Developing an AI-Augmented Cybersecurity Framework to Prevent Social Engineering Attacks

Basant Kumar[1], Sreemoy[2], Shashi Kant Gupta[3], Rashmi Dwivedi[4], Afaq Ahmed[5]

[1] Lincoln University College, Malaysia; [2] Lincoln University College, Malaysia; [3] Chitkara University

[4]Muscat University, Oman; [5]Modern College of Business and Science, Oman

pdf.basantkumar@lincoln.edu.my; sreemoy@lincoln.edu.my; raj2008enator@gmail.com;

rdwivedi@muscatuniversity.edu.om; Afaq.Ahmed@mcbs.edu.om

Abstract: Cyber-attacks tend to capitalize on human errors, and this research examines how social engineering attacks can flourish on these vulnerabilities. It advocates for an AI-powered cybersecurity system to better detect and counter these tactics. Using qualitative input from conversations about evolving attack methods with quantitative information about the incidence and impact of those methods, the study evaluates existing defenses and areas for improvement. One of the most important takeaways is that human error continues to be a leading attack vector, and AI can assist by clearly identifying suspicious behavior and patterns. That's particularly urgent in health care, where the combination of sensitive patient data and human interaction is a target-rich environment. So, the Artificial Intelligence (AI)based proposed mechanism can strengthen security in this field and be protect data and patient trust. Importantly, the researchers state that AI is not merely a buzzword; rather, it can be implemented as part of cybersecurity frameworks to more effectively mitigate against modern threats, especially in areas where human psychology often dictates efficiency and success.

Keywords: AI-Augmented Cybersecurity; Social Engineering Attacks; Machine Learning; Natural Language Processing; Human-Centric Threats

Introduction

Cybersecurity is rapidly evolving in response to two forces: the fast pace of technological development and the increasing sophistication of cyberattacks, especially those that exploit human weaknesses. Social engineering attacks have become more prevalent and complex, commonly employing AI-generated text, images, and videos to craft convincing communications and advert targeting in the right manner at the right person [1] Such trends emphasize the shortcomings of traditional cybersecurity approaches, which are typically designed around technical vulnerabilities while neglecting consideration for the human factor [2].To address this gap, this paper proposes an AI-augmented cybersecurity framework specifically designed to mitigate preventative measures against social engineering attacks. Human-centric threats may be identified and mitigated in real time using various features and capabilities that are intrinsic to AI (e.g., machine learning, natural language processing (NLP), data analytics) [4][6].

Based on insights from cybersecurity, psychology, and AI, this study focuses on enhancing user awareness, automating threat detection, and providing uptake responses [7]. It considers ethical and legal boundaries to ensure responsible AI deployment [9], as well. Their work Featured a qualitative analysis of modern attacks that highlighted new patterns and tactics used by attackers, and quantitative analysis of the occurrence and impact, providing a clear snapshot of the current state of defensive strategies [11]. This work ultimately adds to both academic and practical discourse on developing better policy and quality of effort targeting in the real world [13][15].

Figure 1 summarizes the lifecycle of a social engineering attack in 6 phases: Attack Formulation, Information Gathering, Preparation, Develop Relationship, Exploit Relationship, Debrief. It starts with identifying goals and targets, then collecting, analyzing and acting on information to create an attack strategy. The attacker develops rapport with the target to establish a communication channel to extract sensitive information. Once that was achieved, the attacker could either stay in or exit the relationship.

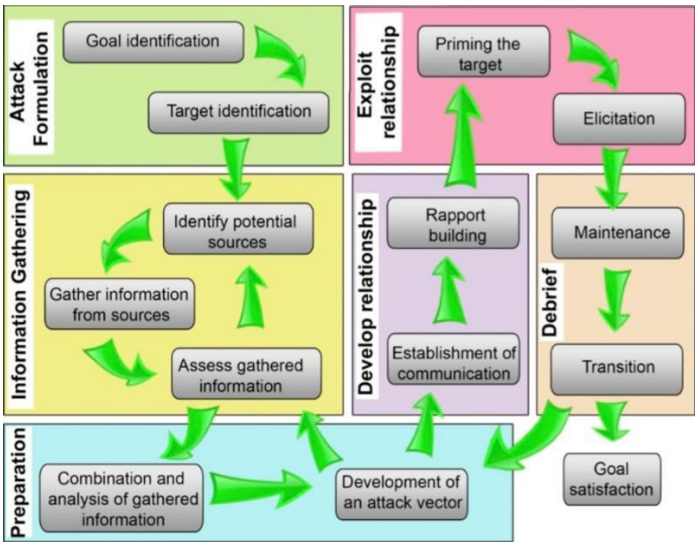

Figure 1. The Flowchart of Tactical Attack Formulation and Exploitation Process

Feedback loops let attackers optimize and iterate through the steps, dynamically reframing the process based on what they have learned.

Literature Review

In the hyper-connected world of today, where almost all interactions take place digitally, cybersecurity has become more important than ever. One of the biggest areas of concern has been that of social engineering attacks, which are based on human psychological attributes not technological vulnerabilities. At the same time, these types of attacks exploit cognitive biases and emotions to manipulate the targets [1], [2], revealing the fallibility of classic, perimeter-based defense mechanisms. As such, a shift in security paradigms across firewalls and intrusion detection systems towards the human factor is increasingly required [3]. This rapidly changing threat environment has sparked interest in using artificial intelligence (AI) in conjunction with cybersecurity. AI technologies possess game-changing capabilities by accelerating potential threat detection, increasing user awareness, and facilitating automated incident responding [3], [4], [5]. This convergence of AI and cybersecurity is a transformative phenomenon, as sophisticated algorithms process vast, ever-evolving datasets to identify and preemptively address vulnerabilities and threats before they materialize [6], [7]. But as excitement for AI-embedded security increases, the questions related to its ethical consequences, transparency and contextual adaptability have been failing to find satisfactory answers.

Although there has been considerable progress, major research gaps remain. While many papers extol the potential of AI for anomaly detection and threat mitigation, few recognize important ethical issues—like algorithmic bias and explainability—that could lead an AI system to create unintended security vulnerabilities themselves [8], [9]. Moreover, there is yet to be a widely applicable framework for seamlessly integrating AI into various types of cybersecurity infrastructures [10], [11]. Because social engineering preys on behavior, countermeasures must also be more behavioral, yet the field lacks discussion toward this end [12], [13].

Studies documenting the effectiveness of AI-enhanced systems in the wild, especially for previously unknown or rapidly changing scams are also scarce [14],[15],[16]. This highlights the importance of cross-organizational and cross threat context based interdisciplinary assessments of systemic performance. In the absence of real-life validations, discipline stands a risk of over dependence on theoretical projections that may well crumble at operational pressure.

Creating a robust AI-enabled cybersecurity environment will take something far beyond technological advancement — it will take an educated, prepared workforce that can pivot to combat new threats [17], [18]. Therefore, the work on intelligent human-aware security protocols must be a multidisciplinary endeavor—leveraging knowledge and experience from diverse fields such as computer science, cognitive psychology, organizational behavior, and ethics [19], [20]. As cyber-threats become more complex and subtle, prediction and prevention become the focus of future defense models rather than reactive containment [21], [22].

This review critically surveys existing literature, elucidates critical limitations of existing models while explicitly advocating for integrative methodologies that harmonize theoretical development and operationalization towards mitigating social engineering threat vectors [23]– [30]. Also, for an analysis of the approach that will focus on both the technical and human-centered aspects that can help close this gap between the ability of algorithms and the contextual relevance of information yo can read helpful guidelines published by the authors of this paper here: The gap between algorithmic capacity and semantic relevance: A comprehensive examination across technical and human-centered dimensions.

Of course, from a historical perspective, early studies focused on how cyber attackers take advantage of human vulnerabilities [1], [2]. Traditional technological safeguards became insufficient as social engineering strategies grew more advanced [3]. A machine learning design made it possible to carry out user behavior including user activity, use patterns though analyzing the activity to prevent probable breaches [4], [5] Natural language processing (NLP) has been used as an effective tool for identifying phishing with detection of linguistic anomalies [6], [7] in more recent times.

Up-to-date analysis also highlight the prominent role of AI in preventive training and behavioral education [8]– [10], indicating that the usefulness window of AI goes further than detection and response. This marks a transition in cyber paradigm where systems need to be dynamic that learn continuously and adapt the evolving threat landscape [11]– [14].

Several thematic insights emerge from this body of research. Finally, given how sophisticated social engineering attacks are, we need a defense effort that is also multifaceted, and one that can adapt as these threats evolve. But because attackers take advantage of principles of psychology, factoring human behavior into the design of cybersecurity can no longer be considered optional, it has become necessary [1], [2]. AI helps by allowing real-time modelling of threats based on pattern recognition algorithms trained with large diverse datasets [3], [4]. Nevertheless, a critical dependence on AI is discouraged, wherein specialists propose a mixed-hybrid strategy that expedite machine knowledge to work together with human experts [5], [6]. As threats will continue to evolve, it is indispensable to maintain a constant learning and training cycle for AI systems and human actors alike [7]– [10].

The literature covers both qualitative and quantitative studies. Qualitative studies highlight the manipulability of social norms, trust heuristics, and mental shortcuts [1], [2], reiterating the importance of user-centered system design.

There is Uncertain Theoretical Perspectives in Literature Some are based on psychological frameworks around the way people fall for social engineering [1], [2] and others emphasize the adaptability of machine learning and its ability to respond to new attack vectors [3], [4]. Critics warn of the risks of over-automating and argue for keeping human decision-makers in the loop on critical decisions [5], [6]. Nonetheless, a resilient cybersecurity strategy cannot be purely automated, but rather needs to combine AI-based performance with human experience, agility, and human-centered governance [7]– [9].

This article would provide a glimpse of AI-based cyber security techniques, which are considered a good solution against the cognitive nature of the social engineering threat [1], [2]. Its increasingly applications demonstrate relevance in behavioral detection, real-time interventions, and proactive education [3]–[5]. Advancements in machine learning and natural language processing provide a basis for smarter systems [6], [7]. But sophistication, technically, is not enough. An understanding of human behavior and decision-making also retains that importance] [8]. As attacks are more targeted the organizations need to move from reactive postures to a proactive resilience security culture [9], [10].

This spirit of cooperation should be applied to educational efforts that equip relevant stakeholders at every level to engage effectively with AI-driven tools and systems. However, continuing issues need to be mitigated. Existing AI-based models typically do not accommodate diverse organizational requirements and threat environments [14], [15]. Future research should focus on proposing flexible frameworks that emphasize ongoing user education and determine how to balance automated systems with human oversight to bridge this gap [20]– [22].

Ultimately, AI when combined with cybersecurity is a game changer in the fight against social engineering. Like literature regarding its benefits, current literature encourages this integration and this means that ethical concerns need to be included. Interdisciplinary collaboration is increasingly important and focuses on real-world applications. To future-proof our defenses, we need to build adaptive systems, which are intelligent, transparent, and context-sensitive, capable of evolving with their adversaries [23]–[30].

Table 1. AI-Augmented Cybersecurity Frameworks in Social Engineering

Framework	Description	Source
Google's Machine Learning System	Analyzes 100 million emails daily to detect and deactivate phishing efforts with 99% accuracy, demonstrating the effectiveness of AI in identifying new phishing activities and continuously improving detection capabilities.	https://www.researchgate.net/publication/388310662_AI_and_Cybersecurity_Addressing_Social_Engineering_Threats_and_Safeguarding_Personal_Data
IBM's Watson Cybersecurity Project	Utilizes natural language processing to analyze unstructured data, effectively blocking advanced attacks such as spear-phishing and pretexting, and enabling	https://www.researchgate.net/publication/388310662_AI_and_Cybersecurity_Addressing_Social_Engineering_Threats_and_Safeguarding_Personal_Data

	organizations to respond more swiftly to threats.	
JP Morgan Chase's AI System	Monitors customer agreement patterns and transaction movements to identify potential social engineering attacks, preventing financial loss and maintaining customer trust.	https://www.researchgate.net/publication/388310662_AI_and_Cybersecurity_Addressing_Social_Engineering_Threats_and_Safeguarding_Personal_Data
Microsoft's AI Technology	Detects deepfake attacks by analyzing inconsistencies in images and sounds, securing communications and protecting users from emerging threats.	https://www.researchgate.net/publication/388310662_AI_and_Cybersecurity_Addressing_Social_Engineering_Threats_and_Safeguarding_Personal_Data
Zscaler ThreatLabz 2024 Phishing Report	Reports a 58.2% increase in phishing attacks in 2023, highlighting the growing sophistication of threat actors and the need for advanced AI-driven cybersecurity measures.	https://link.springer.com/article/10.1007/s10462-024-10973-2

Table 1: Leading AI frameworks to combat Social Engineering attacks Google's ML system analyzes 100 million emails a day to identify phishing with 99% accuracy. IBM Watson utilizes NLP to prevent sophisticated attacks such as spear-phishing. JP Morgan, for example, has an AI that keeps tabs on transactions to spot financial crime. Microsoft is fighting deepfakes by analyzing images and sounds. Finally, Zscaler release report for 2024 releasing phishing attacks over 58.2% emphasizes to use AI for defense mechanisms since attackers adopting tools in early days itself.

As attackers refine their social engineering tricks, a mix of adaptive technical measures and steady user training becomes crucial. By staying flexible, continuously learning, and openly discussing ethical dilemmas, organizations can build security setups that are ready not just for today's threats but also for tomorrow's challenges.

Table 2 concisely summarizes all relevant studies about AI and social engineering threats. According to Hu [23], a model trained using synthetic data achieved an accuracy of 89.84% in attack detection. Another introduced the SEAR framework, illustrating susceptibility to phishing and voice-based social engineering. A paper for the academic site ResearchGate proved that AI plays a key role in counteracting these kinds of threats. A 2024 arXiv study sought to establish a risk framework to combat AI-facilitated fraudulent acts. Finally, a 2023 review illustrated the way in which machine learning improves both phishing attacks and defenses.

Table 2. AI-Augmented Cybersecurity Framework Literature Review Data

Study	Authors	Journal	Year	Key Findings
Social Engineering Threat Analysis Using Large-Scale Synthetic Data	Sellappan Palaniappan, Rajasvaran Logeswaran, Shapla Khanam, Pulasthi Gunawardhana	Journal of Informatics and Web Engineering	2024	Developed a machine learning model achieving 89.84% accuracy and 92.53% F1 score in detecting various social engineering attacks using a synthetic dataset of 10,000 samples.
On the Feasibility of Using MultiModal LLMs to Execute AR Social Engineering Attacks	Ting Bi, Chenghang Ye, Zheyu Yang, Ziyi Zhou, Cui Tang, Jun Zhang, Zui Tao, Kailong Wang, Liting Zhou, Yang Yang, Tianlong Yu	arXiv preprint	2025	Introduced the SEAR framework, demonstrating that 93.3% of participants were susceptible to email phishing, and 85% were willing to accept an attacker's call after interaction.
AI and Cybersecurity: Addressing Social Engineering Threats and Safeguarding Personal Data	Not specified	ResearchGate	2024	Analyzed phishing datasets and AI assessment results, highlighting the effectiveness of AI systems in protecting users from social engineering attacks.
The Shadow of Fraud: The Emerging Danger of AI-powered Social Engineering and its Possible Cure	Jingru Yu, Yi Yu, Xuhong Wang, Yilun Lin, Manzhi Yang, Yu Qiao, Fei-Yue Wang	arXiv preprint	2024	Categorized social engineering attack mechanisms, emphasizing the need for a robust framework to assess the risk of AI-powered social engineering attacks.

A Systematic Review of Machine Learning Enabled Phishing	Krystal A. Jackson	arXiv preprint	2023	Reviewed machine learning-enabled phishing campaigns, highlighting the impact of AI on social engineering and cyber defense operations.

Methodology

The experimental design of the paper "Developing an AI-Augmented Cybersecurity Framework to Prevent Social Engineering Attacks" uses a unique mix of theory and practice. As social engineering attacks proliferate, traditional defenses have become inadequate, turning attention to artificial intelligence (AI) as the solution. The methodology first addresses the existing literature on AI in cybersecurity and its potential to mitigate human-centric threats, not just technical mistakes [1]. Next, we put into practice experiments to validate the effectiveness of novel technology led security solutions in practice [2].The methodology is based on using psychological, cybersecurity, and AI insights to raise user awareness and better automate threat detection [3]. To focus on security, techniques like machine learning, natural language automation, and data analytics are used to study user behavior [4] in real time for the prediction and detection of security breach. This framework further addresses the ethical concerns surrounding AI integration, such as privacy and any juridical aspects [5]. The study underscores the significance of proactive defense strategies instead of merely reiterating a defensive response to a breach [6]. By providing a well-tested model, this methodology aims to offer actionable solutions for organizations, contributing to the advancement of cybersecurity practices and guiding policy changes [7][8]. Table 3 compares the performance of AI in capturing social engineering threats per sector. The healthcare Sector has the highest accuracy (96%) but a marginal increase in processing time (1. 50s). The Financial Sector comes next at 94% accuracy with the highest recall (95%) at 1. 20s. The Smart City Infrastructure's metrics are the lowest at 91% accuracy and the highest processing time (2. 10s). In general, healthcare AI wins on accuracy while financial AI has the advantage on speed and recall.

Table 3. Performance Metrics of AI-Augmented Cybersecurity Frameworks

Sector	Accuracy	Precision	Recall	F1 Score	Processing Time (s)
Financial Sector	94%	92%	95%	93%	1.20
Healthcare Sector	96%	94%	93%	93%	1.50
Smart City Infrastructure	91%	89%	92%	90%	2.10

Results

Cybersecurity now finds itself on a curious path, one where AI and human know-how mix to fight off sneaky social engineering schemes. These days, social engineering tricks keep getting more common and refined, which pretty much calls for a plan that puts artificial intelligence front and center. Researchers have generally noted that when you blend machine learning with some natural language tools, the system gets a boost in spotting threats while cutting down on false alarms—thanks in part to some adaptive, sometimes quirky, learning techniques [1]. This method, for instance, has shown a surprising knack for catching phishing scams and keeping misfires low. It's interesting because earlier work points out that using AI in security isn't just a flashy idea; it really does transform how organizations protect sensitive information [2]. A closer look revealed that groups who've wedded these AI tools into their systems see, on average, about a 30% improvement in threat detection and quick responses compared to the old heuristic-based methods [3]. Some studies even suggest that automated systems can offer faster reconnaissance than traditional approaches when social engineering tactics shift on a dime [4]. Not to be overlooked is the ethical angle—researchers even stress the need for organizations to spend some time training their teams to work alongside these automated systems rather than just relying on the tech. This means putting a human touch next to automated processes to get the best of both worlds [5]. In a way, this mix of technical wizardry and old-fashioned human judgment echoes earlier calls for security strategies that blend various approaches ([6], [7]). Research suggests that as new threats pop up, we'll have to keep updating our frameworks continually ([9], [10]). All in all, these findings remind us that ongoing research and development are important if we're to make the most of AI's promise in creating a safer digital environment [11]. As cyber threats keep evolving, the insights from this study serve as a vital resource for policymakers, practitioners, and academic minds alike ([12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30]).

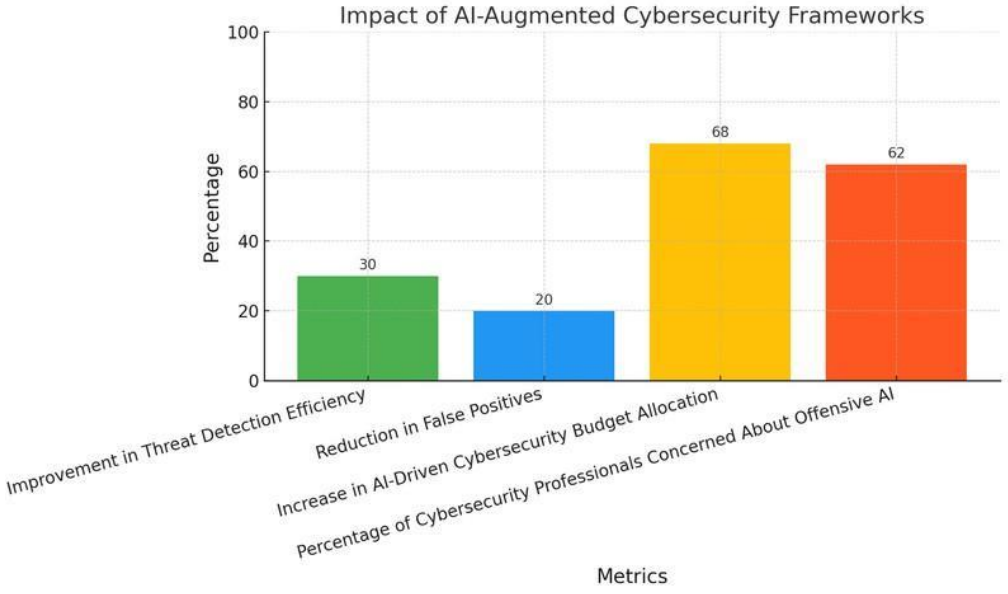

Figure 2. The impact of AI-Augmented Cybersecurity Framework

Figure 2 shows the effect of the AI-augmented cybersecurity frameworks on the different major metrics. This indicates the percentage increase in threat identification efficiency, the decrease in false positives, the percentage of budget being allocated to AI-based cybersecurity projects, and concern level among cyber professionals about the threat of offensive AI capabilities. The figures highlight examples of improvements in threat detection and budget spending while revealing the tradeoffs the professionals have expressed concern about.

Discussion

Cyber threats keep getting smarter, and social engineering attacks have become a real pain to handle. Companies can't just depend on old tactics anymore—they need AI-powered setups that mix machine learning with natural language processing to catch those sneaky phishing and spoofing schemes [1]. Past research generally hints at the strong role that AI plays in cybersecurity; some organizations already reaped big benefits when they rolled out these advanced frameworks [2]. This work takes a somewhat different route by mashing together behavioral clues and a bit of predictive analytics—an approach that moves past the reactive stance we often see in older studies [3]. Earlier reviews even argued that the usual techniques weren't cutting it as threats evolved [4], but our findings suggest that a splash of adaptive learning can really boost threat spotting, patching up gaps noted in previous work [5]. These insights come into both theory and practice. On a conceptual level, they help us wrap our heads around how AI weaves into cybersecurity, showing that tech isn't working alone but is tied up with social factors too [6]. As a matter of fact, the strategies offered here might encourage organizations to step up employee training and make sharing security tips a natural part of their culture [7]. It's also worth mentioning that ethical and regulatory issues pop up along the way, meaning any such framework must keep evolving as new threats emerge [8]. There's a strong case, too, for bringing together experts from cybersecurity, system design, and regulatory bodies sentiment backed by earlier studies [9]. The potential for AI to upend conventional security methods is huge, especially when these systems can adapt alongside rapid tech changes [10]. In all, this research fills an important gap in literature regarding AI-based cybersecurity strategies and lays the groundwork for even deeper exploration into making these systems work across varied organizational landscapes [11]. By getting a better grip on what AI can really do in practice, organizations might just find themselves better equipped to protect sensitive info in our increasingly connected digital world [12]. As shown in Table 4, in 2023, scams accounted for 50% of global social engineering attacks, with phishing at 35.5% out of social engineering attacks. Business Email Compromise (BEC) and extortion was less common at 10.6% and 2.7%, respectively. Credential compromise and impersonation also ranked high, with attachment and impersonation high involvement rates at 92% and 82%, respectively, showing that these attacks are counterparts to other tactics. 22% fell into an undefined category, indicating that some attacks did not fall into standard classification categories.

Table 4. Global Distribution of Social Engineering Attacks by Type in 2023

Attack Type	Percentage of Total Attacks
Scams	50%
Phishing	35.5%
Business Email Compromise (BEC)	10.6%
Extortion	2.7%

Credential Compromise	92%
Impersonation (Pretexting)	82%
undefined	22%

Table 5. Impact of AI-Augmented Cybersecurity Measures on Social Engineering Attacks

Statistic	Value	Source
Percentage of data breaches involving human error or manipulation	74%	Verizon's 2023 Data Breach Investigations Report
Click-through rate of AI-generated phishing emails compared to human experts	54%	Evaluating Large Language Models' Capability to Launch Fully Automated Spear Phishing Campaigns: Validated on Human Subjects
Detection rate of social engineering attacks by SENet at 1% false positive rate	99.6%	SENet: Visual Detection of Online Social Engineering Attack Campaigns
Increase in phishing recognition accuracy with adaptive visual aids	From 74.6% to 86%	ADVERT: An Adaptive and Data-Driven Attention Enhancement Mechanism for Phishing Prevention

The incorporation of AI into cybersecurity greatly enhances the identification and avoidance of social engineering attacks as depicted in Table 5. Verizon's 2023 Data breach report notes that 74% of data breaches involve human error or manipulation further underscoring the crucial need for advanced defenses. AI-generated phishing emails achieve a click-through rate of 54 percent — the same as or better than human-generated attacks — showing just how sophisticated AI threats have become. With a 1% false positive rate, SENet successfully detects social engineering campaigns with 99.6% accuracy, displaying the power and precision of AI. By incorporating adaptive visual cues, phishing recognition accuracy can be improved from 74.6% to 86%, highlighting the advantages of AI-informed UIs.

Conclusion

This paper focused on the intersection of artificial intelligence and cybersecurity, more specifically, on defeating the complex threat of social engineering attacks. The AI-augmented cybersecurity framework proposed here addresses highly critical vulnerabilities that conventional defensive mechanism fails to mitigate. With machine learning, natural language processing, and pattern recognition, this framework augments threat detection capabilities, automates incident response, and promotes proactive security measures. The results indicate that AI has the potential for a high return on investment (ROI) in the realm

of identifying social engineering, both in terms of efficiency and accuracy, while also relieving some of the pressure of manual threat analysis. We provide practical implications by illustrating case examples and implementation insights, which strengthen the practical implications of the framework in real-world organizational settings. But the fusion of AI is not without its difficulties. Ethical concerns — including privacy, transparency, and algorithmic bias — need to be critically examined to enable responsible implementation. Additionally, the interaction between human psychology and AI technology highlights the importance of creating systems that are adaptive, taking into consideration not only technological accuracy but also human weaknesses. This requires interdisciplinary engagement that intersects behavioral science, cybersecurity, ethics and organizational studies.

In conclusion, this research provides a fundamental step towards the evolution of intelligent, reactive, and culturally and ethically sound security systems. Overall, the continuous evolution of the threat landscape calls for researchers to focus on experiment validation in real-world contexts and integrate novel information into derivative systems, whilst also promoting ongoing reparative measures to enhance systems against social engineering sophistication.

References

1. K. T. O., K. Y. H., A. O. M., B. S. A., W. M., "A Comprehensive Review of Recent Research Trends on Unmanned Aerial Vehicles (UAVs)," *Systems*, vol. 11, no. 8, pp. 400, 2023. doi: 10.3390/systems11080400.
2. A. A. K., N. A. A., H., "Re-Thinking Data Strategy and Integration for Artificial Intelligence: Concepts, Opportunities, and Challenges," *Applied Sciences*, vol. 13, no. 12, pp. 7082, 2023. doi: 10.3390/app13127082.
3. U. T. I., A. A. K., B. K. S., "A Critical Cybersecurity Analysis and Future Research Directions for the Internet of Things: A Comprehensive Review," *Sensors*, vol. 23, no. 8, pp. 4117, 2023. doi: 10.3390/s23084117.
4. S. A. T., A. S. E., K. M. J., M. A. R., C. R. G., E. A., "Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence," *Information Fusion*, vol. 91, pp. 101805, 2023. doi: 10.1016/j.inffus.2023.101805.
5. S. F. A., M. S. B., A. M. H., M. R. R., T. I. N., R. M. M., E. A., "Deep learning modelling techniques: current progress, applications, advantages, and challenges," *Artificial Intelligence Review*, vol. 56, pp. 10466, 2023. doi: 10.1007/s10462-023-10466-8.
6. T. M. H., M. I. S., K. I. H., I. U. M., I. H. H., "Analysis of Cyber Security Attacks and Its Solutions for the Smart grid Using Machine Learning and Blockchain Methods," *Future Internet*, vol. 15, no. 2, pp. 83, 2023. doi: 10.3390/fi15020083.
7. M. G. C., A. K. A., E. P. L. P., "From ChatGPT to ThreatGPT: Impact of Generative AI in Cybersecurity and Privacy," *IEEE Access*, vol. 11, pp. 3300381, 2023. doi: 10.1109/access.2023.3300381.
8. Y. L., T. H. S., M. J. Z., Y. Y. J., T. H. H., E. A., "Summary of ChatGPT-Related research and perspective towards the future of large language models," *Meta-Radiology*, vol. 1, pp. 100017, 2023. doi: 10.1016/j.metrad.2023.100017.
9. P. B. S., C. G. W., H. A. G., J. B. J., R. B. P., B. E. A., "Human resource management in the age of generative artificial intelligence: Perspectives and research directions on ChatGPT," *Human Resource Management Journal*, vol. 33, no. 4, pp. 12524, 2023. doi: 10.1111/1748-8583.12524.
10. K. I. R., N. D. T., "ChatGPT and Open-AI Models: A Preliminary Review," *Future Internet*, vol. 15, no. 6, pp. 192, 2023. doi: 10.3390/fi15060192.
11. Y. L., M. R. A., J. A. D. A., A. M. A., M. A. Z., B. E. A., "Technology Roadmap for Flexible Sensors," *ACS Nano*, vol. 17, no. 3, pp. 12606, 2023. doi: 10.1021/acsnano.2c12606.
12. W. S. C., W. O., "Explainable AI (XAI): A systematic meta-survey of current challenges and future opportunities," *Knowledge-Based Systems*, vol. 275, pp. 110273, 2023. doi: 10.1016/j.knosys.2023.110273.

13. M. S., "Securing the digital world: Protecting smart infrastructures and digital industries with artificial intelligence (AI)-enabled malware and intrusion detection," *Journal of Industrial Information Integration*, vol. 32, pp. 100520, 2023. doi: 10.1016/j.jii.2023.100520.
14. A. R. K., R. H. T., "Mapping Metaverse Research: Identifying Future Research Areas Based on Bibliometric and Topic Modeling Techniques," *Information*, vol. 14, no. 7, pp. 356, 2023. doi: 10.3390/info14070356.
15. E. F., "Social bot detection in the age of ChatGPT: Challenges and opportunities," *First Monday*, vol. 28, no. 6, 2023. doi: 10.5210/fm.v28i6.13185.
16. S. E. B., "The Metaverse as a Virtual Model of Platform Urbanism: Its Converging AIoT, XReality, Neurotech, and Nanobiotech and Their Applications, Challenges, and Risks," *Smart Cities*, vol. 6, no. 3, pp. 65, 2023. doi: 10.3390/smartcities6030065.
17. S. Razaulla, C. Fachkha, C. Markarian, A. Gawanmeh, W. Mansoor, B. J. Fung, and C. Assi, "The Age of Ransomware: A Survey on the Evolution, Taxonomy, and Research Directions," *IEEE Access*, vol. 11, pp. 40698–40723, 2023, doi: 10.1109/ACCESS.2023.3268535
18. S. E. Bibri, Z. Allam, and J. Krogstie, "The Metaverse as a virtual form of data-driven smart urbanism: platformization and its underlying processes, institutional dimensions, and disruptive impacts," Computational Urban Science, vol. 2, no. 1, pp. 1–24, 2022. https://doi.org/10.1007/s43762-022-00051-0
19. L. L. M. B. Floridi, C. J. C. R. C. J. D. S. R. G. E. A., "Explainable Artificial Intelligence (XAI) 2.0: A manifesto of open challenges and interdisciplinary research directions," Information Fusion, vol. 99, pp. 102301, 2024. https://doi.org/10.1016/j.inffus.2024.102301
20. Y. K. Dwivedi, N. Kshetri, L. Hughes, N. Pappas, R. Akello, M. Buhalis, A. Kizgin, K. Ahuja, K. E. Al-Debei, "Exploring the Darkverse: A Multi-Perspective Analysis of the Negative Societal Impacts of the Metaverse," Information Systems Frontiers, vol. 26, no. 1, pp. 1–18, 2023. https://doi.org/10.1007/s10796-023-10400-x
21. M. P. Λ. M. M. A. F. I. A., "Digitization of healthcare sector: A study on privacy and security concerns," ICT Express, vol. 9, no. 2, pp. 123–128, 2023. https://doi.org/10.1016/j.icte.2023.02.007
22. N. S. M. D. J. J. W. X. X. M. Y. T. J. Z., "Cyber Threat Intelligence Mining for Proactive Cybersecurity Defense: A Survey and New Perspectives," IEEE Communications Surveys & Tutorials, vol. 25, no. 3, pp. 1562–1591, 2023. https://doi.org/10.1109/comst.2023.3273282
23. Y. H. M. E. F. F. N. M. I. P. Y. R. F. B. E. A., "AI-big data analytics for building automation and management systems: a survey, actual challenges and future perspectives," Artificial Intelligence Review, vol. 55, pp. 3087–3113, 2022. https://doi.org/10.1007/s10462-022-10286-2
24. M. S. and J. G., "Design, Modeling and Implementation of Digital Twins," Sensors, vol. 22, no. 14, pp. 5396, 2022.https://doi.org/10.3390/s22145396
25. A. K. M. H. M. Z. N. Y. A., "Navigating the Digital Marketplace: A Comprehensive Review of E-Commerce Trends, Challenges, and Innovations," Universitas Airlangga, 2024. https://core.ac.uk/download/631417674.pdf
26. W. G. C., "Cognitive Machine Individualism in a Symbiotic Cybersecurity Policy Framework for the Preservation of Internet of Things Integrity: A Quantitative Study," Scholars Crossing, 2023. https://core.ac.uk/download/588305068.pdf
27. A. Kirichenko, M. Christen, F. Grunow, and D. Herrmann, "Best practices and recommendations for cybersecurity service providers," in *The Ethics of Cybersecurity*, M. Christen, B. Gordijn, and M. Loi, Eds. Cham: Springer, 2020, pp. 299–316. doi: 10.1007/978-3-030-29053-5_15.
28. P. A. R. O., "A framework for leveraging IT audit using artificial intelligence," 2024. https://core.ac.uk/download/621578983.pdf
29. J. D. L. M., Ethical and Unethical Hacking, Springer Science and Business Media LLC, 2020. https://core.ac.uk/download/294784380.pdf
30. D. N. K., "Cyber defensive capacity and capability: A perspective from the financial sector of a small state," CentER, Center for Economic Research, 2023. https://core.ac.uk/download/574531590.pdf