Using Generative AI for Cybersecurity Awareness Training in Healthcare

Dmitry Zhdanov[1]
School of Information Technology
Illinois State University
Normal, IL, USA

Thomas M. Caldera
Digital Innovation Development
OSF HealthCare
Peoria, IL, USA

Mary Elaine Califf
School of Information Technology
Illinois State University
Normal, IL, USA

ABSTRACT

Generative AI (GAI) tools have recently become available for broad public use, and the use of AI-augmented systems is on the rise. Simultaneously, the healthcare sector is experiencing a notable increase in cyberattacks, resulting in the need to ramp up cybersecurity awareness among healthcare personnel. This industry-university partnership aims to improve cybersecurity training effectiveness and outcomes in a regional hospital system in the USA.

Our proposed solution is to build and evaluate a GAI-driven training environment that could be used in cybersecurity awareness training (GAI CAT). Using the baseline models of the AI industry and augmenting those with healthcare-focused cybersecurity information, we aim to make cybersecurity awareness more effective, interactive, and user-driven. Differences in knowledge retention and engagement with training will be measured and reported in a hospital setting.

Keywords: Cybersecurity; training and awareness; generative artificial intelligence; design science; healthcare.

[1] Corresponding author. dzhdano@ilstu.edu +1 309 438 3240

INTRODUCTION

This project is an industry-academia collaboration between a public research university and a regional hospital system in the USA. The project's innovative core centers around the strategic incorporation of Generative AI (GAI) within cybersecurity awareness training – overall and in healthcare specifically.

Generative AI (GAI) tools have recently become available for broad public use. Well-known examples include ChatGPT for text generation and MidJourney for image creation. While having a solid potential to augment and transform many human activities, GAI also comes with various concerns – from replacing humans in some occupations to copyright infringement to the potential use of GAI for criminal activities. Recent examples include "evil twins" of ChatGPT such as WormGPT and EvilGPT which can be used to generate phishing campaigns and malware code (Dutta, 2023). The growing consensus though is that the best uses of GAI are in the areas where human activity can be augmented with GAI to resolve complexity while keeping human agency and decision-making authority in place (see, for example, Nah et al., 2023). We propose that cybersecurity awareness training is one of those activities with a significant potential for augmentation by GAI.

Simultaneously, the healthcare sector is being increasingly and repeatedly targeted for cyberattacks, including ransomware; cyberattacks against hospitals has doubled in 2023 (Office of the Director of National Intelligence, 2024). Considering the potential for healthcare professionals to inadvertently expose sensitive data in Generative AI models, adhering to HIPAA and patient privacy policies is imperative. Therefore, situating this study in a healthcare setting is both timely and important.

Some recommendations to improve Cybersecurity Awareness Training (CAT) include 1) breaking it down into smaller but ongoing tasks that can be incorporated into the daily routines (EC-Council, 2020), and 2) gamifying CAT to make the users more engaged (Oroszi ,2020). We believe that GAI can help in both of these ways.

Our proposed solution is to build and evaluate a GAI-driven training environment which could then be used in CAT. The environment would use a chatbot platform wherein users interact with a human-like GAI tool. There are two advantages to using this chatbot in security awareness training: 1) CAT becomes more interactive and engaging, and 2) CAT becomes user-driven and can be used on demand. We develop and test a tool called GAI-CAT in a healthcare setting. Our hypotheses and approach are described in the subsequent sections.

BACKGROUND AND HYPOTHESES

Cybersecurity awareness training (CAT) is an important element of overall cybersecurity programs in organizations of all sizes and industries. However, it is nontrivial to make this training valuable and effective. Common issues with CAT programs include the following:

- CAT is conducted once or twice a year, while for the rest of the year, employees do not think about cybersecurity issues (Hutchinson, 2021)
- CAT is usually passive – it involves watching videos or reading short scenarios followed by a quiz. Jumping to the quiz is an optimal strategy to "pass" for many participants.
- Cybersecurity is an abstract and secondary task to those not working directly in the cybersecurity function (West, 2008). The abstract element is manifest without a clear action-feedback mechanism (i.e., if I comply with security policies, at best nothing happens. If I do not comply, the adverse effects are a probability and may not impact me directly). The secondary task element becomes evident when people bypass security

mechanisms to complete their primary tasks (e.g., "I need to finish this report, thus I will copy data to an unprotected device")

We believe that GAI can help improve CAT by making it more interactive, user-focused, and embedded in the daily operations of an organization. Our proposed solution is to build and evaluate a GAI-driven training environment which could then be used in CAT.

Our broad conjecture is that the use of a GAI-enhanced tool for CAT will improve security training outcomes. This conjecture can be further operationalized in the following hypotheses:

1. Use of GAI CAT will lead to higher engagement with training.
2. Use of GAI CAT will lead to better retention of training information.
3. Use of GAI CAT will lead to a better distribution of training in time (across the year).
4. Use of GAI CAT will improve the overall cybersecurity posture of the organization (e.g., fewer incidents, faster reporting/discovery, more security-conscious culture)

ARTIFACT DESIGN

Our approach is based on the architecture recommended by OpenAI (Jarvis, 2023). We will provide a cybersecurity-specific corpus of data and create a cybersecurity training application on top of it. GAI CAT artifact will be created and fine-tuned using specialized cybersecurity data and prompt engineering. Sample inputs could include guiding cybersecurity documents from NIST 800 series such as NIST 800-53 (NIST, 2022) or MITRE ATT&CK repositories (MITRE, 2023). Inputs specific to HIPAA and related regulations will also be included (such as HealthIT.gov, 2023). These specific textual embeddings will be added on top of the foundational language models of Open AI.

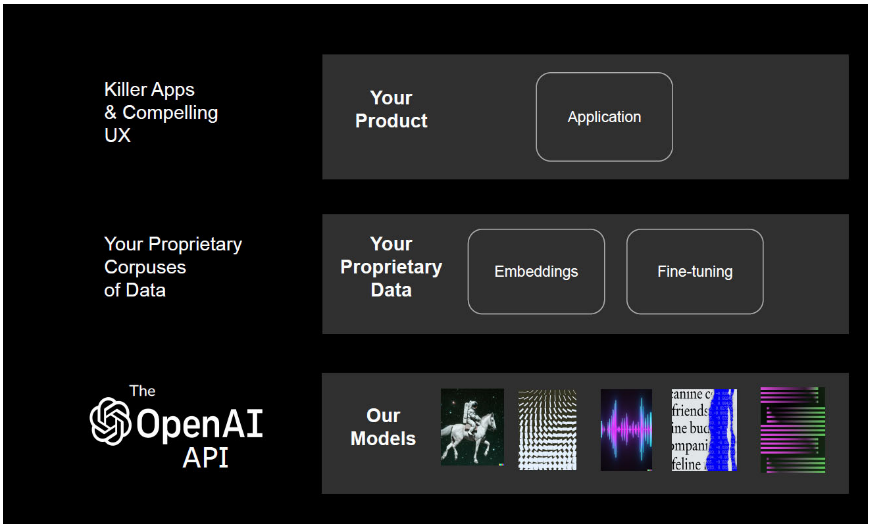

Figure 1. Building Applications on Open AI.

STUDY DESIGN

In order to test the hypotheses, we will perform a baseline assessment of the cybersecurity knowledge of the employees prior to deploying the GAI CAT solution. Knowledge retention will be tested with some delay after the training program is completed. The comparison will be made between those participants who used the GAI CAT tool (treatment group) vs those who did not (control group). The instrument includes two categories of questions: 1) questions regarding a general understanding of cybersecurity issues and policies that are typically included in cybersecurity training (such as password policies; recognizing phishing; and social media security), and 2) questions regarding the GAI CAT experience for the treatment group (feedback regarding ease of use, helpfulness, frequency of interaction and so on). The project workflow is depicted in Figure 2 below.

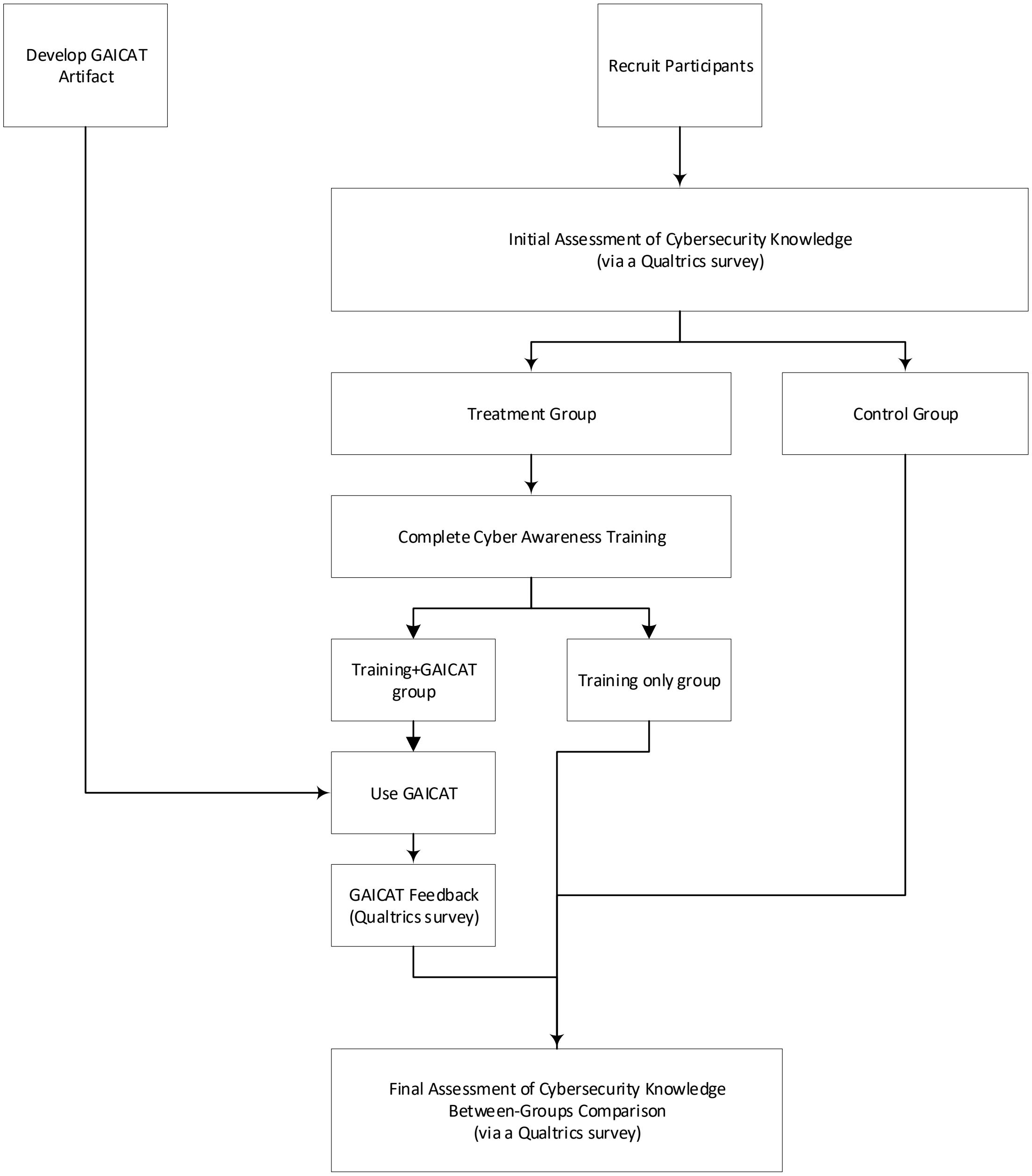
Develop GAICAT Artifact
Recruit Participants
Initial Assessment of Cybersecurity Knowledge (via a Qualtrics survey)
Treatment Group
Control Group
Complete Cyber Awareness Training
Training+GAICAT group
Training only group
Use GAICAT
GAICAT Feedback (Qualtrics survey)
Final Assessment of Cybersecurity Knowledge Between-Groups Comparison (via a Qualtrics survey)

Figure 2. GAI CAT Project Workflow

Study participants will be recruited from the hospital system with an estimated number of ~100 participants. A chance to win a $50 Amazon gift card is provided as an incentive. Participation in the study is voluntary and is not related to employee performance evaluation. This study does not involve the general population.

CURRENT STATUS AND EXPECTED OUTCOMES

As of December 2024, the GAI CAT artifact is developed and ready to deploy. The training platform is undergoing a pilot test. We plan to open the study to the participants in January 2025.

Upon completion, we anticipate that this project will produce significant impacts. It proactively mitigates the risk of sensitive data exposure, ensuring stringent adherence to regulatory standards like HIPAA. Empowering healthcare professionals with the ability to engage in training actively equips them with the knowledge and skills required to safeguard against emerging cyber threats. The study serves as a tangible demonstration of how GAI can augment human activities within the healthcare sector. It showcases the technology's transformative potential in addressing critical aspects, like healthcare cybersecurity.

ACKNOWLEDGEMENTS

This project is supported by a grant from the Connected Communities Initiative, administered by OSF HealthCare and Illinois State University.

REFERENCES

Dutta, T. 2023. “Hackers Released New Black Hat AI Tool Evil-GPT as a Replacement for Worm GPT.” *Cyber Security News*, August 10, 2023, https://cybersecuritynews.com/hackers-released-evil-gpt/ Accessed on 9/5/2023.

EC-Council. 2020. “Security Awareness Training: 6 Important Training Practices”, May 28, https://aware.eccouncil.org/security-awareness-training-6-important-training-practices.html Accessed on 9/6/2023.

HealthIT.gov. 2023. “Health IT Privacy and Security Resources for Providers” https://www.healthit.gov/topic/privacy-security-and-hipaa/health-it-privacy-and-security-resources-providers, Accessed on 9/6/2023.

Hutchinson, E. 2021. “How often should you conduct cybersecurity awareness training?” https://www.cira.ca/en/resources/news/cybersecurity/how-often-should-you-conduct-cybersecurity-awareness-training/ , *Canadian Internet Registration Authority*, October 27. Accessed on 9/6/2023.

Jarvis, C. 2023. "Powering your products with ChatGPT and your own data", *Open AI + Microsoft*, https://drive.google.com/file/d/1dB-RQhZC_Q1iAsHkNNdkqtxxXqYODFYy/view, Accessed on 9/6/2023.

MITRE. 2023. https://attack.mitre.org/ Accessed on 9/5/2023.

Nah, F., R. Zheng, J. Cai, K. Siau & L. Chen. 2023 "Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration", *Journal of Information Technology Case and Application Research*, 25:3, 277-304, DOI: 10.1080/15228053.2023.2233814

NIST. 2022. "Assessing Security and Privacy Controls in Information Systems and Organizations" https://csrc.nist.gov/pubs/sp/800/53/a/r5/final, Accessed on 9/5/2023.

Office of the Director of National Intelligence. 2024. "Ransomware Attacks Surge in 2023; Attacks on Healthcare Sector Nearly Double." https://www.dni.gov/files/CTIIC/documents/products/Ransomware_Attacks_Surge_in_2023.pdf, Accessed on 10/4/2024.

Oroszi, E. 2020. "Using Gamification to Improve the Security Awareness of Users: The Security Awareness Escape Room", *ISACA Journal*, 2020:4, 1-5.

West, R. 2008. "The Psychology of Security", *Communications of the ACM*, 51:4, 34-41.