence and Cyber Security Measures
[17]	The Impact of Generative AI and LLMs on the Cybersecurity Profession.	[18]	Generative AI for Cyber Security: Analyzing the Potential of ChatGPT, DALL-E, and Other Models for Enhancing the Security Space
[1]	GenAI in the Cyber Kill Chain: A Comprehensive Review of Risks, Threat Operative Strategies and Adaptive Defense Approaches.	[19]	Bridging knowledge gap: the contribution of employees' awareness of AI cyber risks comprehensive program to reducing emerging AI digital threats.
[20]	Security and Privacy Perspectives on Using ChatGPT at the Workplace: An Interview Study	[21]	AI for cyber-security risk: harnessing AI for automatic generation of company-specific cybersecurity risk profiles
[22]	Enhancing Cyber Security Enhancement Through Generative AI	[23]	What The Phish! Effects of AI on Phishing Attacks and Defense
[24]	Generative AI in Cybersecurity	[25]	The paradigm of hallucinations in AI-driven cybersecurity systems: Understanding taxonomy, classification outcomes, and mitigations
[26]	Cyber Security Issues and Challenges Related to Generative AI and ChatGPT	[27]	GAI-Driven Offensive Cybersecurity: Transforming Pentesting for Proactive Defence

3.3 Literature Analysis Procedure

To analyse the extracted data, a thematic analysis following the six-step framework outlined by Braun and Clarke [28] was employed to identify, examine, and document the uncovered themes and concepts. As outlined by Maguire and Delahunt [29], a thematic analysis is valuable for identifying and highlighting the key themes and their interrelationships in a meaningful way. The 16 papers that were included for analysis were input into NVIVO software for coding to conduct a comprehensive analysis of the data gathered [30, 31]. An initial set of 208 codes was generated from an overall analysis of the papers. From this, themes were developed based on the recurring patterns identified, and the relevance and accuracy of the defined themes were then assessed, resulting in 3 main themes defined from the papers for interpretation. The following themes were outlined for this study: Understanding Gen AI and cybersecurity, the benefits of Gen AI usage for cybersecurity, and the drawbacks related to Gen AI usage for cybersecurity.

4 Results and Discussions

The findings from this study provide an overview of the themes which reflect the intersection between Gen AI and cybersecurity within organisations. This study explored the capabilities and benefits of Gen AI for cybersecurity, as well as the associated concerns and challenges.

4.1 Understanding Gen AI for Cybersecurity

With the rapid rise of AI and its growing usage within organisations, it is important to understand the emerging technologies that fall under it and their growing usage within organisations. AI is defined by Capodieci et al. [17]p.448 as " the field of study that analyses the creation of intelligent machines" . While Prasad et al. [15]p.107 recognise AI as a "multidisciplinary technology" that can emulate human behaviours and tasks more efficiently, store data, make decisions, and promote machine learning. The advanced nature of AI as a technology has seen it being deployed in many contexts, from organisations to individuals' lives. AI offers several advantages for addressing security concerns, including detecting network vulnerabilities, analysing large datasets to identify suspicious activities, and automating elements of incident response to enable timely countermeasures [3]. The field has evolved over the years, transforming technological capabilities. Schreiber & Schreiber [19] highlight the shift from rule-based systems to the more recent subset, GenAI, which can produce original content and be utilised for security. AI technologies continue to advance, with GenAI and related tools developing rapidly [26].

GenAI models exhibit transformative capabilities that can impact an organisation's cybersecurity space. Large language models (LLMs) fall under GenAI and use deep learning and statistical methods to analyse vast amounts of data and predict patterns [2, 17]. These models can generate realistic content and assist with tasks that improve aspects of an organisation's cybersecurity profile [3, 16]. Their value has been observed in organisational settings, particularly for cybersecurity. In some cases, GenAI models are developed, fine-tuned, or privately deployed to maintain control over data handling and governance, while others rely on enterprise-grade, vendor-hosted options with contractual data protections [17, 20].

However, GenAI has also made it more difficult for organisations to detect malicious threats as attack techniques evolve. Adversaries can leverage GenAI to craft persuasive phishing and social-engineering campaigns, generate deepfakes, and automate parts of exploitation, which can outpace some legacy controls [23, 24]. Such capabilities risk eroding trust, misleading employees, and prompting actions that compromise privacy settings and safeguards [2, 20, 21]. The evolving nature of cyber threats underscores the need for enhanced, AI-aware defences that specifically address GenAI-enabled risks [19]. As such, GenAI presents a double-edged sword for organisational cybersecurity; its defensive benefits are substantial, but so too are the new attack surfaces and modes of deception it enables.

4.2 Benefits of Gen AI Usage for Cybersecurity

Gen AI can analyse data, identify threat patterns, and surface insights that help assess the effectiveness of an organisation's security posture. By using predictive capabilities on large datasets, Gen AI models can improve threat detection and inform security decisions [3, 24]. Organisations also utilise these models to generate realistic scenarios and synthetic data to simulate testing environments and evaluate the effectiveness of their threat detection and risk analysis approaches [3, 18]. Furthermore, adversarial testing is used to probe the vulnerabilities of Gen AI-enabled systems and to develop more robust models and controls [18]. AI-enabled security tooling can augment monitoring and adapt scoring or prioritisation as risks evolve in organisations. By using Gen AI to develop and enhance cybersecurity measures, organisations can strengthen protections while leveraging the opportunities these technologies present. Overall, AI-generated security systems can combine the benefits of Gen AI for cybersecurity by effectively identifying and adapting to evolving cyber risks in organisations [1, 21].

Threat intelligence and detection have improved in efficiency with the introduction of Gen AI for cybersecurity [21]. Gen AI models can protect against cyber-attacks by identifying potential threats and providing updated security insights that support risk mitigation [3, 22]. These systems can also automate elements of incident response, such as enrichment, triage, and playbook steps, typically with human-in-the-loop oversight, to contain breaches more quickly [1]. This increases the chances of minimising the impact of attacks and containing risk efficiently. Where appropriate, automation enables timely adjustments to countermeasures; however, full real-time autonomy is uncommon and is usually governed by policy [2].

Cybersecurity experts can leverage Gen AI models to analyse existing datasets and identify potential threats [3]. The same capabilities can support system response design and employee awareness training. For example, Gen AI can generate realistic training materials or simulated phishing content to increase vigilance and reduce exposure to sophisticated attacks, provided outputs are reviewed for accuracy and context [3, 16, 19]. Cybersecurity testing profiles are generated by these models to familiarize employees with common trends in cyber-attacks. Therefore, the usage of Gen AI to ensure that employees are aware of the evolving cyber threat landscape and to train their system responses enhances the cybersecurity of organisations.

Gen AI has been shown to provide beneficial enhancements to organisational cybersecurity across business domains[21]. Many organisations are implementing Gen AI-enabled tools within their security programmes. Used appropriately, Gen AI can enhance defensive measures by helping to detect and mitigate cyberattacks[2, 3, 22]. In summary, key benefits include improved threat detection, partial automation of incident response, analysis of large datasets to create testing scenarios, and timely adjustments to countermeasures.

4.3 Drawbacks of Gen AI Usage for Cybersecurity

Despite the transformative capabilities of Gen AI for cybersecurity within organisations, attackers are also leveraging these tools to lower barriers to entry and to scale and

customise cyber-attacks [24]. This dynamic contributes to greater sophistication in attack campaigns, even if overall incident rates are influenced by many factors. Organisations face a range of AI-related risks, including large-scale phishing and social engineering, the spread of misinformation, deepfakes, AI-assisted malware development and polymorphism, automated reconnaissance and exploitation, data manipulation and poisoning, prompt injection, and adversarial prompt attacks [17, 19, 21–23, 25, 26]. In parallel, Gen AI deployments introduce system-level risks such as hallucinations that can mislead analysts or users, inadvertent leakage of sensitive data through prompts or outputs, and integrity threats to models and data pipelines (e.g., poisoning or inversion) [17, 22].

Ethical Concerns and Limited Employee Awareness

Gen AI models are trained on large datasets and require this data to generate functional outputs. In practice, employees may inadvertently input sensitive information into Gen AI tools when seeking help to draft security strategies or analyse configurations, which can expose privacy risks if data-handling settings are not well controlled [26]. Many employees have limited experience with AI applications and are often unaware of these risks [19, 20]. As such, confidential information may be unintentionally shared via Gen AI tools. Where retention or model training on user inputs is permitted, sensitive details could be revealed through outputs or logs, increasing the risk of disclosure to unauthorised recipients [17, 22]. Furthermore, some vendor-hosted Gen AI services may retain prompts and outputs by default (unless enterprise controls disable this), which, together with misconfigurations or third-party integrations, can create leakage pathways that compromise the privacy and integrity of organisational data.

Ethical concerns also arise around data accuracy and the provenance of information used by Gen AI systems. Models can inherit biases from their training data, leading to unintentional unfair outcomes [17, 26]. Limited interpretability makes such biases harder to detect and explain, which can undermine transparency and decision quality in cybersecurity use cases [25]. While Gen AI providers implement policies and safety filters, these safeguards can be bypassed through "jailbreaking" prompts, leaving organisations exposed to misuse if internal controls are weak [24]. Because Gen AI content is derived from heterogeneous datasets, comprehensive accuracy checks are challenging. Many organisations further lack clear ethical guidelines for acceptable use, data classification, and review procedures, increasing exposure to data risks [20, 27]. To mitigate these issues, organisations should aim for Gen AI-generated outputs that are reviewed for bias and transparency and governed by explicit ethical policies, covering retention settings, no-training modes, access controls, and human-in-the-loop review to improve the protection and security of sensitive data [21].

Malicious Actors and Fraudsters

Fraudsters and threat actors are leveraging Gen AI to avoid detection and to launch sophisticated, targeted cyber-attacks more efficiently [1]. Malicious actors have identified vulnerabilities in AI-enabled systems that can be exploited through automation of reconnaissance and exploitation steps and through adversarial techniques (e.g., prompt

injection and evasion) [2, 22]. In this way, Gen AI can be misused to generate convincing lures, tailor payloads, and scale attacks, creating defensive challenges for organisations that adopt these tools [21, 24]. The overall impact varies across sectors and environments, but there is clear evidence of increasing sophistication and scale in some attack campaigns that make use of Gen AI [17]. As such, companies that employ Gen AI for cybersecurity are exposed to malicious actors who may use the same technologies to seek unauthorised access to confidential information and to execute cyber-attacks, underscoring the need for strong governance, monitoring, and guardrails.

Deepfakes and Phishing Attacks

Deepfake technology enables malicious actors to generate convincing synthetic content to deceive individuals and spread [17, 19, 26]. Using deepfakes, attackers can misrepresent themselves as trusted employees or executives to induce fraudulent actions under the appearance of legitimate requests [2, 16]. Employees may become susceptible to such attacks because of the increasing realism of deepfakes, which makes it difficult to distinguish authentic from fabricated content and can accelerate the spread of misinformation [19, 24]. Deepfakes can therefore misdirect staff through identity impersonation and fraud, exposing organisations to privacy and integrity risks and compromising security processes [18]. The impact of deepfake technology on organisational cybersecurity highlights the potential limitations of Gen AI for cybersecurity when governance and verification controls are weak.

Phishing attacks can be socially engineered by cyber attackers using Gen AI models to deceptively obtain access to private and sensitive data from employees [2, 17, 26]. Attackers can enhance their tactics with AI to improve targeting, tone, and timing, posing material risks to an organisation's cybersecurity posture [22]. Such misuse allows adversaries to craft sophisticated messages that appear legitimate and manipulate recipients into disclosing sensitive information or executing risky actions [24]. Detecting AI-generated phishing emails is challenging for employees, especially without proper training and robust layered controls, because the messages are often persuasive and consistent with internal styles [23].

4.4 Conceptual Framework

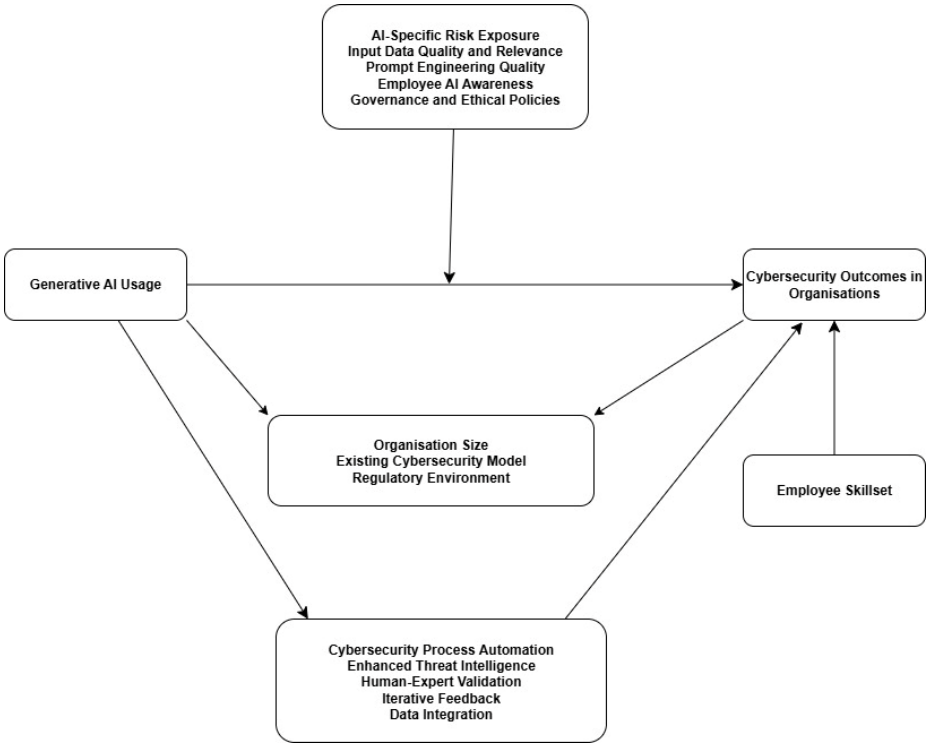

Fig. 2. Conceptual framework linking GenAI usage to organisational cybersecurity outcomes

The figure shows how Gen AI usage relates to cybersecurity outcomes in organisations. Outcomes refer to the ability to detect, respond to, and prevent threats. Gen AI can influence outcomes directly. It can also do so by improving security processes such as process automation, enhanced threat intelligence, human-expert validation, iterative feedback, and data integration. The strength of this influence depends on several factors: AI-specific risk exposure, input data quality and relevance, prompt-engineering quality, employee AI awareness, and governance and ethical policies. Contextual factors can shape both Gen AI adoption and outcomes. These include organisation size, the current cybersecurity model, and the regulatory environment. Employee skillset also affects outcomes and is shown separately in the diagram

5 Conclusion and Future Research Directions

This study provided an understanding of Gen AI and its capabilities for cybersecurity within organisations. Gen AI can analyse large datasets, identify threat patterns, and surface insights that inform security decisions. Additionally, Gen AI models can adapt to the evolving threat landscape and automate incident responses, typically with human

oversight, to ensure an effective response to attacks and minimise the impact of cyberattacks. These models also provide valuable insights for enhancing cybersecurity systems and simulate cyberattacks to raise awareness of potential threats across the workforce.

The study also found important drawbacks. Attackers can use Gen AI to run deceptive campaigns, for example deepfakes and highly tailored phishing, that exploit employee vulnerabilities. A limited understanding of Gen AI among staff can lead to the accidental sharing of sensitive information with large language models, thereby weakening data security and privacy. Because Gen AI learns from varied datasets, it may reproduce biases and create integrity risks if outputs are not reviewed. Gaps in ethical guidance and data-handling rules further reduce organisational control over what is entered into tools and over the accuracy of the outputs. Organisations need to recognise these concerns and put mitigation strategies in place so that benefits can be realised responsibly.

Given this double-edged nature, many organisations struggle to keep pace with rapid Gen AI advances, which can leave them exposed to evolving threats. The practical task is to strengthen governance, configure tools correctly, maintain human review, and build targeted awareness so that Gen AI supports defence rather than creating avoidable risk.

This paper outlined the benefits and drawbacks of Gen AI for organisational cybersecurity based on published academic literature. While prior work covers many practical uses, there is limited coverage of how employees perceive Gen AI in security work, even though those perceptions can create opportunities or additional risks. Future research should examine these perceptions to provide a more complete understanding of Gen AI use for cybersecurity in organisations. Two research questions follow: How do perceptions and uses of Gen AI differ between SMEs and large enterprises, and what does that mean for outcomes? Which security awareness and review approaches help staff detect AI-enabled phishing and deepfakes in practice?

References

1. Deshpande AS, Gupta S (2023) GenAI in the Cyber Kill Chain: A Comprehensive Review of Risks, Threat Operative Strategies and Adaptive Defense Approaches. In: 2023 IEEE International Conference on ICT in Business Industry & Government (ICTBIG). IEEE, Indore, India, pp 1–5
2. Ankalaki S, Atmakuri AR, Pallavi M, Hukkeri GS, Jan T, Naik GR (2025) Cyber Attack Prediction: From Traditional Machine Learning to Generative Artificial Intelligence. IEEE Access 13:44662–44706. https://doi.org/10.1109/ACCESS.2025.3547433
3. Saddi VR, Gopal SK, Mohammed AS, Dhanasekaran S, Naruka MS (2024) Examine the Role of Generative AI in Enhancing Threat Intelligence and Cyber Security Measures. In: 2024 2nd International Conference on Disruptive Technologies (ICDT). IEEE, Greater Noida, India, pp 537–542
4. Prasad Agrawal K (2024) Towards Adoption of Generative AI in Organizational Settings. Journal of Computer Information Systems 64:636–651. https://doi.org/10.1080/08874417.2023.2240744
5. Neupane S, Fernandez IA, Mittal S, Rahimi S (2023) Impacts and Risk of Generative AI Technology on Cyber Defense

6. 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 11:80218–80245. https://doi.org/10.1109/ACCESS.2023.3300381
7. Singla A, Sukharevsky A, Yee L, Chui M, Hall B (2025) The State of how Organizations are Rewiring to Capture Value
8. Jada I, Mayayise TO (2024) The impact of artificial intelligence on organisational cyber security: An outcome of a systematic literature review. Data and Information Management 8:100063. https://doi.org/10.1016/j.dim.2023.100063
9. Abbass H (2021) Editorial: What is Artificial Intelligence? IEEE Trans Artif Intell 2:94–95. https://doi.org/10.1109/TAI.2021.3096243
10. Teo ZL, Quek CWN, Wong JLY, Ting DSW (2024) Cybersecurity in the generative artificial intelligence era. Asia-Pacific Journal of Ophthalmology 13:100091. https://doi.org/10.1016/j.apjo.2024.100091
11. Byrne MD (2023) Generative Artificial Intelligence and ChatGPT. Journal of PeriAnesthesia Nursing 38:519–522. https://doi.org/10.1016/j.jopan.2023.04.001
12. Okoli C (2015) A Guide to Conducting a Standalone Systematic Literature Review. CAIS 37:. https://doi.org/10.17705/1CAIS.03743
13. Wohlin C, Kalinowski M, Romero Felizardo K, Mendes E (2022) Successful combination of database search and snowballing for identification of primary studies in systematic literature studies. Information and Software Technology 147:106908. https://doi.org/10.1016/j.infsof.2022.106908
14. PRISMA 2020 flow diagram. In: PRISMA statement. https://www.prisma-statement.org/prisma-2020-flow-diagram. Accessed 30 Sept 2025
15. Prasad SG, Sharmila VC, Badrinarayanan MK (2023) Role of Artificial Intelligence based Chat Generative Pre-trained Transformer (ChatGPT) in Cyber Security. In: 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC). IEEE, Salem, India, pp 107–114
16. Bethany M, Galiopoulos A, Bethany E, Bahrami Karkevandi M, Beebe N, Vishwamitra N, Najafirad P (2025) Lateral Phishing With Large Language Models: A Large Organization Comparative Study. IEEE Access 13:60684–60701. https://doi.org/10.1109/ACCESS.2025.3555500
17. Capodieci N, Sanchez-Adames C, Harris J, Tatar U (2024) The Impact of Generative AI and LLMs on the Cybersecurity Profession. In: 2024 Systems and Information Engineering Design Symposium (SIEDS). IEEE, Charlottesville, VA, USA, pp 448–453
18. Sai S, Yashvardhan U, Chamola V, Sikdar B (2024) Generative AI for Cyber Security: Analyzing the Potential of ChatGPT, DALL-E, and Other Models for Enhancing the Security Space. IEEE Access 12:53497–53516. https://doi.org/10.1109/ACCESS.2024.3385107
19. Schreiber A, Schreiber I (2024) Bridging knowledge gap: the contribution of employees' awareness of AI cyber risks comprehensive program to reducing emerging AI digital threats. ICS 32:613–635. https://doi.org/10.1108/ICS-10-2023-0199
20. Kimbel A, Glas M, Pernul G (2025) Security and Privacy Perspectives on Using ChatGPT at the Workplace: An Interview Study. In: Clarke N, Furnell S (eds) Human Aspects of Information Security and Assurance. Springer Nature Switzerland, Cham, pp 184–197
21. Schreiber A, Schreiber I (2025) AI for cyber-security risk: harnessing AI for automatic generation of company-specific cybersecurity risk profiles. ICS. https://doi.org/10.1108/ICS-08-2024-0177
22. Krishnamurthy O (2023) Enhancing Cyber Security Enhancement Through Generative AI. Vol No

23. Shreyas Kumar, Anisha Menezes, Sarthak Giri, Srujan Kotikela (2024) What The Phish! Effects of AI on Phishing Attacks and Defense. ICAIR 4:218–226. https://doi.org/10.34190/icair.4.1.3224
24. Metta S, Chang I, Parker J, Roman MP, Ehuan AF (2024) Generative AI in Cybersecurity
25. Sood AK, Zeadally S, Hong E (2025) The paradigm of hallucinations in AI-driven cybersecurity systems: Understanding taxonomy, classification outcomes, and mitigations. Computers and Electrical Engineering 124:110307. https://doi.org/10.1016/j.compeleceng.2025.110307
26. Pasupuleti R, Vadapalli R, Mader C (2023) Cyber Security Issues and Challenges Related to Generative AI and ChatGPT. In: 2023 Tenth International Conference on Social Networks Analysis, Management and Security (SNAMS). IEEE, Abu Dhabi, United Arab Emirates, pp 1–5
27. Zaydi M, Maleh Y (2025) GAI-Driven Offensive Cybersecurity: Transforming Pentesting for Proactive Defence: In: Proceedings of the 11th International Conference on Information Systems Security and Privacy. SCITEPRESS - Science and Technology Publications, Porto, Portugal, pp 426–433
28. Braun V, Clarke V (2006) Using thematic analysis in psychology. Qualitative Research in Psychology 3:77–101. https://doi.org/10.1191/1478088706qp063oa
29. Maguire M, Delahunt B Doing a Thematic Analysis: A Practical, Step-by-Step Guide for Learning and Teaching Scholars.
30. Bhattacherjee A (2012) Social Science Research: Principles, Methods, and Practices. Global Text Project, Place of publication not identified
31. Leech NL, Onwuegbuzie AJ (2011) Beyond constant comparison qualitative data analysis: Using NVivo. School Psychology Quarterly 26:70–84. https://doi.org/10.1037/a0022711