Designing a WebVR and RAG-based Training Service for Medical Device Operation: A Service Design Prototype for Novice Biotech Operators

Sheng-Ming Wang[1,*,†], Yu-Tung Hsiao[1,†], Pin-Yan Shih[1,†] and Chung-Wen Hsiao[1,†]

[1]*Department of Interaction Design, National Taipei University of Technology, Taipei, Taiwan*

Abstract

Medical equipment operation training in pharmaceutical and clinical laboratory settings has long relied on static Standard Operating Procedure (SOP) documents, making it difficult for novice operators to develop spatial understanding of equipment operations and safely practice abnormal scenarios. This study proposes a service design prototype that integrates WebVR-based immersive training and a Retrieval-Augmented Generation (RAG) AI Agent to support the learning of medical equipment operation procedures for novice biotechnology operators. The system was designed through persona development, AT-ONE touchpoint analysis, and Responsible AI governance principles. The primary contribution of this work is a design framework and functional prototype that integrates immersive learning, SOP-grounded AI assistance, and Responsible AI governance. Empirical validation is identified as a direction for future research.

Keywords

WebVR, Medical Training, AI Agent, Service Design, Mixed Reality

1. Introduction

Training in the operation of medical equipment is a high-risk process in pharmaceutical and clinical laboratory environments. Before entering actual operations, novice operators must become familiar with equipment layouts, operational procedures, and handling abnormal situations. However, current training approaches are largely based on static SOP documents and instructor demonstrations, limiting the opportunities for learners to practice high-risk procedures repeatedly in a safe environment [1]. Immersive technologies such as WebVR provide opportunities for scenario-based training in virtual environments, allowing learners to become familiar with operational procedures and spatial relationships without interacting with real equipment [2]. Meanwhile, the development of Retrieval-Augmented Generation (RAG) enables AI systems to retrieve external knowledge sources before generating responses to knowledge-intensive queries. This approach helps reduce hallucination risks associated with relying solely on model parameters and allows the retrieved knowledge sources to be inspected and traced. Therefore, in SOP-based operational training contexts, RAG can serve as a technical foundation for procedural knowledge support and source transparency [3].

This study proposes a prototype of a medical equipment operation training service centered on WebVR and an RAG-based AI Agent. The target users are novice operators in the biotechnology and pharmaceutical industries. The system integrates a 360-degree virtual laboratory, interactive equipment hotspots, AI-assisted question answering based on SOP, and task-based assessments with certification mechanisms.

APMAR'26: The 18th Asia-Pacific Workshop on Mixed and Augmented Reality, Aug. 3-5, 2026, Taipei, Taiwan

*Corresponding author.

†These authors contributed equally.

ryan5885@mail.ntut.edu.tw (S. Wang); amy27378186@gmail.com (Y. Hsiao); vicky131831@gmail.com (P. Shih); amyangel9425@gmail.com (C. Hsiao)

2. Related Work

2.1. Immersive Technologies for Professional Training

VR and WebXR technologies have been widely applied in high-risk training contexts such as surgical procedures, equipment maintenance, and emergency response. Previous studies indicate that immersive environments can effectively support the transfer of training, facilitate the acquisition of procedural skills, and reduce operational errors [1]. Compared to dedicated head-mounted VR devices that often require specialized hardware and complex setup procedures, WebXR can be deployed through standard web browsers, offering cross-device compatibility ranging from mobile devices to advanced VR headsets while significantly reducing organizational adoption barriers [4]. However, existing research suggests that current VR interfaces still lack realistic haptic feedback, which limits the user's ability to perform fine-grained finger-level manipulations and reduces the overall sense of natural interaction and realism [5]. Therefore, this study positions the WebXR prototype as a tool for situational familiarization and procedural rehearsal prior to hands-on training, rather than as a replacement for physical equipment training.

2.2. RAG for Professional Knowledge Retrieval

Large language models face hallucination challenges in medical and regulatory contexts. RAG addresses this issue by grounding generated responses in external knowledge sources and is widely considered an effective approach to improving answer reliability and traceability [3]. Recent studies have applied RAG to medical question-answering systems by integrating external medical literature and professional knowledge sources, thereby reducing the risk of inaccurate AI-generated information while improving the reliability and verifiability of the system [6]. However, research on human–AI collaboration suggests that when AI systems directly provide recommendations or answers, users may become reliant on heuristics and exhibit reduced motivation to engage in analytical thinking, leading to excessive reliance on AI [7]. To mitigate this risk, researchers have proposed Cognitive Forcing Functions, which encourage active user participation in decision-making through mechanisms such as delayed hints, requiring users to make an initial judgment, or revealing supporting evidence before presenting recommendations [7].

2.3. Research Gap

Although previous studies have separately explored immersive training, RAG systems, and Responsible AI governance, a service design framework that integrates these three components within the context of medical equipment operation training remains underexplored. This study seeks to address this gap by proposing a prototype of a medical equipment operation training service.

3. System Design

3.1. Design Context and Target Users

The proposed system is designed for novice operators in the biotechnology and pharmaceutical industries. Although these users typically possess basic laboratory knowledge, they often lack hands-on experience with specific equipment. Their primary training challenges include the inability to safely rehearse high-risk abnormal operations repeatedly, difficulty mapping written SOP procedures to physical equipment locations, and limited access to immediate support from Subject Matter Experts (SMEs). This study adopts the processing of blood samples as the core training scenario and focuses on three types of equipment: centrifuges, hematology analyzers, and biosafety cabinets.

3.2. Service Design Methodology

This prototype was developed as part of a three-week Problem-Based Learning (PBL) project using service design methodologies. The design process included user research, persona development, customer journey mapping, and service planning. In particular, the project employed the AT-ONE touchpoint analysis framework [8] and service blueprinting techniques[9]. These methodologies were selected because prior research has demonstrated their effectiveness in facilitating communication and integration among cross-functional stakeholders, including operators, SMEs, QA managers, and IT teams. Furthermore, they support the coordinated design of front-stage user experiences and back-stage governance systems through visualization of stakeholder interactions and organizational responsibilities [9].

3.3. System Architecture and Training Workflow

The proposed system adopts a linear workflow with branching interactions (Figure 1). The learners progress through four sequential stages: introduction to the training task ⟶ 360° WebVR Laboratory Exploration ⟶ training of the equipment and AI-Assisted Inquiry ⟶ assessment and certification based on tasks. The AI Agent can be accessed throughout the exploration, equipment training, and assessment stages.

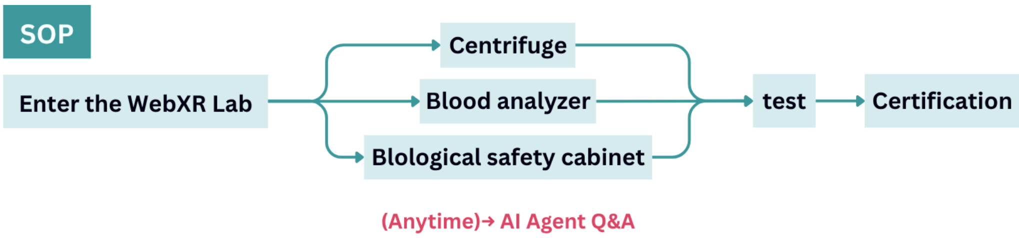

Figure 1: System Architecture

3.3.1. Training Task Introduction

In Figure 2, upon entering the system, learners are first presented with a training task card that outlines the course title (Hematology Analyzer Operation and Blood Sample Processing), estimated completion time (5 minutes), difficulty level (beginner) and the list of equipment covered in the session. This design is informed by Cognitive Load Theory [10], which suggests that human cognitive processing capacity is limited. By providing learners with a clear learning framework and task expectations before entering a highly stimulating immersive virtual environment, the system aims to reduce the likelihood of cognitive overload and enable learners to allocate their limited cognitive resources to subsequent equipment interactions and core skill acquisition.

3.3.2. 360° WebVR Laboratory Environment

In Figure 3, after entering the 360° virtual laboratory, students are free to explore the environment and access detailed information panels for individual equipment through interactive glowing hotspots embedded within the scene. These panels provide information on equipment functions, operational procedures, and risk warnings. For example, the centrifuge panel introduces a three-step procedure involving balanced sample placement, speed configuration, and post-operation sample retrieval, while also highlighting risks such as mechanical damage caused by improper balancing. For high-risk operational areas, the system provides Glow-based visual highlighting as a layered guidance mechanism, encouraging learners to infer procedures independently rather than receiving direct answers. This exploratory design aims to support the development of spatial understanding of the laboratory equipment layouts of the learners. Previous studies have shown that free exploration in virtual environments can effectively facilitate the construction of environmental knowledge and cognitive maps [11]. Furthermore, compared to traditional two-dimensional media, such as documents or maps, immersive spatial experiences have been found to significantly improve user understanding of relative object locations, thus addressing limitations of conventional training approaches in supporting spatial cognition [12].

Figure 2: Training Task Introduction

Figure 3: 360° WebVR Laboratory Environment

3.3.3. SOP-Grounded RAG AI Agent

In the left panel of Figure 3, during scene exploration and assessment activities, learners can access the AI Agent panel at any time to ask questions. The system provides several predefined questions as guided entry points, while also allowing learners to freely submit their own queries. The AI Agent follows a RAG-only principle: all responses must be retrieved from approved SOP documents imported into the knowledge base. For each response, the system displays the corresponding source document name, version number, and section reference (e.g., SOP-001 Blood Sample Handling v1.0, Section 2.1). When no verifiable source can be retrieved from the knowledge base, the Agent refuses to provide an answer and instead directs learners to consult a Subject Matter Expert (SME) or Quality Assurance (QA) personnel. This design ensures the transparency of information sources and the auditability of AI-generated responses while clearly distinguishing the role of AI assistance from the official interpretation and decision-making of SOPs.

From a technical perspective, the proposed RAG pipeline consists of four stages: document preprocessing and indexing, semantic retrieval, context grounding, and response generation. During knowledge base construction, approved SOP documents are segmented into semantically meaningful text chunks and indexed to support efficient semantic retrieval. When a learner submits a question, the system first converts the query into semantic embeddings and retrieves the most relevant SOP document chunks from the indexed knowledge base. The retrieved passages are then provided as the only external context for response generation, ensuring that all answers are grounded in verified SOP documents rather than relying solely on the language model's internal knowledge. Each response presents the corresponding SOP document name, version number, and section reference to facilitate source verification and improve traceability. If no sufficiently relevant evidence can be retrieved, the system refuses to generate an answer and instead advises learners to consult a Subject Matter Expert (SME) or Quality Assurance (QA) personnel. By constraining AI-generated responses to retrieved and verifiable SOP knowledge, the proposed RAG-only mechanism reduces the risk of hallucination while improving the transparency, traceability, and accountability of AI-assisted learning.

3.3.4. Equipment Training and Abnormal Scenario Simulation

In Figure 4, after selecting a specific piece of equipment, the learners enter a equipment-specific training module. Using the Hematology Analyzer as an example, a step-by-step guidance panel is displayed on the right side of the interface, presenting operational procedures from Step 1 to Step 5. The learners are required to complete each step sequentially before progressing to the next stage. For example, in Step 1, the learners must verify the consistency of the sample tube labels, laboratory request forms, and patient information recorded in the system. They are required to confirm that the patient name, identification number, and specimen type match exactly. The system simultaneously presents relevant risk warnings (e.g., incorrect sample identification can result in test results being assigned to the wrong patient) together with the corresponding SOP reference. This module also incorporates learning of abnormal scenarios. When learners trigger a simulated error scenario, the system presents common mistakes, potential risks, and corrective recommendations associated with the current operational step, along with relevant SOP references. By exposing learners to contextualized error cases, the system aims to improve their understanding of operational risks and response strategies while reducing the gap between theoretical knowledge of SOP and practical application. Previous studies have shown that reproducing high-risk or emergency situations in real-world environments is often difficult due to safety and cost constraints [13]. Virtual reality technologies provide a safe, high-tolerance and repeatable training environment in which learners can repeatedly practice abnormal situation handling procedures without exposing actual equipment or personnel to risk [14].

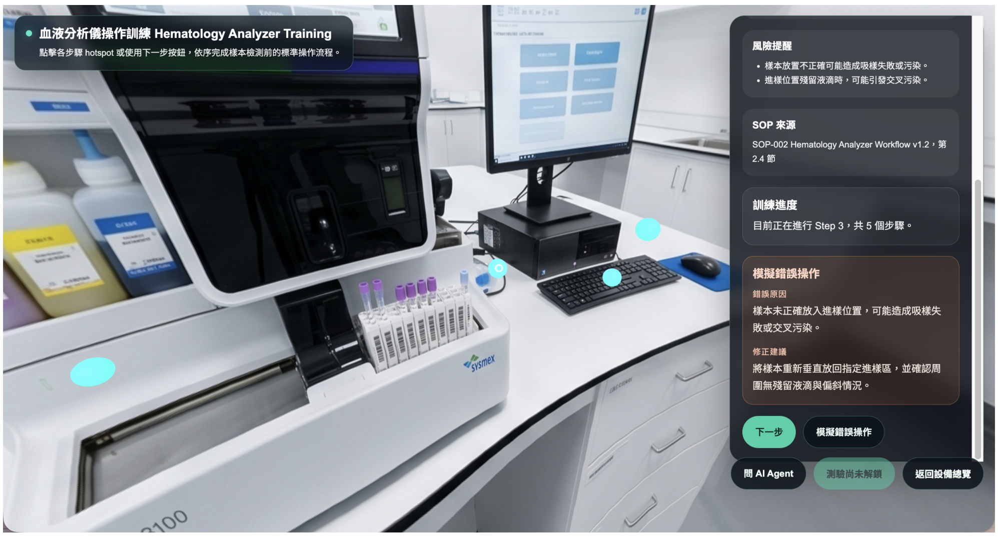

Figure 4: Abnormal Scenario Simulation

Figure 5: Task-Based Assessment and Certification

3.3.5. Task-Based Assessment and Certification

In Figure 5, after completing the equipment exploration and training activities, the learners proceed to a task-based assessment consisting of five multiple-choice questions related to SOP for the selected equipment. Learners who do not pass the assessment are guided back to the training module to review the relevant operational procedures and SOP content before reattempting the assessment. Upon successful completion, the system issues a certification and records the learner's training status, thereby completing the training workflow.

3.4. Responsible AI Design Principles

The AI Agent was designed according to the following governance principles to address accountability requirements in high-risk medical training contexts.

- **Transparency:** Each response displays the corresponding SOP source, version number, and section reference, allowing learners to verify the original documentation and improving the traceability of AI-generated responses.
- **Knowledge Boundaries:** The AI Agent is restricted to retrieving information only from verified SOP documents imported into the knowledge base. When no verifiable source is available, the system refuses to provide an answer and instead directs users to consult a Subject Matter Expert (SME) or Quality Assurance (QA) personnel.
- **Assessment Monitoring:** Learners are permitted to access the AI Agent during assessments; however, all usage records are logged and incorporated into training analytics to prevent AI assistance from masking learners' actual level of understanding.
- **Human Oversight:** The system includes administrative review interfaces for QA managers and SMEs, ensuring that AI-generated responses serve solely as supplementary information and do not replace formal SOP updates or professional judgment.

4. Discussion and Limitations

4.1. Design Contributions

The primary contribution of this study is the proposal of a service design framework that integrates WebVR-based immersive learning, RAG-enabled knowledge retrieval, and Responsible AI governance. Although previous studies have often discussed immersive training and AI assistance separately, this work attempts to address both front-stage learning experiences (e.g., spatial familiarization, task guidance, and real-time inquiry) and back-stage governance requirements (e.g., source traceability, audit records, and human oversight) within a single training service. The browser-based deployment of WebVR reduces organizational adoption barriers and enables immersive training experiences without requiring dedicated VR hardware, making it suitable for large-scale training implementation. In addition, the source disclosure mechanism of the RAG AI Agent responds to concerns about AI overreliance identified in human–AI collaboration research. Users may accept AI recommendations without sufficient verification; therefore, AI systems should not only provide answers but also reveal the evidence supporting those answers and encourage users to engage in active judgment through source verification and layered guidance mechanisms [7]. This study aims to foster critical and responsible use of AI-assisted information by allowing learners to view the source document, version number, and section reference associated with each response.

4.2. Limitations

First, the proposed prototype has not yet undergone formal user evaluation. Current design decisions are primarily informed by existing literature and service design methodologies rather than empirical validation with novice operators. Future work will conduct controlled user studies to compare the proposed WebVR + RAG training service with conventional PDF-based SOP training in terms of task completion time, operational error rate, training confidence, AI reliance, and overall user experience.

Second, the current RAG knowledge base is built upon simulated SOP documents. The prototype has not yet been integrated with real-world enterprise SOP version management systems, and issues related to knowledge base completeness, version conflict management, and multilingual document support remain to be addressed.

Third, the WebVR environment simplifies the complexity of real-world equipment operations. Although the 360-degree virtual environment provides opportunities for spatial familiarization and procedural rehearsal, it cannot reproduce tactile feedback or fine motor interactions. Consequently, the proposed system is intended as a preparatory training tool rather than a replacement for hands-on equipment training.

Finally, future work will include a comprehensive evaluation of both user experience and Responsible AI perceptions. User experience will be evaluated using the User Experience Questionnaire Short Version (UEQ-S), while additional Responsible AI questionnaire items will assess learners' perceptions of source utilization, critical engagement, understanding of AI limitations, active judgment, assessment transparency, and operational risk awareness. This evaluation aims to examine not only the usability and overall user experience of the proposed training service, but also whether the Responsible AI mechanisms effectively promote trustworthy, transparent, and accountable AI-assisted learning.

5. Conclusion

This study presents a prototype of medical equipment operation training service centered on WebVR and an RAG-based AI Agent. Through a combination of a virtual laboratory 360-degrees, interactive equipment hotspots, AI-assisted inquiry based on SOP and task-based assessments, the system provides novice biotechnology operators with a safe, repeatable and auditable training environment. The contribution of this work does not lie in the implementation of a fully validated medical training system. Rather, it proposes a service design framework that integrates immersive learning, AI-assisted SOP knowledge retrieval, and Responsible AI governance.

Future work will focus on integrating real-world SOP knowledge bases, further optimizing the RAG knowledge retrieval mechanism, and conducting user studies with laboratory operators. The evaluation will combine the User Experience Questionnaire Short Version (UEQ-S) with additional Responsible AI questionnaire items to investigate learners' perceptions of source utilization, critical engagement, understanding of AI limitations, active judgment, assessment transparency, and operational risk awareness. These evaluations will provide empirical evidence for assessing not only the usability of the proposed WebVR-based training service, but also the effectiveness of its Responsible AI design in supporting trustworthy, transparent, and accountable AI-assisted learning.

Acknowledgments

This research originated from a Problem-Based Learning (PBL) project conducted in the Service Design for Interactive Technology course. The authors would like to thank the course instructor for guidance on service design methodologies, interactive technology applications, and research direction. We also appreciate the valuable feedback and suggestions provided by classmates during project presentations and discussions. The system prototype and service workflow presented in this paper were developed collaboratively by the project team and will require further validation and refinement through real-world deployment and user testing.

Declaration on Generative AI

Generative AI tools were used during the preparation of this manuscript to assist with structural organization, language refinement, academic English editing, and Chinese–English translation. Generative AI was not used as the sole source of research information and did not replace the authors' judgment regarding the literature, system design, or research arguments. All AI-assisted content was reviewed, revised, and verified by the authors, who assume full responsibility for the accuracy of the citations, research findings, and conclusions presented in this paper. In addition, certain interface texts and visual mockups were developed with the assistance of generative AI tools during the ideation process. All final materials were selected, modified, and integrated by the authors prior to inclusion in this work.

References

[1] H. Stefan, M. Mortimer, B. Horan, Evaluating the effectiveness of virtual reality for safety-relevant training: a systematic review, Virtual Reality 27 (2023) 2839–2869.
[2] M. Mergen, M. Meyerheim, N. Graf, Reviewing the current state of virtual reality integration in medical education–a scoping review protocol, Systematic Reviews 12 (2023) 97.
[3] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W.-t. Yih, T. Rocktäschel, et al., Retrieval-augmented generation for knowledge-intensive nlp tasks, Advances in neural information processing systems 33 (2020) 9459–9474.
[4] F. J. C. Rodríguez, G. Frattini, F. T. P. Meireles, D. A. Terrien, S. Cruz-Leon, M. D. Peraro, E. Schier, D. M. Moreno, L. A. Abriata, Molecularwebxr: Multiuser discussions about chemistry and biology in immersive and inclusive vr, arXiv preprint arXiv:2311.00385 (2023).
[5] Y. Shi, G. Shen, Haptic sensing and feedback techniques toward virtual reality, Research 7 (2024) 0333.
[6] J. Sohn, Y. Park, C. Yoon, S. Park, H. Hwang, M. Sung, H. Kim, J. Kang, Rationale-guided retrieval augmented generation for medical question answering, in: Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers), 2025, pp. 12739–12753.
[7] Z. Buçinca, M. B. Malaya, K. Z. Gajos, To trust or to think: cognitive forcing functions can reduce overreliance on ai in ai-assisted decision-making, Proceedings of the ACM on Human-computer Interaction 5 (2021) 1–21.
[8] S. Clatworthy, Service innovation through touch-points: Development of an innovation toolkit for the first stages of new service development, International Journal of design 5 (2011).
[9] M. J. Bitner, A. L. Ostrom, F. N. Morgan, Service blueprinting: a practical technique for service innovation, California management review 50 (2008) 66–94.
[10] J. Sweller, Cognitive load during problem solving: Effects on learning, Cognitive science 12 (1988) 257–285.

[11] S. H. Creem-Regehr, J. W. Kelly, B. Bodenheimer, J. K. Stefanucci, Virtual reality as a tool to understand spatial navigation, Reference module in neuroscience and biobehavioral psychology (2023).

[12] S. U. König, A. Keshava, V. Clay, K. Rittershofer, N. Kuske, P. König, Embodied spatial knowledge acquisition in immersive virtual reality: Comparison to map exploration, Frontiers in Virtual Reality 2 (2021) 625548.

[13] D. Scorgie, Z. Feng, D. Paes, F. Parisi, T. Yiu, R. Lovreglio, Virtual reality for safety training: A systematic literature review and meta-analysis, Safety science 171 (2024) 106372.

[14] B. Xie, H. Liu, R. Alghofaili, Y. Zhang, Y. Jiang, F. D. Lobo, C. Li, W. Li, H. Huang, M. Akdere, et al., A review on virtual reality skill training applications, Frontiers in Virtual Reality 2 (2021) 645153.