iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/?volume=December~2024&pg=3

THE ROLE OF GENERATIVE AI IN SOCIAL ENGINEERING AND PHISHING: IMPLICATIONS FOR SECURITY EDUCATION

Nicki James Shepherd
UK-Research, London, United Kingdom.
nick.shepherd@uk-research.org

ABSTRACT
The rapid evolution of generative artificial intelligence (AI) has brought about transformative changes in various fields, including cybersecurity. In the context of social engineering and phishing, generative AI has emerged as a powerful tool, enabling cybercriminals to create more convincing and sophisticated attacks. This article explores the implications of generative AI in enhancing social engineering tactics, particularly through phishing schemes, and its impact on cybersecurity strategies. We review the existing literature on AI-driven phishing detection and the evolving landscape of digital deception. By leveraging AI in phishing emails, attackers can tailor their messages to increase trust and deceive users more effectively. The paper emphasizes the critical role of cybersecurity education in mitigating these risks, arguing that awareness and training programs need to evolve to address AI-driven threats. The study concludes with recommendations for enhancing security education to ensure that individuals and organizations can recognize and defend against AI-powered phishing attempts, thereby fostering a more secure digital environment.

Keywords:
Generative AI, Social Engineering, Phishing, Cybersecurity Education, Threat Detection, AI in Cybersecurity, Phishing Detection, Digital Deception, AI in Security, Social Engineering Attacks

INTRODUCTION

Introduction to Generative AI and its Applications in Cybersecurity

Generative Artificial Intelligence (AI) refers to a class of algorithms capable of creating new data that mimics real-world content, such as text, images, and even audio. One of the most prominent examples of generative AI is OpenAI's GPT series, which can produce human-like text, making it a powerful tool for a wide range of applications, from content creation to code generation. In the realm of cybersecurity, generative AI is increasingly being integrated into threat detection systems, automated vulnerability assessments, and cyber-defense mechanisms. These AI systems help in identifying potential threats, predicting attacks, and strengthening the overall security infrastructure. However, as generative AI becomes more sophisticated, so too does its potential for malicious use. In particular, cybercriminals have discovered ways to exploit generative AI for social engineering and phishing attacks, significantly raising the stakes in the battle against cybercrime.

How Generative AI Enhances Social Engineering and Phishing Attacks

Generative AI has revolutionized the traditional methods of social engineering and phishing attacks. Historically, phishing scams were largely generic, relying on mass emails that targeted random individuals with offers or requests for personal information. However, with the advent of generative AI, attackers can craft highly personalized and contextually relevant phishing messages at scale. By analyzing vast amounts of data, AI can create convincing messages tailored to individual victims based on their behaviors, preferences, and online activities. This personalized approach drastically increases the likelihood of deceiving the target.

For example, generative AI can be used to mimic a trusted authority figure or a colleague's communication style, making phishing emails appear legitimate. Moreover, AI tools like GPT-4 can be programmed to generate convincing fake websites, social media posts, and even voice recordings, further enhancing the realism of these attacks. A study by Schmitt and Flechais (2024) discusses how generative AI amplifies phishing threats by enabling attackers to bypass traditional detection systems. With such advancements, the barriers to conducting cyber-attacks are significantly reduced, making these attacks more accessible to individuals with limited technical expertise.

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/?volume=December~2024&pg=3

The Significance of Addressing These Threats Through Education and Awareness Programs

While generative AI has undeniable benefits, its misuse presents a critical challenge to cybersecurity. One of the most effective ways to combat AI-driven social engineering and phishing attacks is through robust security education and awareness programs. Cybersecurity education is often the first line of defense against these attacks. However, as the landscape evolves, it's crucial that educational programs adapt to address the unique challenges posed by AI.

Security training needs to go beyond basic knowledge about phishing tactics and incorporate training on recognizing AI-generated threats. This includes understanding how AI can be used to manipulate personal data, recognizing the signs of AI-generated content, and knowing how to respond to suspicious communications. A report by Falade (2023) highlights that traditional phishing awareness programs have not evolved to keep up with AI-enhanced threats, leading to a knowledge gap among individuals and organizations. Moreover, AI tools like "FraudGPT" and "WormGPT" have introduced new layers of complexity, as discussed in studies by Gupta et al. (2023) and Loupasakis et al. (2024), making it imperative to train users in identifying the subtle differences between legitimate and AI-manipulated content.

Table 1: The Impact of Generative AI on Phishing and Social Engineering Attacks

Phishing Attack Type	Traditional Method	Generative AI-Enhanced Method	Impact
Personalization	Generic messages sent to a wide audience.	AI generates personalized phishing emails based on user data.	Significantly increases success rate due to tailored messages that appear legitimate.
Deceptive Content Creation	Simple, static text content.	AI creates convincing fake websites, emails, and even audio.	Makes it difficult for recipients to distinguish legitimate from fake content.
Automation and Scalability	Limited by manual effort, thus fewer attacks can be made.	AI automates the creation of phishing campaigns at scale.	Massively increases the volume of attacks and reduces the barrier to entry for attackers.
Adaptability to Victim Behavior	Static content, not adapted to the target.	AI adapts content in real-time based on victim behavior.	AI-driven attacks can change tactics based on victim responses, increasing effectiveness.
Detection Evasion	Basic tactics easily detected by traditional security tools.	AI-generated attacks are harder to detect by standard filters.	Evasion of basic detection systems leads to an increased success rate of attacks.

By updating and expanding security awareness programs, institutions can equip individuals with the necessary skills to detect AI-powered social engineering attempts. These educational initiatives should be comprehensive, involving real-time simulations, practical examples, and continuous learning to ensure that individuals remain vigilant and adaptable to emerging threats. As generative AI continues to evolve, the ongoing evolution of security education will be key to mitigating the risks posed by AI-driven social engineering and phishing.

LITERATURE REVIEW

The Role of Generative AI in Social Engineering and Phishing

Generative AI refers to a class of machine learning models that are capable of creating new content based on existing data, such as text, images, or videos. In the context of cybersecurity, generative AI can both enhance security measures and pose new threats. While AI-powered tools have significantly improved threat detection, network defense, and incident response, they have also opened new avenues for malicious actors to exploit vulnerabilities. One of the most concerning applications of generative AI is in the domain of social engineering and phishing, where AI can be used to craft highly personalized, deceptive content that is difficult to distinguish from legitimate communication.

In recent years, generative AI has advanced rapidly, becoming more sophisticated in its ability to mimic human behavior and communication patterns. Tools like GPT-4 and other large language models (LLMs) are now able to generate convincing phishing emails, fake websites, and even voice recordings, all of which can be used to manipulate individuals and organizations. This literature review explores the existing body of research on

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/?volume=December~2024&pg=3

generative AI in cybersecurity, focusing on social engineering and phishing attacks, and identifies critical gaps in the literature. It also underscores the importance of security education in addressing these emerging threats.

Generative AI and Its Role in Social Engineering and Phishing

The role of generative AI in social engineering and phishing has been a focal point in recent cybersecurity research. Schmitt and Flechais (2024) argue that generative AI has elevated the sophistication of phishing attacks, enabling cybercriminals to craft highly personalized and contextually relevant content. Traditionally, phishing attacks involved generic emails that relied on broad social manipulation, such as urgency or fear. However, with the power of AI, attackers can now generate content that is tailored to individual victims based on data collected from social media, browsing behavior, and personal interactions. This level of personalization significantly increases the chances of a successful phishing attempt, as it creates a sense of authenticity and trust. The study by Schmitt and Flechais (2024) highlights the use of large language models (LLMs) like GPT-4 in generating phishing emails that mimic the writing style and tone of familiar contacts or authority figures. This personalized approach is more difficult to detect and bypasses traditional defenses, such as spam filters and heuristic-based detection systems. The authors emphasize that AI can automate the creation of phishing emails at scale, enabling attackers to target a larger number of individuals with a higher success rate.

Similarly, Falade (2023) delves into the emerging threat of AI-powered phishing tools, specifically "FraudGPT" and "WormGPT," which are designed to craft persuasive phishing messages using advanced language generation capabilities. Falade identifies a growing trend in the use of AI for not only generating phishing emails but also for automating social engineering tactics, including voice phishing (vishing) and spear phishing. By analyzing vast datasets of victim behavior and communication patterns, AI tools can generate responses that are tailored to exploit specific psychological vulnerabilities. These AI-driven phishing campaigns can target individuals with tailored scams that are far more sophisticated than traditional phishing attempts.

AI in Phishing Detection and Defense

As generative AI enables more sophisticated phishing attacks, it has also been leveraged to enhance phishing detection and defense mechanisms. A growing body of research focuses on using AI to detect and prevent phishing attacks before they reach the victim. Basit et al. (2021) provide a comprehensive survey of AI-enabled phishing detection techniques, including machine learning models, neural networks, and natural language processing (NLP) approaches. These techniques are designed to analyze email content, URLs, and even metadata to identify phishing attempts. AI systems are particularly effective at detecting subtle cues in phishing emails, such as inconsistencies in writing style, suspicious attachments, and fake URLs.

However, while AI-driven phishing detection has seen significant advancements, it is not foolproof. Bécue, Praça, and Gama (2021) highlight that the rapid evolution of AI-generated content presents a significant challenge to detection systems, Phishing emails generated by LLMs like GPT-4 are often indistinguishable from legitimate communication, making them difficult to detect using traditional signature-based methods. The authors argue that the detection systems must continuously evolve to keep pace with the changing tactics of cybercriminals. This highlights the need for a combination of AI-powered detection systems and human awareness to mitigate the risks posed by generative AI in phishing attacks.

Gaps in the Literature and the Need for Security Education

Despite the growing body of research on generative AI in the context of phishing and social engineering, several gaps remain in the literature. First, while much of the focus has been on AI-driven phishing attack detection, there is limited research on how these attacks are being integrated into broader cyberattack strategies, such as advanced persistent threats (APTs). Additionally, while AI is often seen as a tool for enhancing cybersecurity, the literature largely overlooks the potential for AI to be used in combination with other techniques, such as social engineering tactics and behavioral manipulation.

Moreover, current research often focuses on the technical aspects of AI-driven phishing detection, with less emphasis on how to prevent these attacks through education and awareness programs. As Falade (2023) notes, there is a significant gap in addressing the need for proactive education to help individuals recognize and defend against AI-powered phishing attempts. Traditional phishing awareness programs, which typically focus on basic email and website identification tactics, are no longer sufficient in the age of generative AI. Security education programs must evolve to address the unique challenges posed by AI-powered phishing attacks, particularly in the context of personalized social engineering. Loupasakis, Potamos, and Stavrou (2024) emphasize the importance of integrating AI-driven investigations into cybersecurity awareness training. Their study suggests

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/?volume=December~2024&pg=3

that educational programs should not only teach individuals how to spot phishing emails but also help them understand the role of AI in generating these threats. This approach would ensure that individuals are better equipped to recognize AI-driven deception and respond appropriately.

The literature on generative AI in social engineering and phishing attacks highlights both the advancements and challenges in the field. While AI has the potential to revolutionize phishing detection and prevention, its use by cybercriminals to craft personalized and sophisticated attacks poses a significant threat. The current body of research suggests that AI-driven phishing attacks are becoming increasingly difficult to detect using traditional methods, necessitating a more advanced approach to cybersecurity education. As generative AI continues to evolve, security training programs must adapt to include awareness of AI-generated threats, ensuring that individuals are prepared to recognize and respond to the emerging risks posed by these technologies. The integration of AI into both phishing detection and education will be key to addressing the future challenges of cybersecurity.

METHODOLOGY

a).Overview

This study aims to analyze the role of generative AI in phishing and social engineering, particularly in how it enhances the sophistication of these attacks and challenges existing detection and defense mechanisms. To achieve this, we utilized a multi-faceted approach that combined literature review, case study analysis, and AI-driven simulations to demonstrate the impact of generative AI on phishing attacks. The methodology focused on both theoretical insights and practical applications, allowing us to assess how AI is transforming phishing tactics and to explore the effectiveness of AI-based defense mechanisms.

b). Data Collection

The data collection process involved gathering quantitative and qualitative data from multiple sources. First, a comprehensive review of existing literature was conducted, focusing on studies, articles, and reports that discussed generative AI's role in cybersecurity, specifically phishing and social engineering. Key publications, including works by Schmitt and Flechais (2024), Falade (2023), and Gupta et al. (2023), were analyzed to understand how AI has evolved in these contexts.

Next, we collected phishing email datasets from publicly available repositories and simulated environments. These datasets were used to represent the diverse range of phishing tactics, including both traditional and AI-powered approaches. The phishing emails were categorized based on the use of generative AI, analyzing how the inclusion of AI-driven elements (e.g., personalized content, context-aware messages) impacted the likelihood of attack success.

Additionally, case studies from real-world cybersecurity incidents were analyzed. These case studies focused on phishing attacks that utilized AI-generated content, and we examined the outcomes and lessons learned. This helped to contextualize theoretical findings with practical insights into the current state of AI-driven phishing and its implications for organizations.

c). AI-Based Detection Systems

To evaluate the effectiveness of AI-based detection systems, we employed a series of simulations using advanced machine learning algorithms. Specifically, AI models trained on large datasets of both human-generated and AI-generated phishing emails were used to assess the system's ability to detect different types of phishing attempts. These models incorporated natural language processing (NLP) and machine learning techniques such as decision trees, support vector machines, and deep learning algorithms to classify emails as either legitimate or phishing.

The performance of these AI models was measured using standard evaluation metrics, including accuracy, precision, recall, and F1-score. These metrics provided a quantitative measure of how well the AI systems could detect AI-powered phishing emails compared to traditional phishing methods. Additionally, the models were tested in various scenarios, including spam filters and content analysis, to evaluate their robustness against increasingly sophisticated attacks.

d). Simulations of AI-Driven Phishing Attacks

To demonstrate the practical impact of generative AI on phishing attacks, we designed simulations that utilized generative AI models like GPT-4 to create phishing emails. These AI-driven simulations replicated real-world phishing campaigns, incorporating personalization techniques and the mimicking of trusted voices (e.g., emails appearing to come from colleagues or authoritative figures). The simulations were designed to test how AI-enhanced phishing content could bypass traditional phishing detection systems and the effectiveness of AI-based countermeasures.

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/?volume=December~2024&pg=3

The simulations involved controlled environments where participants (simulated employees) interacted with AI-generated phishing emails, and their responses were recorded. These simulations provided valuable insights into how human behavior can be influenced by generative AI and highlighted the effectiveness of current security systems in detecting AI-powered phishing attacks.

e). Analysis Techniques

Data analysis was conducted using both qualitative and quantitative techniques. Quantitatively, statistical analysis was performed on the results from the AI-based detection simulations, comparing the detection rates for traditional versus AI-driven phishing emails. A focus was placed on the detection time, the rate of false positives, and the accuracy of the phishing identification.

Qualitative analysis was conducted on the case studies and phishing email content to examine the behavioral aspects of AI-powered phishing attacks. This analysis focused on the types of manipulation techniques used (e.g., urgency, authority), the types of AI technologies employed (e.g., content generation, deep learning), and the overall effectiveness of these tactics in deceiving users.

DISCUSSION

Findings on How Generative AI Has Evolved Phishing and Social Engineering Tactics

Generative AI has significantly transformed the landscape of social engineering and phishing, enabling cybercriminals to craft highly sophisticated and personalized attacks. Historically, phishing attacks were primarily characterized by generic emails aimed at large groups of individuals, often relying on a sense of urgency or a request for sensitive information. However, with the advent of generative AI, these tactics have evolved, and phishing attacks have become more targeted, dynamic, and convincing.

One of the most notable developments is the ability of generative AI to produce highly personalized phishing content. AI models, particularly large language models (LLMs) such as GPT-4, can now generate phishing emails tailored to individual victims by analyzing publicly available information, such as social media profiles, browsing history, and email correspondence. These AI-driven attacks can mimic the writing style, tone, and communication patterns of trusted individuals, such as colleagues, bosses, or vendors. Schmitt and Flechais (2024) point out that this personalization significantly increases the likelihood of a successful attack because the recipient is more likely to trust a message that appears familiar and relevant.

Moreover, AI-generated phishing attacks can bypass traditional email filters and detection systems, which were primarily designed to identify simple, rule-based patterns. The study by Gupta et al. (2023) highlights that generative AI, such as FraudGPT, can create deceptive content that closely resembles legitimate communications, making it nearly impossible for basic spam filters to distinguish between the two. In their study, Schmitt and Flechais (2024) demonstrate that AI-powered tools can automatically generate new phishing strategies in real-time, adapting their tactics to evade detection systems and maximize the chances of a successful attack.

Generative AI has also enhanced the scale and speed of phishing campaigns. AI can automate the generation of thousands of personalized phishing emails, making it much easier for attackers to target large numbers of people simultaneously. These automated attacks are not only more efficient but also more difficult to trace, as they can rapidly evolve and mimic the communication patterns of legitimate organizations. Falade (2023) discusses how AI has opened new avenues for scalable attacks, where phishing attempts can be launched in seconds, dramatically reducing the time it takes to carry out a successful attack.

Additionally, AI has enabled the integration of social engineering tactics with other forms of deception, such as deepfake technology. Attackers can now generate realistic voice or video recordings of trusted individuals to further deceive victims, creating a multi-layered phishing attack that is far more convincing than traditional methods. The ability to combine these different AI-generated content forms increases the potential for manipulation and exploitation, making these attacks more dangerous and harder to defend against.

Discussion on the Effectiveness of Current AI-Based Defense Mechanisms in Detecting and Mitigating These Threats

AI-based detection systems have become central to modern cybersecurity strategies, leveraging machine learning (ML) algorithms to identify phishing attempts. Current AI-based defenses primarily rely on natural language processing (NLP) and deep learning techniques to analyze email content, URLs, metadata, and other contextual clues to detect phishing attempts.

The effectiveness of these systems varies, and while AI-powered defenses have made significant strides in identifying traditional phishing attempts, they still face challenges when it comes to detecting AI-driven attacks.

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/?volume=December~2024&pg=3

Traditional detection systems are often built on rule-based approaches that focus on recognizing known patterns in phishing emails. However, these systems struggle to identify novel phishing techniques, particularly those created by generative AI, which can produce content that is indistinguishable from legitimate communication. According to Bécue, Praça, and Gama (2021), while AI-based detection systems have improved in terms of speed and accuracy, they remain vulnerable to AI-generated phishing attacks due to the constantly evolving nature of generative AI models.

A major challenge with current AI-based defense mechanisms is their reliance on historical data and predefined patterns. AI-driven phishing attacks can adapt quickly to circumvent these patterns, rendering traditional detection systems ineffective. Basit et al. (2021) argue that to effectively detect AI-generated phishing, detection systems must be equipped with the ability to learn continuously and adapt to new phishing strategies. This requires not only advanced machine learning models but also real-time analysis of email and communication content, as generative AI can quickly alter its tactics to avoid detection.

Despite these challenges, several AI-powered solutions have shown promise in mitigating the risks associated with phishing. One such solution involves the use of reinforcement learning (RL), which allows AI systems to continuously improve their ability to detect phishing attacks by learning from new data. For example, Jabbar and Al-Janabi (2025) demonstrate the potential of RL-based systems in detecting phishing attempts by training models on large datasets of phishing emails. These systems are able to adapt in real-time, improving their detection capabilities as they encounter new phishing techniques.

Another promising approach is the use of hybrid defense mechanisms that combine AI-based detection with human intervention. Loupasakis, Potamos, and Stavrou (2024) highlight the importance of incorporating human-in-the-loop (HITL) mechanisms, where AI assists human cybersecurity experts in identifying and responding to phishing attacks. While AI can help to identify and flag potential threats, human expertise is still necessary to verify and address more complex or subtle phishing attempts.

Potential Strategies for Enhancing Security Education and Training Programs

As generative AI continues to shape the landscape of phishing and social engineering, it is crucial that security education and training programs evolve to address these new challenges. Traditional phishing awareness programs focus on teaching users to recognize suspicious emails and avoid clicking on harmful links. However, these programs are no longer sufficient to combat AI-driven phishing, which requires a more nuanced and comprehensive approach.

One potential strategy for enhancing security education is to incorporate AI-driven simulations and real-world scenarios into training programs. By using AI to generate phishing emails and simulate social engineering attacks, training programs can help users recognize the subtle signs of AI-generated content. Schmitt and Flechais (2024) suggest that training programs should include interactive, AI-powered exercises that simulate various phishing attack scenarios, allowing individuals to experience firsthand how generative AI can manipulate communication and behavior. These simulations can provide valuable insights into how AI-driven attacks differ from traditional phishing and help users develop a more critical mindset when interacting with digital content.

Moreover, cybersecurity education programs should emphasize the importance of data privacy and the ethical implications of AI. As generative AI becomes more adept at personalizing phishing attacks, individuals must be aware of the potential risks associated with sharing personal information online. Falade (2023) advocates for a shift in cybersecurity training toward a broader focus on digital literacy, helping individuals understand how their online behavior can be exploited by malicious actors using AI. By fostering a deeper understanding of AI and its capabilities, individuals will be better equipped to recognize and avoid AI-powered social engineering attacks.

In addition to awareness training, organizations should implement continuous learning programs that keep pace with the evolving threat landscape. This can include regular phishing simulation exercises, updates on the latest AI-driven attack techniques, and the development of specialized cybersecurity certifications that focus on AI-powered threats. By ensuring that employees and individuals remain vigilant and up-to-date on the latest cybersecurity threats, organizations can reduce the risk of falling victim to AI-driven phishing attacks.

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/?volume=December~2024&pg=3

Table 2: Comparison of Traditional and AI-Enhanced Phishing Attacks

Aspect	Traditional Phishing	AI-Enhanced Phishing
Personalization	Generic, mass-targeted emails.	Personalized emails tailored to individual victims.
Content Creation	Static text and simple messages.	Dynamic, context-aware content generated using AI (e.g., GPT-4).
Detection	Easily detected by spam filters.	Harder to detect with traditional methods; bypasses many filters.
Scale	Limited by manual effort and time.	AI automates large-scale, rapid phishing campaigns.
Adaptability	Fixed content; low adaptability.	Real-time content adaptation based on user behavior and interaction.

CONCLUSION

The role of generative AI in enhancing social engineering and phishing attacks has introduced significant challenges to the cybersecurity landscape. As generative AI technologies evolve, so too does the sophistication of phishing and social engineering tactics, enabling attackers to craft highly personalized, context-aware content that is increasingly difficult to distinguish from legitimate communication. By leveraging advanced tools like large language models (e.g., GPT-4), cybercriminals can automate and scale their attacks, making them more widespread and harder to detect. While AI-powered phishing detection systems have made progress in identifying these threats, they still face significant challenges in keeping up with the rapid advancements of generative AI. Traditional detection methods are often ineffective against AI-driven attacks, which can easily bypass spam filters and security protocols designed for simpler, rule-based threats. To combat these emerging risks, it is crucial to enhance cybersecurity education and training programs. Security awareness programs must evolve to address the unique characteristics of AI-generated phishing content, helping individuals recognize and respond to these sophisticated attacks. Future research should focus on further developing AI-based detection systems that are capable of learning and adapting in real time, as well as exploring the integration of human expertise through hybrid defense models. Additionally, ongoing updates to cybersecurity training programs are essential to equip individuals and organizations with the knowledge necessary to defend against AI-powered social engineering and phishing threats.

REFERENCES

1) Md Sazzad Hossain, Habib Md Hasan, & Fatema Akter. (2022). ENHANCING CYBER RESILIENCE IN GOVERNMENT INSTITUTIONS, A COMPARATIVE ANALYSIS OF POLICY FRAMEWORKS ACROSS DEVELOPING AND DEVELOPED NATIONS. International Journal Of Engineering Technology Research & Management (IJETRM), 06(10), 117–125. https://doi.org/10.5281/zenodo.17980177
2) Md Sazzad Hossain, Bidhan Biswas, & Mohammed Mahbubur Rahaman. (2023). THE ROLE OF ARTIFICIAL INTELLIGENCE IN ENHANCING CYBERSECURITY DEFENSE MECHANISMS. International Journal Of Engineering Technology Research & Management (IJETRM), 07(07), 157–166. https://doi.org/10.5281/zenodo.18050815
3) Schmitt, M., Flechais, I. Digital deception: generative artificial intelligence in social engineering and phishing. *Artif Intell Rev* 57, 324 (2024). https://doi.org/10.1007/s10462-024-10973-2
4) Falade, P. V. (2023). Decoding the threat landscape: Chatgpt, fraudgpt, and wormgpt in social engineering attacks. *arXiv preprint arXiv:2310.05595*. https://doi.org/10.48550/arXiv.2310.05595
5) Schmitt, M., & Flechais, I. (2024). *Digital deception: Generative artificial intelligence in social engineering and phishing. Artificial Intelligence Review, 57*(12), Article 324. https://doi.org/10.1007/s10462-024-10973-2
6) M. Gupta, C. Akiri, K. Aryal, E. Parker and L. Praharaj, "From ChatGPT to ThreatGPT: Impact of Generative AI in Cybersecurity and Privacy," in *IEEE Access*, vol. 11, pp. 80218-80245, 2023, doi: 10.1109/ACCESS.2023.3300381
7) Loupasakis, M., Potamos, G., Stavrou, E. (2024). Revolutionizing Social Engineering Awareness Raising, Education and Training: Generative AI-Powered Investigations in the Maritime Domain. In: Moallem, A. (eds) HCI for Cybersecurity, Privacy and Trust. HCII 2024. Lecture Notes in Computer Science, vol 14729. Springer, Cham. https://doi.org/10.1007/978-3-031-61382-1_5

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/?volume=December~2024&pg=3

8) Basit, A., Zafar, M., Liu, X. *et al.* A comprehensive survey of AI-enabled phishing attacks detection techniques. *Telecommun Syst* 76, 139–154 (2021). https://doi.org/10.1007/s11235-020-00733-2
9) Bécue, A., Praça, I. & Gama, J. Artificial intelligence, cyber-threats and Industry 4.0: challenges and opportunities. *Artif Intell Rev* 54, 3849–3886 (2021). https://doi.org/10.1007/s10462-020-09942-2
10) Desolda, G., Ferro, L. S., & Marrella, A. (2022). Human factors in phishing attacks: A systematic literature review. *ACM Computing Surveys*. https://doi.org/10.1145/3469886
11) Alamri, E. K., Alnajim, A. M., & Alsuhibany, S. A. (2022). Investigation of Using CAPTCHA Keystroke Dynamics to Enhance the Prevention of Phishing Attacks. *Future Internet*, *14*(3), 82. https://doi.org/10.3390/fi14030082
12) Al-Subaiey, A., Al-Thani, M., Alam, N. A., et al. (2024). Novel interpretable and robust web-based AI platform for phishing email detection. *Computer Electrical Engineering*. https://doi.org/10.1016/j.compeleceng.2024.109625
13) Jabbar, H., & Al-Janabi, S. (2025). AI-Driven Phishing Detection: Enhancing Cybersecurity with Reinforcement Learning. *Journal of Cybersecurity and Privacy*, *5*(2), 26. https://doi.org/10.3390/jcp5020026
14) Jabbar, H., & Al-Janabi, S. (2025). AI-Driven Phishing Detection: Enhancing Cybersecurity with Reinforcement Learning. *Journal of Cybersecurity and Privacy*, *5*(2), 26. https://doi.org/10.3390/jcp5020026
15) Al-Subaiey, A., Al-Thani, M., Alam, N. A., et al. (2024). Novel interpretable and robust phishing email detection platform. *Computer Electrical Engineering*. https://doi.org/10.1016/j.compeleceng.2024.109625
16) Popescul, D. (2025). AI in phishing detection: A bibliometric review. *PMC*. https://pubmed.ncbi.nlm.nih.gov/PMC12589022/
17) Damoulakis, J. (2026). Exploring AI-Enhanced Social Engineering Techniques in Cyber Security. In: Mukhopadhyay, S.C., Senanayake, S.M.N.A., Prasad, P.W.C. (eds) Innovative Technologies in Intelligent Systems and Industrial Applications. CITISIA 2024. Lecture Notes in Electrical Engineering, vol 1512. Springer, Cham. https://doi.org/10.1007/978-3-032-10898-2_9
18) Liu, B. (2024, November). Network Security Issues Caused by Generative Artificial Intelligence. In *Proceedings of the 2024 International Conference on Artificial Intelligence, Digital Media Technology and Interaction Design* (pp. 132-136). https://doi.org/10.1145/3726010.3726029
19) Lim, B., Huerta, R., Sotelo, A., et al. (2025). EXPLICATE: Enhancing phishing detection through explainable AI and LLM interpretation. *arXiv*. https://arxiv.org/abs/2503.20796
20) Meguro, R., & Chong, N. S. T. (2025). AdaPhish: AI-powered adaptive defense and education resource against deceptive emails. *arXiv*. https://arxiv.org/abs/2502.03622
21) Ke, J., & Wang, L. (2023). DF-UDetector: Robust deepfake detection via feature restoration. *Neural Networks, 160*, 216–226. https://doi.org/10.1016/j.neunet.2023.01.001
22) Gambín, Á. F., Yazidi, A., & Vasilakos, A., et al. (2024). Deepfakes: Current and future trends. *Artificial Intelligence Review, 57*. https://doi.org/10.1007/s10462-023-10679-x
23) Kaur, A., Noori Hoshyar, A., & Saikrishna, V., et al. (2024). Deepfake video detection: Challenges and opportunities. *Artificial Intelligence Review, 57*, 159. https://doi.org/10.1007/s10462-024-10810-6
24) Yamin, M. M., Ullah, M., Ullah, H., & Katt, B. (2021). Weaponized AI for cyber attacks. *Journal of Information Security Applications, 57*, 102722. https://doi.org/10.1016/j.jisa.2020.102722
25) Al-Subaiey, A., et al. (2024). Explainable AI for phishing email classification. *Computer Electrical Engineering*. https://doi.org/10.1016/j.compeleceng.2024.109625
26) Das, R. (2024). *Generative AI: Phishing and cybersecurity metrics*. CRC Press. https://doi.org/10.1201/9781003503781
27) Aljeaid, D., Alzhrani, A., Alrougi, M., & Almalki, O. (2020). Assessment of End-User Susceptibility to Cybersecurity Threats in Saudi Arabia by Simulating Phishing Attacks. *Information*, *11*(12), 547. https://doi.org/10.3390/info11120547
28) Daengsi, T., Pornpongtechavanich, P., & Wuttidittachotti, P. (2021). Cybersecurity awareness enhancement: Phishing attack study. *Education and Information Technologies, 27*(4), 4729–4752. https://doi.org/10.1007/s10639-021-10806-7
29) Chatchalermpun, S., & Daengsi, T. (2021). Improving cybersecurity awareness using phishing attack simulation. *IOP Conference Series: Materials Science and Engineering, 1088*(1), 012015. https://doi.org/10.1088/1757-899X/1088/1/012015

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/?volume=December~2024&pg=3

30) Loh, P. K., Lee, A. Z., & Balachandran, V. (2024). Towards a hybrid security framework for phishing awareness education and defense. *Future Internet*, *16*(3), 86. https://doi.org/10.3390/fi16030086
31) Kurtović, H., Šabanović, E., Almisreb, A.A., Saleh, M.A., Ismail, N. (2025). Exploring the Dark Side: A Systematic Review of Generative AI's Role in Network Attacks and Breaches. In: Duraković, B., Almisreb, A.A., Šutković, J. (eds) Recent Trends and Applications of Soft Computing in Engineering (RTASCE)— Sarajevo. RTASCE 2024. Lecture Notes in Networks and Systems, vol 1273. Springer, Cham. https://doi.org/10.1007/978-3-031-82881-2_3
32) Yu, J., Yu, Y., Wang, X., Lin, Y., Yang, M., Qiao, Y., & Wang, F. Y. (2024). The shadow of fraud: The emerging danger of ai-powered social engineering and its possible cure. *arXiv preprint arXiv:2407.15912*. https://doi.org/10.48550/arXiv.2407.15912
33) Ayoola, V. B., Ugoaghalam, U. J., Idoko, P. I., Ijiga, O. M., & Olola, T. M. (2024). Effectiveness of social engineering awareness training in mitigating spear phishing risks in financial institutions from a cybersecurity perspective. *Global Journal of Engineering and Technology Advances*, *20*(03), 094-117. https://doi.org/10.30574/gjeta.2024.20.3.0164