Learning Engineering by Design: An Agentic AI Application for Rapid, Personalized Health & Safety Training in Disaster Response and Hazardous Environments

Henry Ryng & Daniel Suhr

Agentic AI | cmi5 | Competency-Based Training | Disaster Response Training | HAZMAT Cleanup | ICICLE Standards | learning engineering | Personalized Learning | Rapid Content Development | Total Learning Architecture

Disaster response and hazardous materials (HAZMAT) cleanup require rapid preparation of diverse, transient workers for site-specific risks. Traditional methods fail to meet tight timelines, creating safety gaps. The Health & Safety Training (HST) Copilot, funded by NIEHS SBIR Phase I (Award 1R43ES037510-01), applies learning engineering principles to embed evidence-based design in an AI platform. Integrating ICICLE best practices and IEEE TLA frameworks, an 11-agent architecture generates personalized primers and assessments from site documents. Development follows a nested learning engineering cycle: human expertise assembly, agentic AI

simulation, evidence/decision tracking (LEED-inspired), and refinement. Phase I SME rubric evaluation confirmed technical feasibility, instructional quality, and rapid generation, with stakeholder commitment to Phase II. Future studies will assess learning transfer and safety outcomes.

Introduction

In disaster response (e.g., hurricanes, chemical spills, terrorism incidents) and HAZMAT cleanup, mobilizing workers occurs in hours or days, yet each site presents unique hazards requiring tailored preparation. Workers arrive with varied backgrounds, yet limited contact training time must prioritize hands-on practice over basic hazard review. Inadequate preparation contributes to preventable incidents (OSHA, 2023). Traditional development timelines (6-8 weeks) fail here, while credentialing often relies on attendance certificates offering no insight into competency or currency amid skills decay.

Structured training often fails to address learner variability (varying reading levels, multilingual needs, and prior experience gaps), limiting effectiveness for mobile workforces and emergency responders. Specific contexts include hurricane/wildfire disaster response requiring rapid team mobilization, HAZMAT remediation with site-specific chemical protocols, terrorism/conflict preparedness for emergency response teams, and pandemic response requiring broad multi-language communication of infection control protocols to diverse health worker populations.

The HST Copilot addresses this via learning engineering—a transdisciplinary, iterative process applying learning sciences, human-centered design, systems engineering, and data-informed decisions to create effective, scalable learning solutions (Goodell, Kessler, & Schatz, 2023; Baker, Boser, & Snow, 2022). Unlike traditional instructional design, LE mandates ongoing evidence collection and iteration to improve outcomes systematically.

This project embeds LE "by design": principles become automatic in the platform, benefiting users without requiring their expertise. Development followed a nested LE cycle (Craig et al., 2025): human interdisciplinary teams defined challenges and decisions, agentic AI simulated specialist collaboration, rubric-based evaluation provided evidence, and Learning Engineering Evidence and Decision (LEED) tracking mechanisms documented rationale for traceability and refinement. The platform integrates Industry Consortium for Innovation and Collaboration in Learning Engineering (ICICLE) best practices and guidelines, along with alignment to IEEE Total Learning Architecture (TLA) competency frameworks in the underlying technology, throughout the agent architecture.

Learning Engineering by Design: Human Teams, Evidence Tracking, and Nested Cycles

LE centers on assembling diverse expertise, applying evidence-based practices, and tracking decisions systematically to ensure rigor and iterability (Goodell & Kolodner, 2023; Thai et al., 2023). HST Copilot development mirrored this: a human LE team—including construction supervisors, disaster response coordinators, safety managers, instructional technologists versed in xAPI/cmi5, and operational stakeholders—collaborated to define requirements and validate outputs.

Design decisions were systematically documented using Kessler's Learning Engineering Evidence and Decision (LEED) approach (Totino & Kessler, 2024), tracking what evidence informed which decisions and why, enabling retrospective analysis

("Why did we do that?") and continuous refinement. This human process informed the core innovation: an agentic AI "team" simulating specialist collaboration.

Agent 1 (Instructional Designer) defines objectives via backward design. Agent 2 (Training Specialist) classifies knowledge, skills, and abilities (KSAs) using frameworks from MIL-HDBK-29612-2A (U.S. Department of Defense, 2001) and Bloom's Taxonomy (Bloom, 1956)—standards HST stakeholders were already familiar with from their domain expertise. Subsequent agents personalize readability, motivation, visuals, and structure per ICICLE principles (cognitive load management, dual coding, learner engagement). This mimics human LE teams but scales rapidly.

The nested cycle—human expertise → AI simulation → evaluation → refinement—aligns with LE's iterative, data-informed nature (Goodell et al., 2023).

The Training Readiness Challenge in NIEHS Contexts

NIEHS-focused domains (disaster response, HAZMAT remediation, emergency preparedness) demand site-specific, time-critical preparation amid diverse workforces. Challenges include:

1. Timelines: Hours to days for deployment vs. weeks or months for traditional training development
2. Unique hazards: Each incident site presents distinct dangers requiring customized content
3. Learner variability: Reading levels (6th-16th grade equivalents), multilingual needs, prior experience gaps, and frequent inclusion of unskilled workers and day laborers rapidly recruited into post-disaster operations
4. Limited contact time: Must prioritize hands-on practice over didactic review of basic hazards
5. Credential trust: Traditional attendance certificates provide no verifiable evidence of competency or skills currency
6. Skills decay: Mobile and intermittently deployed workers experience competency deterioration without systematic refresher mechanisms

The HST Copilot Solution

Users upload available documents (training materials, site assessments, safety protocols, SOPs); the 11-agent pipeline generates personalized primers (Agents 1-10) aligned to ICICLE standards, then assessments (Agent 11) structured for IEEE Total Learning Architecture (TLA) competency frameworks. cmi5 instrumentation enables analytics for decay detection and competency tracking across deployments. TLA structures produce verifiable digital badges with evidence of specific capabilities, trusted by regulators and employers.

Multi-Agent Architecture with Embedded Standards

Each agent embeds specific learning engineering principles from ICICLE standards, creating a collaborative architecture where outputs build progressively toward complete, standards-aligned training materials:

Agent	LE/ICICLE/TLA Function
1	Extract objectives (backward design)
2	Classify KSAs (MIL-HDBK-29612-2 + Bloom's; TLA alignment)
3	Personalize readability (cognitive load management)
4	Motivational framing (learner engagement)

5	Visual icons (dual coding theory)
6-9	Scenario-based practice items (authentic assessment)
10	Compile with TLA/cmi5 structure (competency evidence)
11	Generate TLA-aligned assessments (performance-based evaluation)

SME review ensures accuracy and safety compliance; human-centered iteration builds stakeholder trust and operational validity.

Phase I Evaluation: Rubric-Based Platform Assessment

Per NIH SBIR Phase I constraints (no human subjects research without IRB approval), evaluation focused on platform technical capability to generate accurate, appropriate pre-training content. A structured 27-criterion rubric assessed primer/quiz quality (content accuracy, clarity, standards alignment, readability appropriateness), platform usability, and overall operational readiness across domains including HAZMAT remediation and disaster response.

Stakeholders (director-level representatives from disaster response and safety organizations) provided high ratings on problem-solution fit (M=4.7/5.0), customization capability (M=4.7/5.0), and overall value proposition (M=4.6/5.0); all (4/4, 100%) committed to Phase II field testing. Preliminary rubric-based evaluation yielded high ratings for content quality (M=4.6/5.0) with zero critical safety errors identified by expert reviewers. Technical validation demonstrated 95% reduction in development time (4.5 hours vs. 6-8 weeks for traditional instructional design approaches), enabling rapid response timelines required for dynamic disaster contexts.

Rubric feedback validated agentic outputs' instructional integrity while identifying areas for enhancement (e.g., assessment distractor realism via future retrieval-augmented generation). This LE-aligned evaluation—systematic evidence collection, decision tracking, and iterative refinement—demonstrates platform feasibility for rapid, personalized training generation.

Addressing Measurement, Trust, and Future Outcomes

ICICLE/TLA frameworks link performance data to competencies via structured evidence (cmi5 instrumentation, scenario authenticity, knowledge checks), supporting construct validity and credibility for regulators and employers. Trust builds through SME oversight, adherence to established standards (MIL-HDBK, Bloom's, ICICLE, TLA), and transparent documentation of design decisions via LEED tracking.

Current work establishes promising infrastructure for rapid, personalized preparation. Phase I validated platform feasibility and content quality through SME evaluation. Phase II will measure actual learner outcomes: transfer to hands-on performance, behavior change in operational contexts, and safety impact (incident/near-miss rates, supervisor performance ratings, competency retention patterns) via IRB-approved protocols in authentic disaster response and HAZMAT environments.

Conclusion

HST Copilot demonstrates LE excellence made automatic: nested cycles combining human expertise with agentic AI simulation, systematic evidence and decision tracking via LEED, and embedded learning sciences standards (ICICLE, TLA) democratize rigorous instructional design for underserved high-risk contexts. By embedding learning engineering excellence into automated processes, the platform democratizes access to rigorous instructional design for underserved populations—

disaster workers, mobile construction crews, and emergency health responders—who traditionally lack such resources. By advancing rapid, verifiable preparation, it supports safer disaster response, HAZMAT operations, and emergency preparedness in contexts where training failures have life-or-death consequences.

Acknowledgments

This work was supported by the NIEHS SBIR Phase I (Award 1R43ES037510-01). The authors thank beta stakeholders for evaluation participation and the ICICLE and TLA communities for standards development that enabled this work.

References

Baker, R. S., Boser, U., & Snow, E. L. (2022). Learning engineering: A view on where the field is at, where it's going, and the research needed. Technology, Mind, and Behavior. https://doi.org/10.1037/tmb0000058

Bloom, B. S. (1956). Taxonomy of educational objectives: The classification of educational goals. Longmans, Green.

Craig, S. D., Avancha, K., Malhotra, P., C., J., Verma, V., Likamwa, R., Gary, K., Spain, R., & Goldberg, B. (2025). Using a Nested Learning Engineering Methodology to Develop a Team Dynamic Measurement Framework for a Virtual Training Environment. In International Consortium for Innovation and Collaboration in Learning Engineering (ICICLE) 2024 Conference Proceedings: Solving for Complexity at Scale (pp. 115-132). https://doi.org/10.59668/2109.21735

Goodell, J., Kessler, A., & Schatz, S. (2023). Learning Engineering at a Glance. Journal of Military Learning, 7, 42-52.

Goodell, J., & Kolodner, J. (Eds.). (2023). Learning engineering toolkit: Evidence-based practices from the learning sciences, instructional design, and beyond. Taylor & Francis.

Occupational Safety and Health Administration. (2023). Commonly used statistics. U.S. Department of Labor. https://www.osha.gov/data/commonstats

Thai, K. P., Craig, S. D., Goodell, J., Lis, J., Schoenherr, J. R., & Kolodner, J. (2023). Learning Engineering is Human-Centered. In J. Goodell (Ed.), The Learning Engineering Toolkit (pp. 83-124). Routledge.

Totino, L., & Kessler, A. (2024). "Why did we do that?" A Systematic Approach to Tracking Decisions in the Design and Iteration of Learning Experiences. The Journal of Applied Instructional Design, 13(2). https://doi.org/10.59668/1269.15630

U.S. Department of Defense. (2001). MIL-HDBK-29612-2A: Instructional Systems Development/Systems Approach to Training and Education, Part 2: Acquisition. Department of Defense.