Original Scientific Paper/Original naučni rad UDC/UDK: 004.8:004.738.5.056
Paper Submitted/Rad primljen:31.12.2025.
Paper Accepted/Rad prihvaćen:20.01.2026.
DOI: 10.5937/SJEM2600020N

Fišing napadi zasnovani na veštačkoj inteligenciji: Novi izazovi i bezbednosne strategije

Toni Nakovski[1], Natasha Blazheska-Tabakovska[2], Mimoza Bogdanoska Jovanovska[3]

[1] University "St. Kliment Ohridski", Bitola, Republic of North Macedonia, toni.nakovski@gmail.com
[2] University "St. Kliment Ohridski", Bitola, Republic of North Macedonia, natasa.tabakovska@uklo.edu.mk
[3] University "St. Kliment Ohridski", Bitola, Republic of North Macedonia, mimoza.jovanovska@uklo.edu.mk

Summary in Serbian: Brzi napredak generativne veštačke inteligencije doveo je do nove generacije izuzetno sofisticiranih fišing napada. Rad razmatra evoluirajući pejzaž pretnji u kojem zlonamerni akteri koriste veštačku inteligenciju za oblikovanje obimnih, personalizovanih i veoma uverljivih fišing kampanja. Analiza je usmerena na nove tehnike napada, uključujući automatizovano generisanje sadržaja, i njihove šire implikacije. Nalazi ukazuju na to da su tradicionalni mehanizmi odbrane sve manje efikasni protiv AI-vođenih napadačkih taktika. Kao odgovor, rad ističe potrebu za slojevitom strategijom sajber odbrane koja integriše bezbednosna rešenja zasnovana na veštačkoj inteligenciji sa kontinuiranom edukacijom i podizanjem svesti korisnika. Takođe se razmatraju etičke, društvene i regulatorne dimenzije zloupotrebe veštačke inteligencije, naglašavajući značaj globalne saradnje između industrije, akademske zajednice i donosioca odluka. Pružanjem sveobuhvatnog pregleda trenutnih izazova i budućih rizika, rad doprinosi jasnijem razumevanju fišinga unapređenog veštačkom inteligencijom i predlaže proaktivan, perspektivan okvir usmeren ka jačanju otpornosti sajber bezbednosti u 21. veku.

Keywords: Veštačka inteligencija, Fišing napadi unapređeni veštačkom inteligencijom, Tehnike fišing napada, Bezbednosna rešenja zasnovana na veštačkoj inteligenciji

AI-Driven Phishing Attacks: Emerging Threats and Security Strategies

Abstract in English: The rapid advancement of generative artificial intelligence has ushered in a new wave of highly sophisticated phishing attacks. This paper examines the evolving threat landscape in which malicious actors leverage AI to craft large-scale, personalised, and highly convincing phishing campaigns. The analysis focuses on emerging attack techniques, including automated content generation, and their broader implications. Findings demonstrate that conventional defence mechanisms are increasingly inadequate against AI-driven adversarial tactics. In response, the paper underscores the need for a multi-layered cyber defence strategy that integrates AI-powered security solutions with continuous user education and awareness initiatives. Furthermore, the ethical, societal, and regulatory dimensions of malicious AI deployment are explored, emphasising the importance of global cooperation among industry, academia, and policymakers. By presenting a comprehensive overview of current challenges and future risks, the paper contributes to a deeper understanding of AI-augmented phishing and proposes a proactive, forward-looking framework aimed at strengthening cybersecurity resilience in the 21st century.

Keywords: Artificial Intelligence, AI-augmented phishing**,** Attack Techniques, AI-powered security solutions

1. Introduction

The rapid advancement of generative artificial intelligence (AI) has fundamentally transformed the cybersecurity landscape, giving rise to an era of highly sophisticated and scalable phishing attacks that challenge traditional defence mechanisms. As AI tools become increasingly accessible, malicious actors exploit these technologies to automate and enhance the creation of personalised and convincing phishing campaigns capable of deceiving even trained security professionals. A central concern is that attackers can now generate deceptive content, including text, audio, and video, at an unprecedented scale and level of realism, making it increasingly difficult to distinguish legitimate communications from fraudulent ones. Audio creation techniques, such as voice cloning, allow

attackers to mimic an individual's voice convincingly. Live filters can falsify an attacker's voice during calls, further enhancing the illusion of authenticity. Deepfake videos enable attackers to impersonate an employee's face during video calls, creating highly realistic scenarios for targeted social engineering attacks. Collectively, these AI-driven capabilities significantly increase the effectiveness of phishing campaigns and present new challenges for traditional security measures. This paper aims to provide an analysis of these emerging threats and to examine strategies that address the risks associated with AI-augmented phishing.

This study addresses several critical questions that are reshaping the field of cybersecurity: What are the primary techniques employed in AI-driven phishing attacks? How do these techniques differ from traditional phishing methods? And what are the broader ethical and societal implications of malicious AI deployment in this context? This paper will explore a range of AI-driven attack techniques, from the use of Large Language Models (LLMs) for automated content generation to the deployment of deepfakes and voice cloning for highly targeted social engineering. It will be argued that the increasing sophistication of these attacks renders traditional, signature-based security measures insufficient. Consequently, this paper will advocate for a holistic security paradigm that integrates AI-powered detection and response systems with robust user awareness and training programs. Furthermore, the ethical and social dimensions of malicious AI will be examined, with a call for a collaborative, multi-stakeholder approach to the development of effective countermeasures and regulatory frameworks. This paper will conclude with a summary of key findings, a discussion of the limitations of the current research, and recommendations for future research directions.
The paper is structured as follows: Following the **introduction,** the related work section provides a comprehensive overview of AI-driven phishing techniques and defense mechanisms. Section 3 outlines the methodological framework applied in the study. Section 4 presents and discusses the results, comparing them with the existing body of research (Eze, Shamir, 2024; Schmitt, Flechais, 2024; Gallagher, 2024*)*. Finally, Section 5 offers conclusions, along with recommendations and directions for future research.

2. Related Work – Literature Review (Theoretical Framework)

Traditional phishing attacks have historically relied on mass-produced, generic messages designed to deceive users into disclosing sensitive information (Basit, Zafar, Liu, & al., 2021). Such messages frequently exhibit identifiable flaws, including grammatical errors or suspicious links, which often allow rule-based filters and attentive users to detect and mitigate the attacks. In contrast, the integration of artificial intelligence (AI), particularly Generative AI (GenAI) and Large Language Models (LLMs), has marked a significant turning point in the phishing landscape (Jabir, Le, & Nguyen, 2025). AI-driven phishing now leverages sophisticated techniques to create highly personalised, contextually relevant, and grammatically flawless messages at an unprecedented scale, evading traditional detection methods (Khalil, 2025) and posing substantial challenges to conventional cybersecurity defences.

Modern AI-driven phishing attacks expand capabilities across multiple dimensions. Beyond crafting tailored text, attackers can generate persuasive voice and video content through techniques such as voice cloning, live voice-modification filters during calls, and deepfake video impersonation during video conferences. These advancements enhance the credibility of malicious communications and make them harder to detect. Furthermore, the speed and scale of execution have dramatically increased: whereas a human attacker may require more time to craft a single phishing message, LLMs can produce hundreds of slightly varied versions in the same timeframe. AI also enables timely and contextual targeting by incorporating real-time news, corporate developments, and personal information into phishing messages, making them highly believable and contextually relevant (Letain-Mathieu, 2025).

The scope of AI-powered phishing techniques is rapidly expanding, marking a significant escalation in the sophistication and impact of contemporary cyber threats. These techniques include:

- **Automated Content Generation:** LLMs generate high-quality, context-aware phishing emails, which are significantly more convincing than traditional messages (Chen, Wu, Nguyen, & Rudolph);
- **AI-Enhanced Spear Phishing:** AI algorithms gather and analyse large amounts of data on specific targets, enabling the creation of hyper-personalised messages that closely mimic legitimate communications (Arntz, 2025);
- **Polymorphic Attacks:** AI continuously alters email characteristics—such as sender information, subject lines, and URLs—to evade signature-based detection systems (IRONSCALES (n.d.), 2025);
- **Deepfakes and Voice Cloning:** Synthetic audio and video content enable advanced attacks, particularly in Business Email Compromise (BEC) and voice phishing (vishing) scenarios, adding a new dimension to phishing threats (Fitzgerald, 2025).

The emergence of AI-driven phishing has necessitated a parallel evolution in defensive strategies, as traditional static detection methods are increasingly insufficient to address the dynamic and adaptive characteristics of these threats (Basit, Zafar, Liu, & al., 2021). In response, there has been a pronounced shift towards AI-powered security solutions, which can be broadly categorised into several key approaches. AI-based email filters leverage advanced machine learning and deep learning techniques to analyse linguistic and contextual patterns in messages, enabling the identification of subtle social engineering cues (Fitzgerald & Bonnie, 2025). Complementing these, deepfake detection models utilise specialised deep learning algorithms to uncover inconsistencies in synthetic media, such as anomalies in facial expressions or audio artefacts. User Behaviour Analytics (UBA) systems further enhance security by continuously monitoring activity patterns to detect anomalies indicative of compromised accounts or phishing attempts (Fitzgerald L. , 2025). In addition, AI-driven threat intelligence platforms automate the aggregation and analysis of threat data, delivering real-time insights into emerging campaigns and attack vectors (Miller, 2025). Finally, user-initiated reporting mechanisms, such as “Report Phishing” buttons within email clients, empower individuals to flag suspicious content directly, thereby creating a feedback loop that strengthens organisational threat intelligence and facilitates rapid response (Harrington, 2025).

Despite the growing body of literature on AI-driven phishing, several critical gaps remain. Empirical evidence on the effectiveness of these AI-powered defence mechanisms, particularly in operational, real-world environments, is limited. Moreover, the ethical and social ramifications of malicious AI deployment in phishing contexts remain underexplored.

3. Methodology

Research Design and Methodology. This study employs a mixed-methods approach, integrating a qualitative analysis of existing literature and case studies with a quantitative evaluation of a real-world phishing simulation. This methodology was selected to provide a comprehensive understanding of the AI-driven phishing phenomenon, capturing both the technical dimensions of attack and defence, as well as the human factors that influence the success of such campaigns.

Data Collection. Data were collected from multiple sources, including peer-reviewed academic journals, industry reports, and technical documentation. The literature review specifically focused on articles and papers published between 2023 and 2025, ensuring the inclusion of the most recent and relevant research on AI-driven phishing. Case study data were obtained through a phishing simulation conducted by the authors, providing empirical insights into the real-world implications of AI-generated attacks.

Phishing Simulation Case Study. To assess the practical effectiveness of AI-generated phishing, a targeted simulation was conducted within a logistics company. Notably, this marked the organisation’s third phishing simulation, with staff having previously undergone security awareness training and exposure to prior simulated attacks. This context is significant, as it illustrates the potential of AI-generated phishing to circumvent even trained and partially vigilant user populations.

For the simulation, a Microsoft SharePoint email template was generated using a single prompt in ChatGPT and distributed to 125 employees. The study was designed to measure the following key metrics:

- Open Rate: The percentage of recipients who opened the phishing email;
- Click-Through Rate: The percentage of recipients who clicked on the phishing link within the email;
- Report Rate: The percentage of recipients who used the "Report Phishing" button to flag the suspicious email; and
- Compromise Rate: The percentage of recipients who entered their credentials on the phishing website.

As part of the organisation’s security awareness program, a “Report Phishing” button had been integrated into the email client, enabling users to promptly flag suspicious messages. The outcomes of this simulation are subsequently presented and analysed in the following section.

4. Analysis of Simulation Outcomes

This section presents the results of the phishing simulation case study and provides an in-depth analysis of its implications in the context of the broader research on AI-driven phishing. The findings are interpreted to address the paper's core objectives, compared with existing studies, and used to derive practical and theoretical contributions.

A phishing simulation was conducted to empirically evaluate the effectiveness of a phishing email generated using a single prompt from a Large Language Model (ChatGPT). The email, crafted to closely resemble a legitimate Microsoft SharePoint notification, was distributed to 125 employees within a logistics company. Importantly, this represented the organisation's third phishing simulation, indicating that the staff had already received security awareness training and had been exposed to previous simulated phishing attacks, thereby highlighting the capacity of AI-generated emails to challenge even a trained user base. This makes the results even more significant, as they demonstrate the effectiveness of AI-generated content against a trained user base. The results of this simulation are summarised in the table below.

Table 1: Phishing Simulation Results

Metric	Count	Percentage
Total Targets	125	100%
Emails Opened	64	51.2%
Phishing Link Clicked	50	40%
Emails Reported	48	38.4%
Credentials Compromised	15	12%

Source: (Data generated by the authors based on the phishing simulation, 2025)

The simulation results are striking and highlight the substantial threat posed by AI-generated phishing content. The observed click-through rate of 40% is notably high, substantially exceeding the typical rates reported for traditional phishing campaigns in comparable organisational settings. What is particularly concerning is that these results were obtained against a trained user base that had already participated in two previous phishing simulations, both of which employed traditional phishing emails. Despite prior exposure and security awareness training, a significant portion of the staff fell victim to the AI-generated phishing attempt. This outcome strongly supports the central hypothesis of this study: AI can be leveraged to produce highly convincing and effective phishing attacks with minimal effort, successfully bypassing the defences of security-aware users. Moreover, the fact that a single, simple prompt was sufficient to generate an email template that deceived a substantial portion of a trained target group underscores both the low barrier to entry for creating advanced social engineering attacks and the limitations of traditional security training against AI-generated content.

The 38.4% reporting rate (48 users) provides a nuanced perspective on the effectiveness of the "Report Phishing" button as a defensive measure. While this demonstrates that a notable portion of the trained user base identified the email as suspicious, it also highlights a critical gap: 51.2% of users (64 users) opened the email, yet only 38.4% (48 users) reported it, meaning that 25% of those who opened it (16 users) neither reported the email nor clicked the link, potentially reflecting uncertainty or passive dismissal. Even more concerning is that 40% of users (50 users) clicked the phishing link despite the availability of the reporting mechanism, indicating that the AI-generated content was sufficiently persuasive to override established caution.

The 12% compromise rate further underscores the severity of the threat. This indicates that a substantial subset of users who engaged with the email proceeded to enter their credentials, revealing a fundamental breakdown in security awareness and a high level of trust in the deceptive content. These findings are consistent with recent research, which has shown that AI-generated phishing emails can achieve click-through rates of up to 54%, compared to only 12% for human-written messages (Fitzgerald & Bonnie, 2025). The present simulation further corroborates these studies, demonstrating that AI-driven phishing represents a tangible, real-world threat rather than a merely theoretical concern (KnowBe4, 2025).

These findings directly address the primary objectives of this paper. They provide compelling evidence of the effectiveness of AI-driven attack techniques and highlight the limitations of existing defensive measures—which in this case included prior security awareness training, standard email security filters, and a user-initiated 'Report Phishing' button—when confronted with high-quality, AI-generated content. The elevated click-through and compromise rates indicate that traditional user awareness programs alone may be insufficient to mitigate the risks posed by personalised and contextually relevant lures crafted by AI. Importantly, the reporting data demonstrate that, while user-initiated mechanisms such as the "Report Phishing" button are valuable, they cannot function as a standalone defence: 38.4% of users reported the email, yet 40% still clicked the malicious link. This underscores the necessity of a multi-layered defence strategy that integrates advanced technical solutions with ongoing,

behaviour-focused user training. In this context, the reporting mechanism should be considered a critical feedback channel that strengthens organisational threat intelligence rather than serving as the primary line of defence.

Comparison with Other Studies. The 40% click-through rate observed in our simulation aligns with the broader trend of increasing phishing success rates facilitated by AI. Although some studies report even higher rates, our findings are particularly noteworthy for two key reasons. First, the phishing email was generated from a single, non-expert prompt, highlighting the minimal effort required to launch highly effective attacks. Second, the results were obtained against a user base that had previously participated in two phishing simulations and completed associated security training. These factors demonstrate that AI-generated phishing can successfully circumvent defences established through traditional security awareness programs. Furthermore, the findings corroborate research on human factors, which indicates that even with formal training and prior exposure to simulated attacks, users remain vulnerable to well-crafted, contextually appropriate AI-generated phishing emails (Basit, Zafar, Liu, & al., 2021). The present simulation thus serves as a practical case study, complementing laboratory-based and statistical research while providing critical real-world evidence of the limitations of current training methodologies.

Practical Implications and Theoretical Contribution. The practical implications of these findings are profound. Organizations must recognize that the threat landscape of phishing evolves beyond the scope of traditional training programs. Defences can no longer rely on detecting grammatical errors or generic greetings, and standard phishing simulations may no longer provide adequate preparation. The fact that a trained user base, which has previously completed two phishing simulations, still exhibits a 40% click-through rate underscores that conventional security awareness training alone is insufficient against AI-generated threats.

Similarly, the 38.4% reporting rate (48 users) demonstrates that the "Report Phishing" button is utilised by a notable portion of trained users, yet it also reveals its limitations as a standalone defence. The observation that 40% (50 users) click the phishing link despite having access to the reporting mechanism indicates that user-initiated reporting tools must be complemented by automated detection systems. Organisations should therefore regard the "Report Phishing" button not as a primary defence, but as a critical component of a layered security strategy that feeds threat intelligence into AI-powered detection systems.

Security strategies must now anticipate that phishing emails are well-crafted, personalised, and highly convincing, capable of deceiving even trained users. Addressing this requires investment in AI-powered email security solutions capable of analysing linguistic patterns and behavioural anomalies, as well as a fundamental shift towards continuous, adaptive, and AI-aware security awareness training. Such training extends beyond traditional simulation-based approaches and provides immediate, contextual feedback (Chen, Wu, Nguyen, & Rudolph; Miller, 2025). Furthermore, reporting mechanisms should be optimised through integration with automated threat intelligence platforms, allowing for rapid analysis of reported emails and real-time updates to defences.

5. Conclusion

The emergence of AI-driven phishing represents a fundamental and enduring shift in the cybersecurity threat landscape. This study demonstrates that the malicious use of generative AI is not a future concern but a present and rapidly escalating reality. The research highlights the techniques, impacts, and defence strategies associated with this new generation of phishing attacks, underscoring the urgent need for a redefined security paradigm.

This research identifies several critical findings. First, AI-powered tools, particularly Large Language Models, have democratised the creation of sophisticated phishing attacks, enabling malicious actors to produce personalised, contextually relevant, and grammatically flawless content. The case study conducted in this research illustrates this risk in practical terms: a single ChatGPT prompt generated a phishing email that achieved a 40% click-through rate and a 12% compromise rate, despite targeting a user base that had previously completed two phishing simulations. These results provide clear, real-world evidence of the effectiveness of AI-generated phishing and the inadequacy of traditional training approaches.

Second, traditional security defences, including conventional phishing awareness training, are increasingly insufficient against the dynamic and adaptive nature of AI-generated threats. The fact that trained employees who had been exposed to multiple previous simulations still fell victim at high rates demonstrates a fundamental limitation of current defensive strategies. The research clearly indicates that a multi-layered defence, integrating AI-powered detection systems with fundamentally reimagined, AI-aware user awareness training, is essential.

Third, the human element remains a critical vulnerability. Traditional training methods alone cannot address it. Security awareness initiatives must shift focus from simple compliance and pattern recognition to fostering

adaptive security resilience. While user-initiated reporting mechanisms such as the "Report Phishing" button demonstrate value—38.4% of users in our study utilised this tool—they cannot serve as a standalone defence. The observation that 40% of users still clicked the malicious link despite access to the reporting mechanism underscores the necessity of automated, AI-powered detection systems that complement user vigilance. Embedded, contextual training that provides immediate feedback is significantly more effective than periodic awareness campaigns; however, these programs may need to be further enhanced with AI-specific modules to teach users how to recognise the subtle cues of AI-generated content.

Finally, the ethical and social implications of malicious AI are profound. The erosion of trust in digital communications, coupled with challenges in regulation and attribution, demands a collaborative, global response involving policymakers, industry leaders, and the cybersecurity community. Addressing AI-driven phishing effectively will require both technical innovation and a reassessment of human factors in cybersecurity.

This paper provides clear answers to the initial research questions, demonstrating that AI-driven phishing employs techniques that differ markedly from traditional approaches in terms of scale, personalisation, and ability to evade detection, making them not only more efficient but also qualitatively more dangerous. The findings further indicate that effective security strategies require an integrated approach, combining AI-powered email filters, deepfake detection, user behaviour analytics, user-initiated reporting mechanisms such as "Report Phishing" buttons, and continuous, behaviour-focused security awareness training, as no single measure is sufficient on its own. Finally, the study underscores the broader ethical and social implications of AI-driven phishing, including the erosion of privacy, regulatory challenges, difficulties in attributing attacks, and a general decline in trust across digital ecosystems.

6. Limitations of the Paper

This paper, while comprehensive, has several limitations. The case study was confined to a single organization and focused on a specific type of phishing email, limiting the generalizability of the findings. Further research is necessary to validate these results across diverse industries and a broader range of AI-generated content. Moreover, given the rapidly evolving nature of AI, new attack vectors and defensive mechanisms are continuously emerging. Consequently, the findings presented here should be viewed as a snapshot of the current landscape rather than definitive or exhaustive conclusions.

7. Directions for Future Research

The fight against AI-driven phishing remains an ongoing challenge. Future research should prioritise several key areas. First, there is a pressing need to develop more advanced, real-time deepfake detection technologies. Second, further investigation into the psychological and cognitive factors that make individuals susceptible to AI-generated content is essential. Finally, the creation of international legal and regulatory frameworks to address the challenges of attribution and liability in the age of AI represents a critical priority.

In conclusion, the era of AI-driven phishing necessitates a proactive, adaptive, and collaborative approach to cybersecurity. The insights and strategies presented in this paper provide a foundation for strengthening defences and building a more resilient and secure digital environment.

References

1. Basit, A., Zafar, M., Liu, X., & al., e. (2021). A comprehensive survey of AI-enabled phishing attacks detection techniques. Telecommun Syst 76, 139–154.
2. Jabir, R., Le, J., & Nguyen, C. (2025). Phishing Attacks in the Age of Generative Artificial Intelligence: A Systematic Reviewof Human Factors. AI, 6(8) https://doi.org/10.3390/ai6080174, 174.
3. Khalil, M. (2025, 04 29). Phishing Statistics 2025: AI-Driven Attacks, Costs & Trends. Retrieved from DeepStrike: https://deepstrike.io/blogs/Phishing-Statistics-2025
4. Letain-Mathieu, G. (2025, 10). AI-Generated Phishing: The Top Enterprise Threat of 2025 . Retrieved from StrongestLayer: https://www.strongestlayer.com/blog/ai-generated-phishing-enterprise-threat-2025
5. Chen, F., Wu, T., Nguyen, V., & Rudolph, C. (n.d.). SoK: Large Language Model-Generated Textual Phishing Campaigns End-to-End Analysis of Generation, Characteristics, and Detection. Retrieved from 10.48550/arXiv.2508.21457. .

6. Arntz, P. (2025, 1 7). AI-supported spear phishing fools more than 50% of targets. Retrieved from Malwarebytes : https://www.malwarebytes.com/blog/news/2025/01/ai-supported-spear-phishing-fools-more-than-50-of-targets
7. IRONSCALES (n.d.). (2025, 11). What are Polymorphic Attacks? Retrieved from IRONSCALES: https://ironscales.com/glossary/polymorphic-attacks
8. Fitzgerald, L. (2025, 3 13). How Deepfake Voice Detection Works. Retrieved from Pindrop: How Deepfake Voice Detection Works
9. Fitzgerald, A., & Bonnie, E. (2025, 08 14). 60+ Phishing Attack Statistics: The Facts You Need To Know for 2026. Retrieved from Secureframe: https://secureframe.com/blog/phishing-attack-statistics
10. Miller, J. (2025, 10 15). Real-World Examples of AI in Cyber Threat Detection. Retrieved from BitLyft: https://www.bitlyft.com/resources/real-world-examples-of-ai-in-cyber-threat-detection
11. Harrington, L. (2025, 04 08). Data from 2024 Phishing Tests Reveals How Human-Targeted Threats Are Evolving. Retrieved from proofpoint: https://www.proofpoint.com/us/blog/email-and-cloud-threats/phish-tests-reveal-human-targeted-threats-evolving)
12. Eze, C. S., & Shamir, L. (2024). Analysis and prevention of AI-based phishing email attacks. Electronics, 13(10), 1839.
13. Schmitt, M., & Flechais, I. (2024). Digital deception: generative artificial intelligence in social engineering and phishing. Artificial Intelligence Review, 57(12), Article 324.
14. KnowBe4. (2025, 05 13). KnowBe4 Report Reveals Security Training Reduces Global Phishing Click Rates by 86%. Retrieved from KnowBe4: https://www.knowbe4.com/press/knowbe4-report-reveals-security-training-reduces-global-phishing-click-rates-by-86
15. Gallagher, S. (2024). Phishing and Social Engineering in the Age of LLMs. In The Ethics of AI and Cybersecurity. Springer, Cham.