

WIP: Clinic-in-a-Box: AI-Powered Stakeholder Simulation for Cybersecurity Education

Ryan Straight¹, Paul Wagner², Rob Honomichl³ and Shengjie Xu⁴

Abstract—This work-in-progress innovative practice category paper presents Clinic-in-a-Box (CiaB), an AI-powered stakeholder simulation system designed to strengthen communication skill development in cybersecurity clinic education. Cybersecurity clinic programs require students to engage with diverse organizational stakeholders, yet meaningful practice opportunities remain limited before real client interactions. CiaB addresses this gap by using large language models to create persistent, sector-specific stakeholder personas that maintain consistent organizational profiles across student interactions. The system currently supports four client types representing common consulting contexts and incorporates sector-specific security frameworks, persistent organizational memory, and comprehensive interaction logging for assessment. This paper describes the system architecture, its grounding in clinic-based learning and design science research, and preliminary development status. Planned evaluation includes expert persona review, comparative analysis against traditional role-play, and student perception surveys.

Index Terms—cybersecurity education, stakeholder simulation, large language models, clinic-based learning, experiential learning

I. INTRODUCTION

This work-in-progress innovative practice paper presents Clinic-in-a-Box (CiaB), an AI-powered stakeholder simulation system for cybersecurity clinic education. Cybersecurity clinic programs have emerged as a significant model for experiential education, connecting students with real organizations that need cybersecurity assessment and guidance [1]. Modeled after legal and medical clinic traditions, these programs position students as consultants who conduct risk assessments for under-resourced organizations, including school districts, small businesses, nonprofits, and municipal governments. The Consortium of Cybersecurity Clinics now coordinates efforts across multiple institutions, reflecting that authentic stakeholder engagement develops professional competencies that coursework does not.

A persistent challenge in clinic-based cybersecurity education is preparation. Students entering client engagements must navigate complex interpersonal dynamics with stakeholders of varying technical understanding, organizational priorities, and communication styles. Traditional preparation approaches rely on instructor-led role-play exercises, which suffer from interrelated limitations: inconsistency across sessions and instructors, limited scalability beyond small cohorts, and an inability to systematically capture and assess interaction data.

CiaB addresses these limitations through large language model (LLM)-based persistent stakeholder personas. CiaB creates simulated organizational clients that maintain consistent profiles, apply sector-appropriate security knowledge, and generate structured interaction logs suitable for instructional

assessment. The system complements authentic client engagements with scalable, repeatable preparation.

This paper makes two contributions: a system architecture that pairs persistent LLM stakeholder personas with sector-specific knowledge bases and assessment-oriented interaction logging, and a phased evaluation design for testing whether such simulation develops stakeholder-communication competency relative to traditional role-play.

The remainder of this paper is organized as follows: Section II reviews related work in clinic-based learning, AI educational agents, and stakeholder simulation. Section III presents the CiaB system design. Section IV describes the preliminary development status and planned evaluation. Section V discusses implications for the engineering education community.

II. RELATED WORK

A. Clinic-Based Learning in Professional Education

Clinic-based learning models originated in legal education, where law clinics provide supervised client representation that develops professional judgment under faculty oversight [2]. Medical education employs standardized patients, trained actors who simulate clinical encounters, enabling consistent and repeatable practice with built-in assessment capabilities [3]. Both traditions demonstrate that supervised, authentic client engagement cultivates competencies that classroom instruction alone cannot develop, including communication under uncertainty, ethical judgment in ambiguous situations, and the management of ongoing client relationships.

Cybersecurity education has begun adapting these models through programs where students conduct security assessments for community organizations [1]. Nti Asare et al. [4] report early lessons from implementations across multiple institutions, identifying scalable assessment models and the mutual benefits of student-community engagement. Cybersecurity clinic implementations, however, face scalability challenges distinct from legal and medical analogues. Unlike law clinics where cases develop over weeks, cybersecurity engagements require students to rapidly understand diverse organizational contexts, from school district IT infrastructure to municipal government compliance requirements. The breadth of sector-specific knowledge required strains traditional preparation methods, and client availability constrains both the frequency and variety of practice opportunities.

B. AI-Mediated Educational Environments

Research on intelligent tutoring systems demonstrates that AI-mediated practice environments can provide personalized, scal-

able learning experiences [5]. A meta-analysis of simulation-based learning in higher education found large positive effects across professional disciplines, with scaffolding and reflection components significantly moderating outcomes [6].

More recently, LLM-based educational agents have shown promise for role-play scenarios across disciplines. Studies suggest that LLMs offer new opportunities for personalized and scalable educational interaction, including simulated professional communication [7]. In medical education, LLM-powered simulated patient systems have demonstrated performance comparable to human standardized patients in clinical communication training while offering advantages in availability, consistency, and scalability [8]. Affordances include on-demand availability and behavioral consistency across interactions, along with interaction capture; however, limitations around hallucination, authenticity, and pedagogical alignment require deliberate instructional design attention [9].

C. Stakeholder Simulation in Cybersecurity Education

Simulation-based approaches in cybersecurity education have focused primarily on technical environments: cyber ranges, capture-the-flag competitions, and network defense exercises [10], [11]. These tools develop technical competencies effectively but do not address the communication and stakeholder management dimensions of cybersecurity practice, despite these skills being emphasized by the NICE Workforce Framework [12] and consistently identified by employers as critical workforce competencies [13].

The medical education parallel is instructive. Standardized patients emerged precisely because clinical communication skills could not be developed through technical training alone, and the transition from human actors to LLM-powered simulations is already underway in that field [8]. Cybersecurity education has no analogous system for stakeholder communication practice. Existing approaches rely on instructor-led role-play, written case studies, and tabletop exercises, which, while valuable, remain inconsistent and difficult to scale. This gap, between extensive technical simulation infrastructure and limited stakeholder communication tools, motivates the CiaB system.

III. SYSTEM DESIGN

CiaB is designed as a modular stakeholder simulation system that integrates with existing cybersecurity clinic curricula. The system architecture (Fig. 1) centers on four components: organizational persona profiles, sector-specific knowledge bases, a persistent interaction engine, and an instructor assessment dashboard.

A. Organizational Persona Profiles

The system supports four client types representing common cybersecurity consulting contexts:

1. **K-12 school districts** facing student data protection requirements, limited IT staffing, and compliance mandates under FERPA and state regulations.

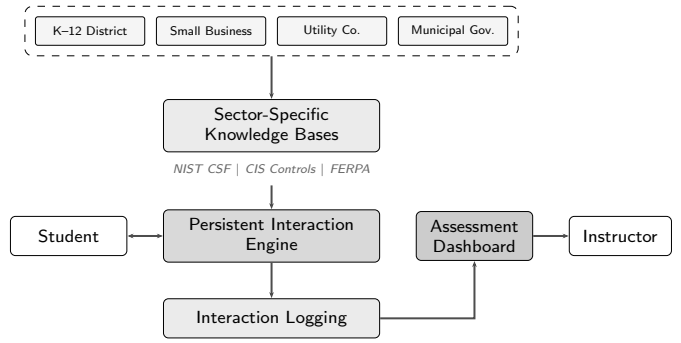


Fig. 1: CiaB system architecture.

2. **Small businesses** with constrained budgets, minimal cybersecurity awareness, and reliance on consumer-grade technology infrastructure.
3. **Utility companies** operating industrial control systems alongside standard IT environments, subject to sector-specific regulatory frameworks.
4. **Municipal governments** managing citizen data, public-facing services, and cross-departmental coordination challenges.

Each client type is instantiated through a structured organizational profile that specifies the organization's mission, size, technology environment, regulatory context, prior incident history, and stakeholder communication preferences. The K-12 school district persona, for example, represents a mid-sized district with approximately 8,000 students, a three-person IT team, a recent phishing incident that compromised staff email accounts, and a superintendent whose primary concern is protecting student data under severe budget constraints. The superintendent persona communicates in non-technical language, expresses frustration with unfunded mandates, and responds with appropriate wariness when students propose solutions requiring significant capital expenditure.

These profiles are generated through a template system informed by the eight-part risk assessment framework developed as part of the broader CiaB toolkit. The framework structures clinic activities across a complete engagement arc: (1) orientation and context setting, (2) client/organization scoping, (3) threat identification, (4) vulnerability identification, (5) risk analysis and prioritization, (6) control and mitigation identification, (7) reporting and communication, and (8) reflection and workforce alignment. Each phase offers multiple modular activity options at varying complexity levels, enabling instructors to customize the experience for secondary, post-secondary, or advanced learners.

B. Sector-Specific Knowledge Bases

Rather than relying on generic LLM knowledge, CiaB incorporates curated knowledge bases aligned to each organizational context. Knowledge bases include mappings to relevant cybersecurity frameworks: NIST Cybersecurity Framework [14] for utility companies, CIS Critical Security Controls [15] for small businesses, FERPA-aligned controls for school districts,

and multi-framework compliance requirements for municipal governments. Simulated stakeholders therefore reference appropriate standards and articulate concerns consistent with their sector’s regulatory environment.

C. Persistent Interaction Engine

The interaction engine maintains organizational memory across student encounters, enabling persistent client relationships that develop over multiple sessions. Each persona is instantiated through a structured system prompt that encodes the organizational profile, sector-specific knowledge base references, and communication register constraints. Conversation history is retained across sessions using a managed context window, enabling the LLM to reference prior disclosures and maintain behavioral consistency without re-prompting. When a student conducts an initial scoping interview with a simulated school district superintendent, the system retains that conversation’s context. Subsequent interactions, including follow-up questions, preliminary findings presentations, and final reporting, build on established rapport and previously disclosed information.

This persistence supports the full arc of a clinic engagement, from initial client scoping through threat identification, risk analysis, control recommendation, and final reporting. Students experience the cumulative nature of consulting relationships rather than isolated, stateless exchanges. CiaB is designed for use in the early weeks of a clinic course, before students begin authentic client engagements.

D. Interaction Logging and Assessment

All student-stakeholder interactions are logged with structured metadata including timestamps, conversation topics, communication quality indicators, and framework alignment tags. Instructors access an assessment dashboard that surfaces patterns across student cohorts: common communication gaps, frequently missed organizational context cues, and progression in professional communication competency over time.

The logging architecture supports both formative and summative assessment approaches, enabling instructors to provide targeted feedback on specific interaction segments rather than relying solely on final deliverable evaluation.

E. Example Interaction Sequence

To illustrate the system’s operation, consider a student team assigned to conduct a risk assessment for a simulated K–12 school district. The engagement proceeds through three CiaB interaction sessions aligned with the risk assessment framework.

In the *scoping session* (Part 2), students interview the simulated superintendent to understand the district’s technology environment. The persona discloses that the district uses a shared Google Workspace instance, maintains student records in a legacy student information system, and recently experienced a phishing incident that compromised three staff email accounts. When students ask about the district’s cybersecurity budget, the superintendent responds with realistic frustration: the district has no dedicated cybersecurity line item, and IT spending

competes directly with instructional materials. The system logs whether students identify FERPA compliance as a concern, whether they ask about third-party vendor access, and how they adapt their language to a non-technical stakeholder.

In the *findings presentation* (Part 7), students present their preliminary risk analysis to the superintendent. The persona responds to recommendations based on its organizational profile: proposals requiring significant expenditure receive skeptical pushback, while low-cost administrative controls receive more receptive responses. Students must defend their prioritization and adapt their communication in real time.

In the *reflective debrief* (Part 8), the instructor reviews the interaction logs, identifying specific moments where students missed organizational context cues, used excessive technical jargon, or effectively translated risk into business impact terms. This structured feedback loop, grounded in logged interaction data rather than instructor recollection, enables targeted communication skill development.

F. NICE Workforce Framework Alignment

The CiaB system’s eight-part framework maps directly to NICE Workforce Framework task, knowledge, and skill (TKS) statements across four work roles: Risk Management Specialist, Cybersecurity Analyst, Security Control Assessor, and Cybersecurity Consultant [12]. The AI stakeholder interaction components specifically develop competencies in stakeholder communication, requirements gathering, and professional reporting, competencies that are difficult to cultivate through technical exercises alone. This alignment enables programs to document how clinic-based activities contribute to nationally recognized workforce competency development.

G. Planned Voice Integration

Future iterations of CiaB will incorporate voice-based interaction through simulated voicemails and audio channels, since consultants frequently conduct client contacts by phone. Voice interaction would enable students to practice interpreting tone, urgency, and implicit organizational concerns conveyed through vocal cues that text-only interaction cannot capture.

IV. PRELIMINARY DEVELOPMENT AND PLANNED EVALUATION

A. Development Status and Methodology

The CiaB project launched in January 2026 within the University of Arizona with a faculty team and five undergraduate researchers. The project follows a design science research (DSR) methodology [16], [17], iterating through problem identification, artifact design, demonstration, and evaluation cycles. Current development is in the artifact design phase, with three concurrent workstreams: organizational persona development, toolkit documentation, and integration planning with the academy’s existing cybersecurity clinic infrastructure.

The persona development workstream has produced initial profile templates for all four client types. Each template underwent review against publicly available organizational data from the corresponding sector to ensure that technology

environments, regulatory references, staffing patterns, and communication styles reflect authentic conditions. The research team also conducted a market analysis of existing cybersecurity simulation and training platforms, identifying that no current product combines persistent LLM-based stakeholder personas with sector-specific knowledge bases and structured interaction logging for assessment.

The toolkit documentation workstream produced the training guidance manual, an instructor resource that operationalizes the eight-part risk assessment framework with modular student output options at three complexity tiers. This document provides the curricular scaffolding within which the AI stakeholder simulation operates.

Development artifacts are iteratively refined through instructor walkthroughs, alignment checks against NICE Workforce Framework task, knowledge, and skill statements [12], and feedback from cybersecurity practitioners serving as domain experts.

B. Evaluation Framework

The evaluation framework addresses three research questions aligned with the system’s design objectives:

- **RQ1:** To what extent do CiaB personas maintain contextually appropriate and consistent responses across extended student interactions?
- **RQ2:** How do student interactions with CiaB personas compare to traditional role-play exercises in developing stakeholder communication competencies?
- **RQ3:** What are student and instructor perceptions of CiaB persona authenticity and pedagogical value?

C. Evaluation Phases and Methods

The evaluation proceeds in three phases aligned with the DSR build-evaluate cycle:

Phase 1: Expert Validation (Spring 2026). Six to eight cybersecurity practitioners who regularly conduct organizational risk assessments will evaluate persona responses across a structured set of interaction scenarios. Reviewers will assess whether simulated stakeholders exhibit sector-appropriate concerns, respond consistently to follow-up questions, and maintain plausible organizational profiles. Evaluation criteria include technical accuracy of framework references, behavioral consistency across sessions, and communication register appropriateness. This phase addresses RQ1 and informs iterative persona refinement before student deployment.

Phase 2: Pilot Deployment (Fall 2026). Students enrolled in clinic opportunities at the University of Arizona will engage with CiaB personas as part of their preparation for real client engagements. A within-subjects design will compare communication artifacts produced during CiaB interactions with those produced during traditional instructor-led role-play exercises conducted in the same course. Artifacts will be evaluated using a rubric adapted from professional communication competency frameworks, assessing dimensions including stakeholder needs identification, technical-to-non-technical

translation, recommendation feasibility, and professional tone. This phase addresses RQ2.

Phase 3: Perception and Impact Assessment (Fall 2026).

Following CiaB interactions, students will complete perception surveys measuring authenticity, engagement, perceived learning value, and self-reported communication confidence using both Likert-scale and open-ended response items. Instructor observations and interaction log analysis will supplement survey data. This phase addresses RQ3.

The study protocol is planned for institutional review board submission at the University of Arizona. Anticipated participants include 20–30 students enrolled in clinic opportunities during the Fall 2026 semester.

V. IMPLICATIONS AND FUTURE WORK

The CiaB architecture provides a replicable model for developing domain-specific LLM educational agents. The approach of pairing structured organizational profiles with curated knowledge bases and persistent interaction memory is transferable to other professional education contexts where stakeholder communication is critical, including engineering consulting, public health, and social work.

The eight-part risk assessment framework and associated modular activity design offer a structured curriculum model that can operate with or without the AI simulation component. Institutions can adopt the toolkit’s scaffolded activities even before implementing the technology layer, making the broader CiaB approach accessible across resource contexts. The tiered activity design, with options for secondary, post-secondary, and advanced learners, supports adoption across a range of educational settings.

The planned evaluation also contributes methodological insight into assessing AI-mediated professional communication practice. As LLM-based educational tools proliferate, validated approaches for measuring their pedagogical effectiveness remain underdeveloped. The phased evaluation framework described here may inform assessment designs for similar systems in other professional disciplines.

A. Ethical Considerations

The use of AI-simulated stakeholders raises considerations that the research team is actively addressing. Students must understand that CiaB interactions serve as preparation for authentic client engagement, and there is a risk that students develop communication patterns optimized for AI interlocutors rather than human stakeholders. The comparative evaluation design (Phase 2) is intended to detect such divergence. The system must also avoid reinforcing stereotypical representations of under-resourced organizations; persona profiles are reviewed by practitioners from the represented sectors to mitigate this concern.

B. Limitations

As a work-in-progress, this paper describes a system in active development at time of acceptance. The persona profiles have not yet been evaluated by external experts, and no student

interaction data is available. CiaB personas are designed to develop baseline communication competency; the variability and unpredictability of real stakeholder interactions require additional preparation through supervised authentic engagements. The initial evaluation is based at a single institution, though multi-site pilots with K–12 partner districts are in planning, which may strengthen generalizability in subsequent work. The system’s dependence on commercial LLM APIs introduces cost, availability, and content policy dependencies that may affect reproducibility across institutions.

Future development will expand persona diversity beyond the initial four client types, refine assessment rubrics based on evaluation findings, and investigate multi-institution adoption pathways through the Consortium of Cybersecurity Clinics network.

VI. REFERENCES

- [1] “Consortium of Cybersecurity Clinics.” 2026 [Online]. Available: <https://cybersecurityclinics.org/>. [Accessed: 17-Feb-2026]
- [2] D. A. Maranville, L. Bliss, C. W. Kaas, and A. Sedillo Lopez, “Building on Best Practices: Transforming Legal Education in a Changing World.” Social Science Research Network, Rochester, NY, June-2015 [Online]. Available: <https://papers.ssrn.com/abstract=2553452>. [Accessed: 17-Feb-2026]
- [3] H. S. Barrows, “An overview of the uses of standardized patients for teaching and evaluating clinical skills,” *Academic Medicine*, vol. 68, no. 6, pp. 443–51, June 1993 [Online]. Available: <http://journals.lww.com/00001888-199306000-00002>. [Accessed: 17-Feb-2026]
- [4] I. Nti Asare, S. Shackleford, J. Chun, and S. Powazek, “Protecting Communities while Training Future Cybersecurity Professionals: Lessons from the Consortium of Cybersecurity Clinics,” *The Cyber Defense Review*, vol. 10, no. 2, pp. 241–265, Dec. 2025 [Online]. Available: <https://cyberdefensereview.army.mil/CDR-Content/Articles/Article-View/Article/4344290/protecting-communities-while-training-future-cybersecurity-professionals-lesson/>. [Accessed: 17-Feb-2026]
- [5] A. C. Graesser *et al.*, “AutoTutor: A tutor with dialogue in natural language,” *Behavior Research Methods, Instruments, & Computers*, vol. 36, no. 2, pp. 180–192, May 2004 [Online]. Available: <http://link.springer.com/10.3758/BF03195563>. [Accessed: 17-Feb-2026]
- [6] O. Chernikova, N. Heitzmann, M. Stadler, D. Holzberger, T. Seidel, and F. Fischer, “Simulation-Based Learning in Higher Education: A Meta-Analysis,” *Review of Educational Research*, vol. 90, no. 4, pp. 499–541, Aug. 2020 [Online]. Available: <https://journals.sagepub.com/doi/10.3102/0034654320933544>. [Accessed: 17-Feb-2026]
- [7] E. Kasneci *et al.*, “ChatGPT for good? On opportunities and challenges of large language models for education,” *Learning and Individual Differences*, vol. 103, p. 102274, Apr. 2023 [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1041608023000195>. [Accessed: 18-May-2023]
- [8] H. Yu *et al.*, “Simulated patient systems powered by large language model-based AI agents offer potential for transforming medical education,” *Communications Medicine*, vol. 6, no. 1, p. 27, Dec. 2025 [Online]. Available: <https://www.nature.com/articles/s43856-025-01283-x>. [Accessed: 17-Feb-2026]
- [9] L. Yan *et al.*, “Practical and ethical challenges of large language models in education: A systematic scoping review,” *British Journal of Educational Technology*, vol. 55, no. 1, pp. 90–112, Jan. 2024 [Online]. Available: <https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjjet.13370>. [Accessed: 17-Feb-2026]
- [10] D. H. Tobey, P. Pusey, and D. L. Burley, “Engaging learners in cybersecurity careers: Lessons from the launch of the National Cyber League,” *ACM Inroads*, vol. 5, no. 1, pp. 53–56, Mar. 2014.
- [11] J. Vykopal, M. Vizvary, R. Oslejsek, P. Celeda, and D. Tovarnak, “Lessons learned from complex hands-on defence exercises in a cyber range,” in *2017 IEEE Frontiers in Education Conference (FIE)*, 2017, pp. 1–8 [Online]. Available: <http://ieeexplore.ieee.org/document/8190713/>. [Accessed: 17-Feb-2026]
- [12] R. Petersen, D. Santos, K. Wetzel, M. Smith, and G. Witte, “NIST SP 800-181 Rev. 1: Workforce Framework for Cybersecurity (NICE Framework),” National Institute of Standards and Technology, NIST Special Publication (SP) 800-181 Rev. 1, 2020 [Online]. Available: <https://csrc.nist.gov/pubs/sp/800/181/r1/final>. [Accessed: 27-Sept-2024]
- [13] ISC2, “2025 ISC2 Cybersecurity Workforce Study.” 2025 [Online]. Available: <https://www.isc2.org/Insights/2025/12/2025-ISC2-Cybersecurity-Workforce-Study>. [Accessed: 17-Feb-2026]
- [14] National Institute of Standards and Technology, “The NIST Cybersecurity Framework (CSF) 2.0,” National Institute of Standards and Technology, Gaithersburg, MD, NIST CSWP 29, Feb. 2024 [Online]. Available: <https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.29.pdf>. [Accessed: 20-Mar-2024]
- [15] Center for Internet Security, “CIS Critical Security Controls Version 8.1.” 2025 [Online]. Available: <https://www.cisecurity.org/controls/v8-1>. [Accessed: 17-Feb-2026]
- [16] A. R. Hevner, S. T. March, J. Park, and S. Ram, “Design Science in Information Systems Research,” *MIS Quarterly*, vol. 28, no. 1, pp. 75–106, Mar. 2004 [Online]. Available: <https://misq.umn.edu/misq/article/28/1/75/261/Design-Science-in-Information-Systems-Research1>. [Accessed: 17-Feb-2026]
- [17] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, “A Design Science Research Methodology for Information Systems Research,” *Journal of Management Information Systems*, vol. 24, no. 3, pp. 45–77, Dec. 2007 [Online]. Available: <https://www.tandfonline.com/doi/full/10.2753/MIS0742-1222240302>. [Accessed: 17-Feb-2026]