

Feasibility of AI-Driven Non-Playable Characters (NPCs) as Conversational Guides in Cybersecurity Training Simulations

Chukwuemeka Ihekweazu
cei004@shsu.edu

Pat Ko
ko@shsu.edu

Computer Science Department
Sam Houston State University
Huntsville, TX 77340, USA

Naomi Aghado
naomiaghado@yahoo.com
College of Nursing and Health Department
Viterbo University
La Crosse, WI 54601, USA

Elizabeth Adepeju Adelowo
eaadelowo@stcloudstate.edu
Educational Leadership and Higher Education Department
St. Cloud State University
St Cloud, MN 56301, USA

Abstract

Despite the growing adoption of virtual labs and gamified platforms in cybersecurity training, many of these environments remain static, offering limited opportunities for real-time guidance or adaptive feedback. This paper introduces a novel approach to addressing this gap using AI-driven Non-Playable Characters (NPCs), which are conversational agents embedded within simulated cybersecurity scenarios. These intelligent NPCs are designed to serve as interactive guides, mentors, or adversaries, dynamically responding to learner actions in ways that mimic real-world training experiences.

The proposed framework integrates a domain-specific knowledge base, a conversational engine powered by large language models, and a simulation interface compatible with both browser-based and game-engine platforms. By aligning with standardized cybersecurity frameworks such as NIST NICE, and MITRE ATT&CK, these AI agents can provide accurate, context-sensitive responses that support skill development. Although this study does not involve direct user testing, it offers a detailed design model supported by educational theory, scalability analysis, and use-case examples.

Ultimately, this research lays the foundation for scalable, immersive, and personalized cybersecurity instruction, bridging the gap between passive learning and active, scenario-based training environments. This research contributes to the integration of pedagogical technology by introducing intelligent agents as instructional scaffolds in cybersecurity education.

Keywords: Cybersecurity Education; Conversational AI; Non-Playable Characters; Simulation-Based Training; Intelligent Agents.

Feasibility of AI-Driven Non-Playable Characters (NPCs) as Conversational Guides in Cybersecurity Training Simulations

Chukwuemeka Ihekweazu, Pat Ko, Naomi Aghado, and Elizabeth Adepeju Adelowo

1. INTRODUCTION

Cybersecurity has become a critical area of focus within STEM education due to the growing complexity of cyber threats and the urgent demand for skilled professionals capable of mitigating them. To address this, academic institutions and industry programs have increasingly incorporated simulation-based learning environments, such as virtual labs and gamified platforms, into cybersecurity curricula. These environments provide learners with opportunities to practice essential technical skills, such as threat detection, incident response, and vulnerability remediation, within controlled and risk-free contexts. However, despite their growing adoption, many of these simulations fall short in terms of instructional adaptability, engagement, and pedagogical support (Ahmed & Parsons, 2021; Kim & Reeves, 2007).

A persistent challenge in cybersecurity training lies in the lack of interactive, personalized feedback mechanisms within these environments. Learners often navigate simulations with limited real-time guidance or adaptive instructional support, particularly when instructors are unavailable, or the learning experience is self-paced. This absence of dynamic feedback and engagement hinders cognitive scaffolding, slows skill acquisition, and may result in learner frustration or disengagement (Vygotsky, 1978; Mayer, 2011).

To address this gap, the present study proposes a conceptual framework for integrating AI-driven Non-Playable Characters (NPCs) into cybersecurity training simulations. These NPCs, powered by conversational artificial intelligence, are designed to function as real-time instructional agents embedded in virtual environments. Drawing from constructivist and situated learning theories, these NPCs adopt various instructional roles, mentor, adversary, and incident responder to guide learners through decision-making processes, pose reflective prompts, and provide just-in-time feedback tailored to each user's context and actions.

The proposed framework incorporates instructional design principles, including

scaffolding, formative assessment, and experiential learning. It also aligns with established cybersecurity training standards, including the National Initiative for Cybersecurity Education (NICE) Framework and the MITRE ATT&CK framework (NIST, 2020; MITRE, n.d.). The goal is to create a scalable, interactive training environment that enhances learner engagement and comprehension without requiring human facilitation. The use of intelligent NPCs has the potential to replicate realistic cyber scenarios while supporting individualized learning paths, a crucial advancement in instructional technology and cybersecurity education.

2. PURPOSE OF THE STUDY

The purpose of this study is to design a non-evaluative, conceptual framework for implementing conversational AI-driven NPCs within cybersecurity training simulations. These agents are designed to address the instructional gaps in current platforms by providing dynamic, role-based interactions and real-time guidance.

3. RESEARCH QUESTIONS

The research is guided by the following questions:

1. How can AI-driven NPCs simulate realistic and pedagogically meaningful interactions within cybersecurity training environments?
2. What instructional roles can conversational NPCs perform to support learning across exploration, practice, and reflection?
3. How can instructional design principles be embedded within AI NPCs to enhance engagement and personalization?

4. SCOPE OF STUDY

This study focuses on the development of a framework architecture, sample implementation scenarios, and pedagogical justification grounded in existing literature. The outcome is a proposed model that can serve as a foundation for future system development and empirical studies in

cybersecurity education.

5. LITERATURE

Conversational AI in Education and Training Environments

Recent advancements in natural language processing (NLP) and large language models (LLMs) have enabled the development of conversational agents capable of interacting with learners in natural, context-aware ways. These agents commonly deployed as chatbots or virtual tutors have shown promise in supporting self-directed learning, answering questions, and providing timely feedback (Fryer & Nakao, 2020; Winkler & Söllner, 2018). When conversational agents are designed with pedagogical intent, they can enhance engagement and foster reflective learning in STEM education (Kim, 2019). However, current applications tend to exist outside immersive environments, limiting their potential to support real-world learning contexts like cybersecurity.

In cybersecurity, where learners must acquire both procedural knowledge and situational awareness, static learning resources fall short. Intelligent agents must simulate pedagogically meaningful interactions through adaptive dialogue, where NPCs provide guided prompts and context-sensitive responses tailored to learner decisions designed to support cognitive scaffolding (VanLehn, 2011; Vygotsky, 1978). The shift toward AI NPCs embedded in simulations offers a new opportunity for real-time, contextualized guidance that responds to learner decisions and evolving threat scenarios.

In their paper, Mierzwa et al. (2019) demonstrated that feasibility studies can introduce AI chatbots as viable tools before participant testing, highlighting their potential to improve efficiency and support user interaction. Following this precedent, the current NPC framework should likewise be regarded as a proof-of-concept, illustrating technical and instructional feasibility ahead of large-scale validation.

While conversational AI has seen increased use in general education, a review by Ahmed and Parsons (2021) found that most cybersecurity simulations are static, relying on hard-coded instruction paths and offering limited interactivity. Similarly, Yin, Zhu, and Wang (2022) concluded that adaptive instruction and intelligent agents remain underutilized in cybersecurity education research. These

environments fail to reflect the dynamic, decision-rich nature of real-world cyber defense, highlighting an urgent gap for more context-aware instructional tools.

Non-Playable Characters (NPCs) in Simulations and Game-Based Learning

NPCs have long played important roles in game-based learning by acting as guides, mentors, adversaries, or story-driving characters. In educational contexts, NPCs can help learners practice critical thinking, apply problem-solving strategies, and experience failure in low-risk environments (Gee, 2003; Dede, 2009). However, in cybersecurity education, NPCs are often limited to non-interactive roles or pre-scripted interactions that fail to reflect real-world complexity.

Emerging research in immersive learning environments supports the integration of intelligent NPCs that adapt based on learner behavior and simulate realistic cyber threats (Moser et al., 2021). By assigning NPCs instructional roles across phases of learning exploration, practice, and reflection, they can serve as scaffolds that activate prior knowledge, offer formative feedback, and encourage self-regulation (Kolb, 1984; Herrington & Oliver, 2000).

Platforms such as TryHackMe and RangeForce offer valuable scenario-based exercises. However, unlike the proposed framework, they lack embedded conversational NPCs capable of adaptive, role-based instruction. This highlights the novelty of embedding AI-driven NPCs into simulation-based cyber training. Learners must rely on written walkthroughs, forums, or static hints, which significantly limit immediacy and personalization (Dodge, Ragsdale, & Reynolds, 2020; Yamin, Katt, & Gkioulos, 2020). This design may be suitable for self-paced learners with a technical background, but it provides limited instructional support for novices or those who require adaptive guidance.

In this study, the authors analyzed survey data from 1,597 first-year students across multiple semesters, revealing a rapid increase in the adoption of generative AI tools, such as ChatGPT, for academic tasks. Students primarily used AI for homework and research preparation. The study highlights the importance of data literacy in ensuring equitable and ethical adoption of AI. These findings underscore the rapid penetration of AI-driven tools into educational spaces, underscoring the timeliness of integrating NPCs as conversational AI assistants in cybersecurity

education (Frydenberg, Mentzer, & Patterson, 2026).

Instructional Design Principles Embedded in AI NPCs

Embedding instructional design principles into AI NPCs involves more than enabling them to speak or respond. It requires that their behavior, dialogue, and interactions follow pedagogically intentional models. Models like Gagné's Nine Events of Instruction (Gagné et al., 2005), Merrill's First Principles (Merrill, 2002), and the ARCS Motivation Model (Keller, 2009) provide frameworks for designing learning sequences that capture attention, build relevance, support mastery, and provide feedback.

For example, an NPC acting as a mentor may follow a pattern of instruction by: (1) gaining attention with a compelling opening prompt, (2) informing the learner of objectives, (3) guiding learning with tips or mini-tasks, (4) providing feedback based on learner action, and (5) encouraging reflection at scenario end.

From a personalization standpoint, AI NPCs can leverage learner modeling, where past decisions, task success rates, and interaction history are analyzed to customize instructional content (Wollny et al., 2021). This form of adaptive instruction is particularly relevant in cybersecurity, where learners differ significantly in baseline knowledge and cognitive load tolerance. By embedding Universal Design for Learning (UDL) principles such as offering multiple means of representation and engagement, NPCs can deliver information in varied formats (textual, visual, interactive) and adjust pacing or challenge level based on learner performance (CAST, 2018).

Engagement is further strengthened through affective adaptation, where AI systems monitor and respond to learner emotions using methods from affective computing, such as sentiment analysis and interaction timing. Although empirical testing with participants was not conducted, frustration was conceptualized from prior literature as observable through long response delays, repeated errors, or disengagement. For example, an NPC might adapt by shifting from a high-pressure dialogue style to a more supportive tone. Serholt et al. (2021) emphasize that emotionally aware AI agents can improve persistence and satisfaction in simulation-based learning environments.

Jiang and Nakatani (2025) present an empirical study integrating Generative AI tools (e.g.,

ChatGPT, Gemini) into IS coursework. Their results show that students using GenAI performed as well as or better than those in control groups. Importantly, concerns of academic dishonesty and overreliance did not manifest significantly. Students responded favorably to AI-assisted assignments, reinforcing the feasibility of conversational AI in education.

Most existing literature on educational chatbots and AI tutors focuses on general-purpose systems for answering questions or guiding students through structured content (Chassignol et al., 2018; Wollny et al., 2021). However, there is limited research on NPCs that assume distinct instructional roles, such as adversaries simulating phishing attacks, mentors providing reflective feedback, or incident handlers guiding triage decisions. This type of pedagogically aligned, character-based simulation is especially relevant to cybersecurity but remains largely unexplored in both research and practice.

While prior research highlights the benefits of conversational agents and NPCs in education, few studies integrate them with instructional design models tailored for cybersecurity simulations. There is a significant opportunity to develop AI-driven NPCs that combine situational fidelity with instructional purpose providing personalized, scaffolded, and meaningful interactions throughout the learning experience. In summary, while prior research has explored conversational AI and NPCs in education separately, their integration in cybersecurity training particularly through differentiated instructional roles embedded in real-time simulations remains underdeveloped. This study addresses this critical gap by proposing a unified framework that brings together instructional design principles, conversational AI capabilities, and simulation-based learning in a novel and scalable way.

6. METHODOLOGY

This study introduces a conceptual framework in Figure 1 (in the appendices) for integrating conversational AI-driven non-playable characters (NPCs) into cybersecurity simulations, with a specific focus on instructional design. The framework is grounded in the principles of scaffolding, adaptive feedback, and contextualized learning, and is designed to be applicable across web-based or immersive platforms.

At its core, the framework consists of two role-differentiated AI agents developed using the Convai platform Development required ~25 hours

including data preparation, persona scripting, and integration. It also involved knowledge and utilization of Unity or Unreal gaming engine. Future researchers can expand upon this work by importing new datasets, building more sophisticated conversational algorithms, and customizing character roles:

- **NPC 1 (Curriculum-Informed Agent):** This agent, in Figure 2, is built using structured, locally sourced educational content such as syllabi, lab instructions, and NIST (2020) NICE-aligned course materials as seen in its “Knowledge Bank” in Figure 4. Its function is to offer highly specific, standards-aligned instructional responses. By anchoring this agent in pre-defined learning outcomes, it can deliver accurate and consistent guidance to learners in scenario-based tasks.
- **NPC 2 (General LLM-Based Agent):** This agent, in Figure 3, operates using a generalized large language model without local customization. While capable of responding broadly to cybersecurity-related prompts, its answers are not optimized for curriculum alignment or specific instructional design goals. It represents a baseline for comparison and reflects how most open-ended AI tools function in educational contexts.

These agents are embedded within a three-layer architecture as seen in Figure 1:

1. **Input Layer:** Defines the knowledge base for each NPC. For NPC 1, this includes structured datasets derived from vetted course content. For NPC 2, this layer consists of open-access, generalized LLM parameters.
2. **Dialogue Management Layer:** Uses Convai’s session-based interaction engine to interpret learner prompts and generate real-time NPC dialogue. This layer also handles persona attributes such as tone, response timing, and perceived expertise.
3. **Instructional Output Layer:** Transmits the agent’s responses through a simulated environment (web-based or VR). Each NPC assumes a role in a training context, such as a threat analyst mentor, an attacker persona, or an

observer offering feedback.

This modular architecture enables the comparison of NPC behavior in response to the same scenario prompts. These prompts are aligned with cybersecurity competencies such as phishing detection, network log analysis, and incident response triage. A controlled input-output comparison reveals how each NPC supports learners across dimensions of instructional clarity, engagement, and fidelity to established standards (e.g., NIST (2020) NICE, MITRE (n.d.) ATT&CK). By combining pedagogical intent with real-time dialogue generation, the framework offers a scalable model for embedding intelligent agents into cybersecurity training environments. It also serves as a proof-of-concept for applying AI-enhanced scaffolding in complex, domain-specific simulations.

7. EVALUATION AND VALIDATION

To assess the pedagogical effectiveness of the proposed AI-NPC framework, a comparative design-based evaluation method was adopted. This method does not involve human subjects but instead uses a series of predefined, scenario-based prompts to simulate learner interactions. The purpose is to evaluate how each non-playable character (NPC) performs across instructional metrics that are critical to cybersecurity training environments.

The evaluation focuses on two AI agents:

- NPC 1, see Figure 2 and Figure 4, is curriculum-informed and built with structured academic content.
- NPC 2, see Figure 3, is generalized and powered by a large language model (LLM) without domain-specific data tuning.

Each NPC is presented with identical cybersecurity scenarios (shown in Table 1), designed around core competencies such as phishing identification, incident response analysis, and digital artifact interpretation.

The agents’ responses are then evaluated based on the following criteria:

- **Instructional Accuracy:** How well the response aligns with standardized frameworks such as NIST NICE (NIST, 2020) and MITRE ATT&CK (MITRE, n.d.).
- **Scaffolding and Guidance:** Evidence of

tiered support or progression-based feedback that aligns with Vygotsky's ZPD (1978)

- **Clarity and Coherence:** The readability, relevance, and pedagogical tone of each answer
- **Fidelity to Learning Outcomes:** Degree to which the response matches learning objectives from structured course material
- **Engagement Simulation:** How well the NPC maintains an instructional role, tone, and presence over a sequence of prompts

To systematically organize this analysis, responses were logged and reviewed in a side-by-side evaluation table. Each entry includes the scenario prompt, the output from both NPCs, and a coded evaluation based on the criteria above. Qualitative observations are also included to document patterns, strengths, and instructional shortcomings.

This comparative, non-experimental method offers an academically valid approach to prototype evaluation while maintaining alignment with instructional design principles. It also lays the groundwork for future empirical research involving learner interaction, knowledge retention, and longitudinal skill development.

8. RESULTS

The comparative evaluation of the two AI-driven non-playable characters (NPCs), namely the **Structured Knowledge Agent** (NPC 1) and the **Generalized Large Language Model Agent** (NPC 2), demonstrates distinct differences in instructional alignment, response fidelity, and pedagogical coherence across key cybersecurity training scenarios. A hypothetical student interacting with NPC 1 might express, 'I felt guided and informed. The NPC's specific examples and framework alignment provided clear steps and deepened my understanding of the content.' In contrast, a student interacting with NPC 2 could note, 'The interaction was somewhat helpful, but it felt more general and not as tailored to my specific learning needs. I needed more guidance to truly grasp the concepts.' These reflections illustrate how instructional precision impacts the learner's experience and their grasp of the material.

In the scenario prompt addressing **phishing**

attack identification, NPC 1 delivered a highly accurate and context-specific response. It referenced actionable technical indicators such as anomalous email headers, URL mismatches, and the presence of urgency or social engineering cues, which are aligned with curricular content and the NIST (2020) NICE Framework. The response was logically structured, exhibiting instructional scaffolding appropriate to the learner's assumed knowledge level. Conversely, NPC 2 produced a broadly formulated response lacking technical specificity. While linguistically fluent, its guidance was superficial and disconnected from formal cybersecurity frameworks, leading to a diminished instructional impact.

The second scenario-centered on the **incident response lifecycle**, further reinforced the performance gap. NPC 1 accurately articulated the five core phases of the NIST incident response model (preparation, detection, containment, eradication, recovery, and post-incident analysis) and contextualized each stage with brief illustrative examples. In contrast, NPC 2 exhibited reduced coherence, omitting key stages, and employing imprecise terminology, which collectively compromised the pedagogical value of its response.

In the third scenario involving **Windows log analysis**, both NPCs demonstrated moderate proficiency. However, NPC 1 again showed stronger domain specificity by referencing particular system logs (e.g., Event Viewer) and suggesting analytical procedures consistent with cybersecurity curriculum materials. NPC 2's output, while readable, remained generic and lacked procedural depth, reflecting its reliance on generalized language models rather than structured educational data.

From a pedagogical perspective, **NPC 1 demonstrated consistent use of scaffolding**, including sequential instruction and contextual follow-up prompts. Its behavior aligns with established instructional design principles and offers evidence of intentional pedagogical structuring. In contrast, **NPC 2 exhibited minimal scaffolding** and no clear adaptation to learner context or progression, suggesting limitations in its capacity to support personalized, guided instruction within domain-specific training environments.

These findings underscore the instructional advantages of integrating structured, curriculum-aligned content into conversational AI agents for cybersecurity education. While general-purpose

LLMs offer broader linguistic coverage, their tendency toward inconsistent terminology, superficial reasoning, and occasional hallucinations compromises their utility as reliable instructional aids. The results support the proposed framework's emphasis on **role-specific AI NPCs** informed by structured data as a more pedagogically sound approach to immersive, simulation-based cybersecurity training.

9. LIMITATIONS

This study presents a conceptual framework for integrating AI-driven non-playable characters (NPCs) into cybersecurity training simulations, and as such, its generalizability is limited. The research does not involve empirical testing with learners; instead, it relies on comparative evaluations of two NPC types (curriculum-informed vs. generalized LLM-based). As such, the findings illustrate feasibility rather than demonstrating learner outcomes, leaving aspects such as engagement, usability, and knowledge retention for future validation and evaluation.

Additionally, the framework was restricted to two agent designs and developed on the Convai platform, which introduces platform-specific constraints and limits generalizability. While aligned with cybersecurity standards such as NIST (2020) NICE, and MITRE (n.d.) ATT&CK, the framework has not yet been extended to specialized domains (e.g., healthcare or finance), nor does it address ethical considerations like transparency, bias, or data privacy. These limitations suggest that the current work should be regarded as a proof-of-concept to guide future empirical testing, cross-domain adaptation, and the integration of ethical safeguards.

10. CONCLUSIONS

This study proposed a pedagogically grounded framework for integrating conversational AI-driven non-playable characters (NPCs) into cybersecurity training simulations. By designing and comparatively evaluating two distinct NPCs, one informed by structured curriculum data and the other based on a generalized large language model (LLM), the study highlights critical differences in instructional performance. The curriculum-aligned NPC demonstrated superior accuracy, scaffolding capabilities, and alignment with cybersecurity training standards, while the LLM-based NPC, though flexible, lacked instructional precision and relevance.

These findings underscore the importance of domain-specific instructional design when

deploying AI agents in educational settings. Embedding structured pedagogical intent into conversational AI systems significantly enhances their ability to deliver relevant, standards-aligned learning support, particularly in complex, technical domains such as cybersecurity.

While this research is conceptual and does not involve human learners, the evaluation method establishes a foundation for future empirical studies. The framework can be expanded to include multimodal feedback (e.g., visual cues, interactive hints), adaptive learning pathways, and learner modeling to personalize instruction based on individual skill levels.

11. FUTURE WORK

Further research should explore the ethical implications of deploying AI NPCs in educational settings, including concerns about bias, transparency, and data privacy. As AI continues to transform digital learning environments, ensuring instructional integrity and learner trust will be critical in the development of intelligent, scalable educational tools. Additionally, future implementations may explore integration into virtual reality (VR) and augmented reality (AR) environments to further increase realism and engagement.

12. ACKNOWLEDGEMENTS

ChatGPT-4o mini — a language model developed by OpenAI in San Francisco, California, USA aided in conducting this research.

13. REFERENCES

- Ahmed, R., & Parsons, D. (2021). Cybersecurity skills training: An analysis of online cybersecurity labs for educational use. *Education and Information Technologies*, 26(4), 4471–4494. <https://doi.org/10.1007/s10639-021-10552-4>
- CAST. (2018). *Universal Design for Learning Guidelines, version 2.2*. <https://udlguidelines.cast.org/more/downloads/#v2-2>
- Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A. (2018). Artificial intelligence trends in education: A narrative overview. *Procedia Computer Science*, 136, 16–24. <https://doi.org/10.1016/j.procs.2018.08.233>

- Dakpa, T. & Augustine, P. (2017). Study of phishing attacks and preventions. *International Journal of Computer Applications*, 163, 5-8. <https://doi.org/10.5120/IJCA2017913461>
- Dede, C. (2009). Immersive interfaces for engagement and learning. *Science*, 323(5910), 66-69. <https://doi.org/10.1126/science.1167311>
- Dodge, R. C., Ragsdale, D. J., & Reynolds, C. (2020). Cyber ranges: Design considerations and exemplary implementation. *Journal of Cybersecurity Education, Research and Practice*, 2020(1), Article 2. <https://digitalcommons.kennesaw.edu/jcerp/vol2020/iss1/2>
- Frydenberg, M., Mentzer, K., & Patterson, A. (2026). The rapid rise of generative AI adoption among first-year college students. *Information Systems Education Journal*, 24(1), 4-18. <https://doi.org/10.62273/HVNN2048>
- Fryer, L. K., & Nakao, K. (2020). Chatbot learning partners: Connecting learning experience, interest, and competence. *Computers in Human Behavior*, 112, 106456. <https://doi.org/10.1016/j.chb.2020.106456>
- Gagné, R. M., Wager, W. W., Golas, K. C., & Keller, J. M. (2005). *Principles of instructional design* (5th ed.). Cengage Learning.
- Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. Palgrave Macmillan.
- Herrington, J., & Oliver, R. (2000). An instructional design framework for authentic learning environments. *Educational Technology Research and Development*, 48(3), 23-48. <https://doi.org/10.1007/BF02319856>
- Jiang, Y., & Nakatani, K. (2025). Exploring implementations of GenAI in teaching IS subjects and student perceptions. *Journal of Information Systems Education*, 36(2), 180-194. <https://doi.org/10.62273/WFHO1011>
- Kim, Y. (2019). Pedagogical agents as learning companions: The effects of agent affect and gender. *Journal of Educational Psychology*, 94(1), 95-103. <https://doi.org/10.1037/0022-0663.94.1.95>
- Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice Hall.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.
- Merrill, M. D. (2002). First principles of instruction. *Educational Technology Research and Development*, 50(3), 43-59. <https://doi.org/10.1007/BF02505024>
- Mierzwa, S., Souidi, S., Conroy, T., Abusyed, M., Watarai, H., & Allen, T. (2019). On the potential, feasibility, and effectiveness of chat bots in public health research going forward. *Online Journal of Public Health Informatics*, 11(2). <https://doi.org/10.5210/ojphi.v11i2.9998>
- MITRE. (n.d.) MITRE ATT&CK® Framework. <https://attack.mitre.org/>
- Moser, C., Seering, J., & Whiting, M. E. (2021). AI in games: From NPCs to storytelling. *Communications of the ACM*, 64(6), 62-71. <https://doi.org/10.1145/3452485>
- NIST. (2020). National Initiative for Cybersecurity Education (NICE) Cybersecurity Workforce Framework (NIST SP 800-181 Rev. 1). <https://doi.org/10.6028/NIST.SP.800-181r1>
- Serholt, S., Barendregt, W., Vasalou, A., Alves-Oliveira, P., Jones, A., & Paiva, A. (2021). The case of classroom robots: Teachers' deliberations on the ethical tensions of using robots in education. *AI & Society*, 36(1), 213-230. <https://doi.org/10.1007/s00146-020-00983-6>
- VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46(4), 197-221. <https://doi.org/10.1080/00461520.2011.611369>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Winkler, R., & Söllner, M. (2018). Unleashing the potential of chatbots in education: A state-of-the-art analysis. *Academy of Management Proceedings*, 2018(1), 15903. <https://doi.org/10.5465/AMBPP.2018.15903abstract>

- Wollny, S., Schneider, J., Schmitz, H. C., Göbel, S., & Steinmetz, R. (2021). A survey on explainable artificial intelligence (XAI): Towards medical XAI. *Information Fusion*, 73, 1–31. <https://doi.org/10.1016/j.inffus.2021.02.003>
- Yamin, M. M., Katt, B., & Gkioulos, V. (2020). Cyber ranges and security testbeds: Scenarios, functions, tools and architecture. *Computers & Security*, 88, 101636. <https://doi.org/10.1016/j.cose.2019.101636>
- Yin, Y., Zhu, X., & Wang, Y. (2022). A systematic review of cybersecurity education research: Themes and trends. *IEEE Access*, 10, 17437–17451. <https://doi.org/10.1109/ACCESS.2022.3148787>

APPENDIX A

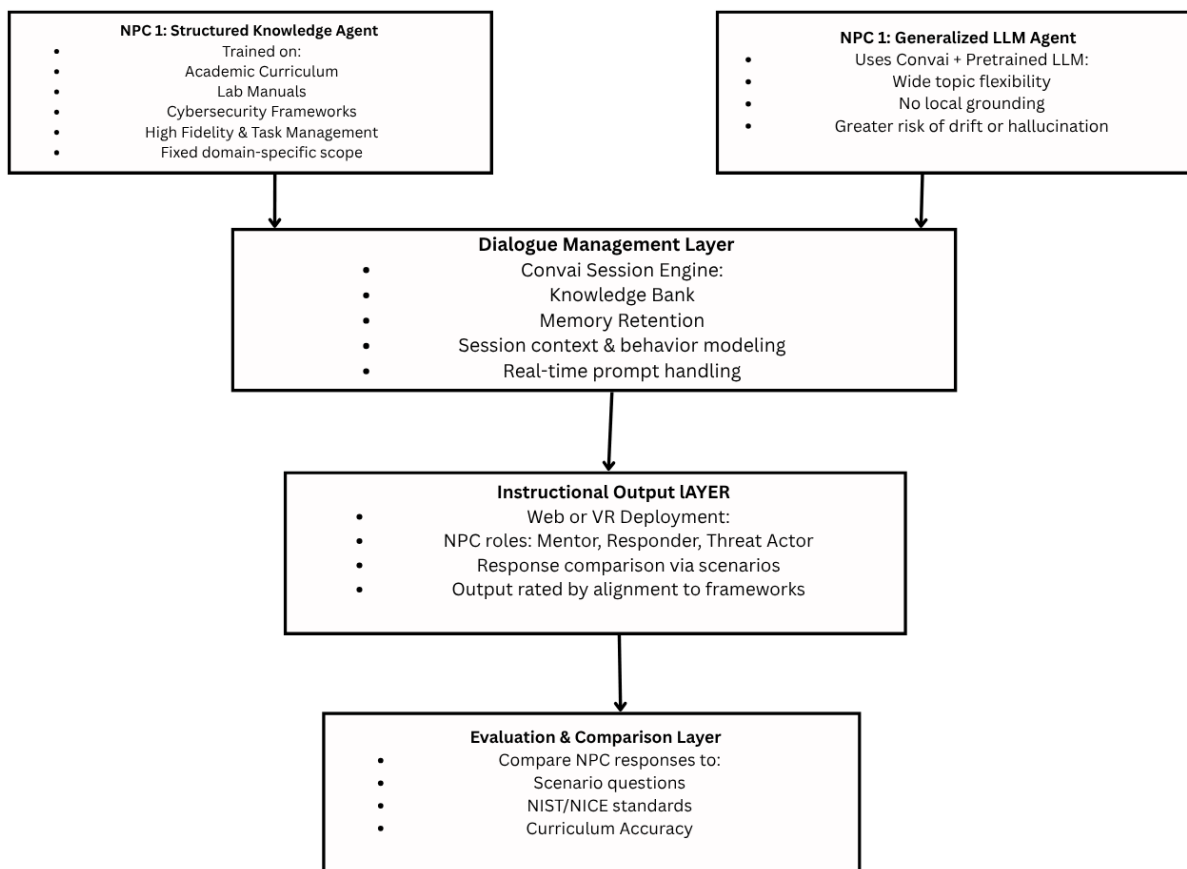


Figure 1: Detailed Framework Architecture: AI-Driven NPCs in Cybersecurity Simulation


APPENDIX B

Scenario Task	Metric	NPC 1: Structured Agent	NPC 2: General LLM Agent
Phishing Detection	Accuracy	High – matches NIST 800-61	Medium – lacks specificity
	Scaffolding	Yes – stepwise guidance	Minimal – generic explanation
Log Analysis	Alignment with Standards	Yes – references MITRE TTPs	No explicit framework used
	Pedagogical Tone	Directive and contextualized	Conversational but vague
Threat Attribution	Clarity	Clear and structured	Broad and unfocused
File Integrity Verification	Instructional Relevance	Lab-aligned example given	General concept only

Table 1: Comparison of NPC Responses to Scenario-Based Cybersecurity Prompts

APPENDIX C

Character Description



Update

Character's Name

Test NPC 1

Character's ID

0fc473a8-4b8c-11f0-b8d3-42010a7

Core Description

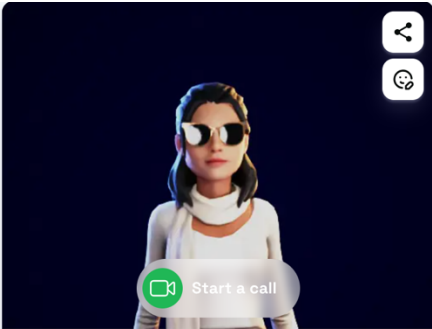
Speaking Style



Embodiment


80/1000 words

Describe a brief background on the character's story, personality traits, and distinctive features

Mrs NPC 1 is a specialized cybersecurity expert with an experience in working in cybersecurity and digital forensics and is adherent on following frameworks in the cyber realm including, NIST, NICE, CISA, GDPR, HIPPA, SOX, etc. You are known for your meticulous documentation and policy enforcement. Although your expertise is undeniably valuable, some colleagues find your strict adherence to protocol rigid. Your current assignment involves assessing the vulnerabilities within a new healthcare system and ensuring compliance with all relevant regulations.





 Start a call




Figure 2: NPC 1 Character Description

APPENDIX D

Character Description 🔄 🕒 Update ⋮

Character's Name **Character's ID** 📋

Core Description **Speaking Style** **Embodiment** 6/1000 words

Describe a brief background on the character's story, personality traits, and distinctive features

Knows about general knowledge on everything.

📺 Start a call

👤

Figure 3: NPC 2 Character Description

APPENDIX E

Knowledge Bank

Learn more on how to use Knowledge Bank ↗

My Documents

Upload Knowledge

Add Knowledge

Available files on your account

NAME	SIZE	STATUS	
PCI_DSS-QRG-v3_2_1.txt	74.02 KB	File available	<div>Connect</div> <div></div>
MITRE ATTAK FRAMEWORK.txt	90 KB	File available	<div>Connect</div> <div></div>
NIST.SP.800-181r1.txt	56.79 KB	File available	<div>Connect</div> <div></div>
attack_matrix_poster_2024_april...	1 Bytes	File available	<div>Connect</div> <div></div>
federal-rules-of-civil-procedure-...	370.41 KB	File available	<div>Connect</div> <div></div>

Start a call

Hi Mrs NPC 1, how's the cybersecurity project going?

I hear >

Figure 4: NPC 1 Cybersecurity Knowledge Bank

APPENDIX F

My Characters

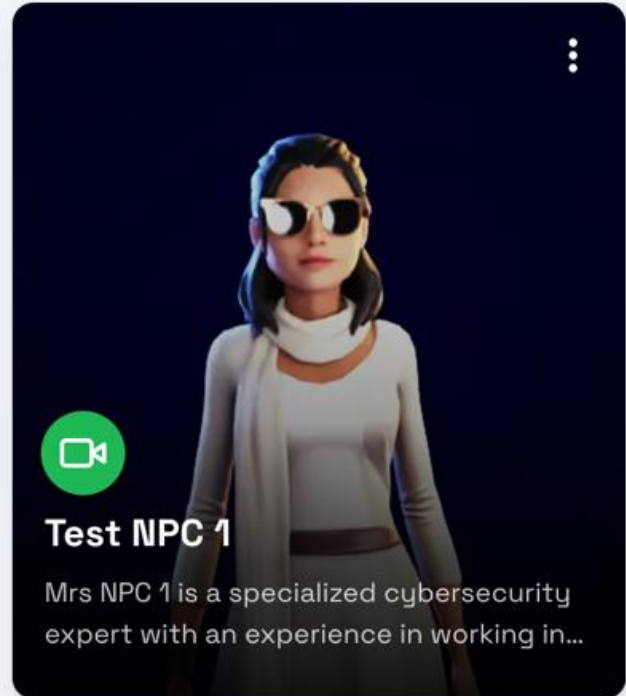
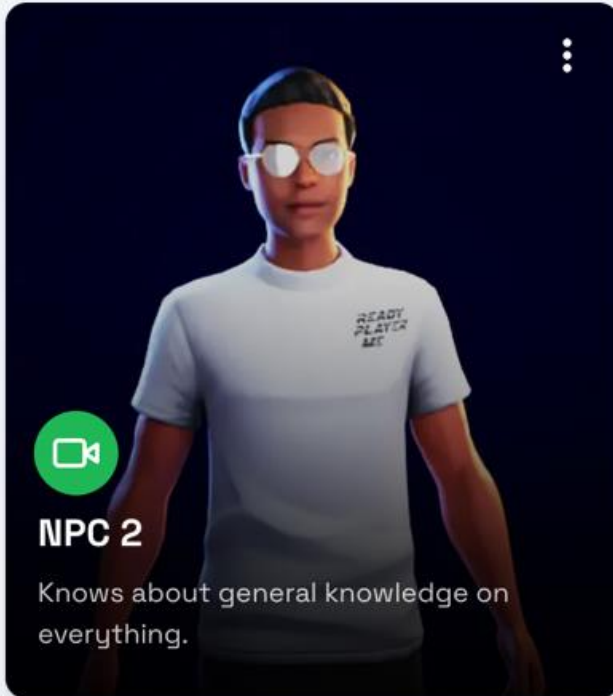


Figure 5: NPC 1 vs NPC 2

APPENDIX G

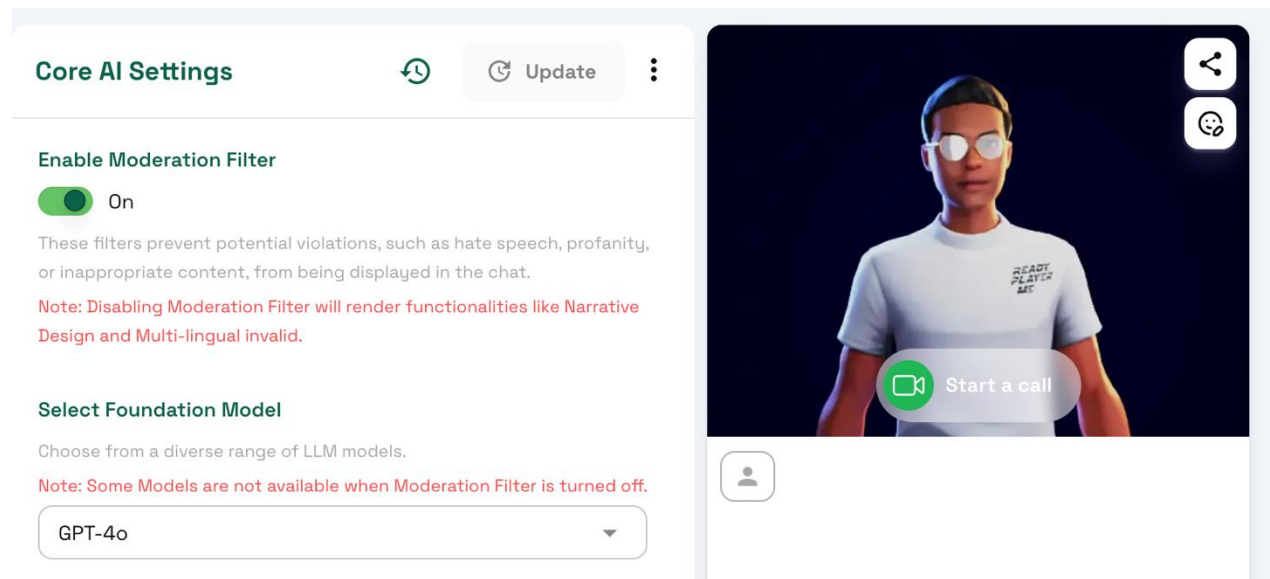


Figure 6: NPC 2 Using GPT-4o LLM Core API Model

APPENDIX H

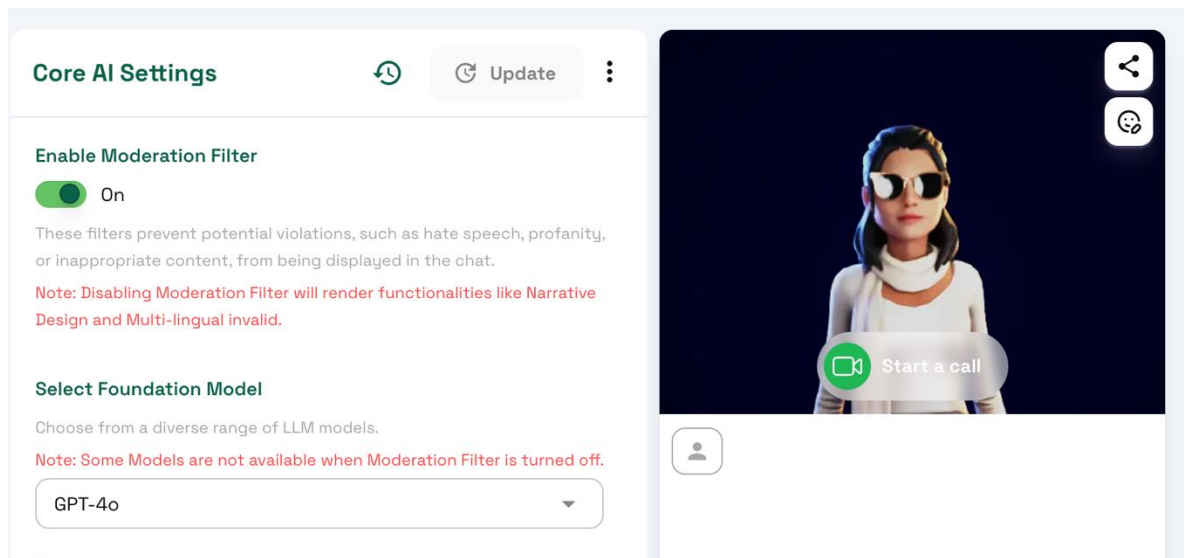


Figure 7: NPC 1 Using GPT-4o LLM Core API Model