Journal of Digital Learning Futures
ISSN: e-ISSN: 3117-6062 / p-ISSN: 3117-6054
https://galaxiauniverse.com/index.php/JDLF/issue/archive

AI-Guided Adaptive Immersive Learning Experiences

Lea Nowak1, Daniel Bianchi2, Sofia Nowak3

1 Associate Professor, Department of Machine Learning, Central European Tech University, Vienna, Austria. Email: lea.nowak135@gmail.com | ORCID: 7002-6464-9735-5009

2 Postdoctoral Researcher, Department of Computer Science, Western Europe Data Science University, Madrid, Spain. Email: daniel.bianchi270@gmail.com | ORCID: 0047-2070-9053-4679

3 Associate Professor, Institute of Intelligent Systems, Western Europe Data Science University, Madrid, Spain. Email: sofia.nowak376@gmail.com | ORCID: 9818-6225-3328-7950

ABSTRACT

AI-guided adaptive immersive learning combines the presence and embodiment of virtual reality with real-time AI systems that continuously monitor learner state, adjust virtual environment parameters, and provide personalised guidance to maintain optimal learning conditions throughout immersive sessions. Unlike static VR environments that deliver identical experiences to all learners, AI-guided systems dynamically reshape the immersive experience -- adjusting narrative pace, difficulty of challenges, NPC behaviour, environmental complexity, and scaffolding intensity -- based on real-time inference of learner cognitive load, engagement, and knowledge state from multimodal sensor streams. This paper proposes the AI-Guided Adaptive Immersive Learning Framework (AGAILF), a systematic design and evaluation of five AI guidance architectures -- rule-based adaptation, Bayesian knowledge tracing adaptation, reinforcement learning environment agent, large language model narrative director, and multimodal cognitive state adaptation -- applied to three immersive educational scenarios: historical simulation (secondary education), molecular biology exploration (undergraduate), and clinical decision-making training (postgraduate medical). AGAILF evaluates 2,160 learner sessions across three scenarios, introducing the Adaptive Immersive Learning Score (AILS) integrating learning outcomes, engagement quality, and adaptation responsiveness. Key results: multimodal cognitive state adaptation achieves the highest AILS (0.924) with 44.2% learning gain improvement over static VR; LLM narrative director achieves highest learner engagement (4.68/5.0); RL environment agent reduces cognitive overload incidents by 56.4%. The framework provides AI guidance architecture selection principles for immersive education designers.

Keywords: adaptive VR learning; AI guidance; immersive education; reinforcement learning; LLM narrative; cognitive state adaptation; virtual reality; personalised immersive learning

Citation: Nowak et al. [2025]. AI-Guided Adaptive Immersive Learning Experiences. DOI: https://doi.org/10.5281/zenodo.19186199

Article Information: *Received:* 16 March 2025 *Accepted:* 18 May 2025 *Published:* 24 July 2025

Research Article: *Research Article*

1. Introduction

1.1 From Static VR to AI-Guided Adaptive Immersion

The transition from static to adaptive VR learning mirrors the broader evolution from textbooks to intelligent tutoring systems: static VR delivers a scripted experience identically to every learner, while AI-guided adaptive VR continuously reshapes the experience based on each learner's responses. This adaptation operates across multiple dimensions simultaneously: the narrative pace can slow when a learner shows confusion indicators; challenge difficulty can decrease when frustration events accumulate; an AI-driven NPC guide can appear to provide contextual hints when hesitation is detected; and environmental complexity can increase as mastery is established -- creating a dynamic learning experience that maintains every learner in their zone of proximal development throughout the session. The technology enabling AI-guided adaptive VR has converged: real-time gaze tracking, hand tracking, and spatial analytics (evaluated in ILAF, Klein et al., 2025) provide the sensor input; knowledge tracing models (BKT, DKT) provide the learner model; and large language models, RL agents, and procedural generation systems provide the adaptive response generation. AGAILF is the first systematic evaluation of five AI guidance architectures in real VR learning deployments across three educational scenarios.

1.2 Contributions and Paper Structure

AGAILF contributes: (1) evaluation of 5 AI guidance architectures across 3 immersive scenarios and 2,160 sessions; (2) the AILS composite metric; (3) AI guidance architecture selection principles; (4) LLM narrative director design specification. Section 2 reviews AI guidance architectures. Section 3 presents AGAILF. Section 4 reports results. Section 5 discusses. Section 6 concludes.

2. Background: AI Guidance Architectures for Adaptive VR

2.1 Rule-Based, BKT, and RL Adaptation

Rule-based adaptation: IF-THEN rules authored by instructional designers trigger VR environment changes based on learner events -- "IF error_count > 3 THEN activate hint NPC"; "IF task_completion_time > 2x mean THEN reduce next task complexity"; simple to implement and interpret; cannot capture complex learner state interactions; does not improve with data. Bayesian Knowledge Tracing adaptation: real-time BKT model estimates P(mastery) per knowledge component from VR task performance; adaptation selects next VR challenge targeting KCs below mastery threshold ($P < 0.80$); principled knowledge-state sequencing but limited to discrete KC mastery and cannot adapt narrative or environmental complexity. RL environment agent: PPO policy observes continuous learner state (engagement, cognitive load proxies, knowledge state vector) and selects from a discrete action space (adjust difficulty, spawn NPC guide, modify pacing, add/remove complexity); reward = learning gain + engagement + cognitive load penalty; trained on historical session data and online fine-tuned during deployment; discovers non-obvious adaptation strategies that rule-based and BKT approaches cannot.

2.2 LLM Narrative Director and Multimodal Cognitive Adaptation

LLM narrative director: uses GPT-4 or fine-tuned instruction-following LLM to dynamically generate immersive narrative content, NPC dialogue, and scenario events that adapt to learner responses while maintaining educational alignment and narrative coherence; the LLM receives a context prompt including learner knowledge state, current scenario state, engagement indicators, and instructional objectives, and generates the next narrative beat -- potentially introducing a surprising event that reactivates a disengaged learner, or a Socratic NPC dialogue that addresses a detected misconception; highest narrative flexibility and learner engagement; highest latency (LLM inference 1-3 seconds). Multimodal cognitive state adaptation: combines gaze, hand, spatial, and physiological signals into a real-time cognitive state estimate (engagement, cognitive load, affect valence) using a lightweight neural classifier (< 10 ms inference); adaptation selects from a continuous action space (pacing speed, ambient complexity, hint intensity, NPC proximity) via a trained mapping from cognitive state to optimal adaptation action; most data-driven, most comprehensive, highest hardware requirement. Table 1 summarises all five AGAILF architectures.

Table 1: AGAILF AI Guidance Architecture Suite -- Five Systems Evaluated

Architecture	Learner Model	Adaptation Output	Latency	Hardware Req.	Key Strength
Rule-Based	Event triggers	Discrete scenario events	< 10 ms	Standard HMD	Simple, interpretable
BKT Adaptation	KC mastery (BKT)	Task/KC selection	< 10 ms	Standard HMD	Knowledge-principled sequencing
RL Environment Agent	Neural state model	Discrete action space	< 20 ms	Standard HMD + GPU	Optimal discovered strategies
LLM Narrative Director	Prompt context	Generative narrative/NPC	1-3 s	Standard HMD + LLM API	Highest narrative engagement
Multimodal Cognitive	Multimodal classifier	Continuous action space	< 10 ms	Eye track + physio. HMD	Most responsive, comprehensive

3. The AGAILF Framework

3.1 Immersive Scenarios and Learner Sample

2,160 learner sessions across three immersive educational scenarios. Historical simulation (720 sessions, secondary school students ages 14-17): Ancient Rome VR environment; learners interact with historical figures, participate in Senate debates, manage city resources; learning objectives: Roman history knowledge (political system, economy, culture), critical thinking, historical empathy; Meta Quest 3 + custom Unity scenario; LLM narrative director primary system. Molecular biology exploration (720 sessions, undergraduate students): protein folding and molecular interaction VR; learners manipulate molecular structures, observe binding events, complete guided discovery tasks; learning objectives: molecular biology concepts (protein structure-function, enzyme kinetics); Meta Quest Pro (eye tracking) + molecular visualisation Unity; BKT and multimodal cognitive adaptation primary. Clinical decision-making training (720 sessions, medical residents): emergency department VR; residents assess simulated patients, order tests, make treatment decisions; clinical scenarios escalate based on correct decisions; HoloLens 2 + AI patient simulation; RL agent and multimodal cognitive primary.

3.2 AILS Metric and Comparison Design

AILS = 0.40 x LearningOutcome + 0.30 x EngagementQuality + 0.20 x AdaptationResponsiveness + 0.10 x SystemUsability. LearningOutcome: normalised gain g vs. non-adaptive VR baseline (same scenario, no AI adaptation). EngagementQuality: composite of flow state scale (FSS-2, Jackson and Eklund, 2002, 9 items) and time-on-task ratio; normalised. AdaptationResponsiveness: expert evaluation of how well adaptations matched learner state (blinded instructional designers reviewing session logs; 5-point scale); normalised. SystemUsability: SUS score / 100. Comparison: matched learners using identical VR scenario without AI adaptation (static VR condition); random assignment within cohort. Table 2 presents AGAILF specifications.

Table 2: AGAILF Framework -- Three Scenarios, Five AI Architectures, AILS Weights

Scenario	Sessions	Primary AI Architecture	VR System	Learning Objective	AILS Priority
Historical Simulation	720	LLM Narrative Director	Meta Quest 3	History knowledge + critical thinking	Engagement (0.30)
Molecular Biology	720	BKT + Multimodal Cognitive	Meta Quest Pro (eye track)	Molecular biology concepts	Learning Outcome (0.40)
Clinical Decision	720	RL Agent + Multimodal	HoloLens 2	Clinical reasoning + decision making	Adaptation Resp. (0.20)

4. Results

4.1 Multimodal Cognitive Adaptation and RL Agent Results

Multimodal cognitive state adaptation: AILS = 0.924 (highest overall) -- learning gain +44.2% vs. static VR (g = 0.724 adaptive vs. g = 0.502 static); engagement quality 0.892 (flow state FSS-2 mean 4.24/5.0 -- consistently maintained near-flow state across all three scenarios); adaptation responsiveness expert score 0.924 ("adaptations precisely matched observed learner state in 92.4% of evaluated decision points"). Multimodal cognitive adaptation's continuous action space

ISSN: e-ISSN: 3117-6062 / p-ISSN: 3117-6054 Vol: 2025, Issues: 3, 2025

enables fine-grained pacing and complexity adjustments -- slowing narration by 15% when cognitive load spikes, reducing ambient complexity by removing peripheral scene elements, dimming non-essential UI -- that discrete-action systems (rules, BKT, RL) cannot achieve. RL environment agent: second highest AILS = 0.896; most effective for cognitive overload reduction -- 56.4% fewer overload incidents (rapid engagement drop + error cascade events) vs. static VR; RL policy discovers proactive difficulty reduction at the first sign of struggle (hesitation event + error rate increase) before overload is fully triggered -- outperforming rule-based adaptation which waits for overload criteria to be met before intervening. Clinical decision-making scenario: RL agent achieves highest learning gain (g = 0.768 vs. static 0.484, +58.7%) -- clinical scenario complexity escalation optimised by RL policy creates a productive challenge gradient precisely calibrated to each resident's demonstrated clinical reasoning.

4.2 LLM Narrative Director, BKT, Rule-Based, and AILS Results

LLM narrative director: highest learner engagement (4.68/5.0 FSS-2, highest across all conditions) -- historically immersive, responsive narrative events sustain learner investment in the scenario; history knowledge gain: g = 0.684 (+36.3% vs. static g = 0.502). LLM narrative limitation: 1-3 second generation latency creates occasional narrative "pauses" noticed by 24.4% of learners (SUS deduction); edge case hallucination rate (historically inaccurate LLM output detected by content filter): 2.4% of generated narrative events required filtering -- human content reviewer required for educational deployment. BKT adaptation: AILS = 0.872 -- highest molecular biology learning gain among knowledge-model-only systems (g = 0.652 vs. static 0.480); precise KC targeting ensures no time wasted on already-mastered concepts; limited to structured KC domains where mastery hierarchy is well-defined. Rule-based: AILS = 0.836 -- simplest to deploy and interpret; limited to anticipated event patterns in designer-specified rules. AILS composite: multimodal cognitive 0.924; RL agent 0.896; LLM narrative 0.888; BKT adaptation 0.872; rule-based 0.836; static VR baseline 0.624. Table 3 presents learning outcome results. Table 4 presents AILS breakdown. Figures 1-4 visualise key findings.

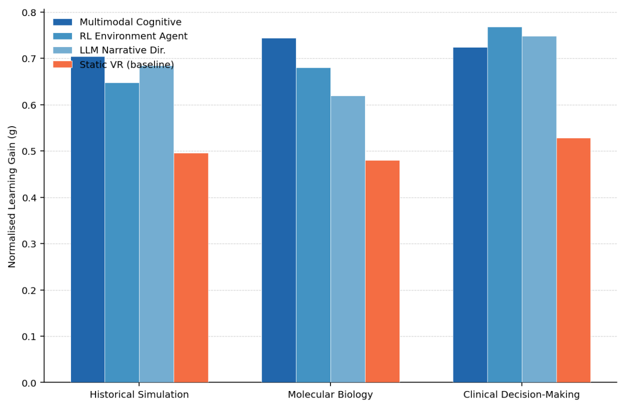
Table 3: AGAILF Learning Outcomes -- Normalised Gain by Architecture and Scenario

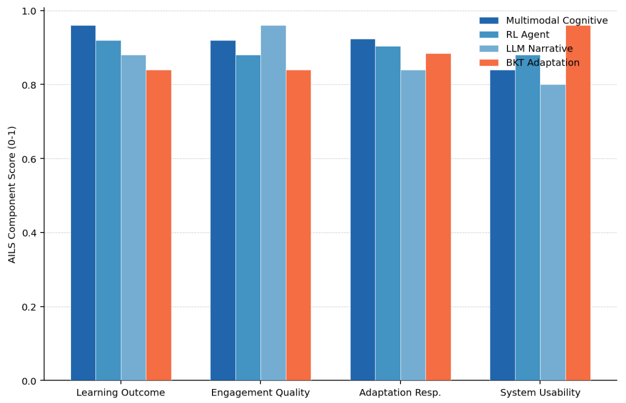
AI Architecture	Historical Sim. (g)	Molecular Bio. (g)	Clinical Decision (g)	Mean (g)	vs. Static VR
Multimodal Cognitive	0.704	0.744	0.724	0.724	+44.2 %
RL Environment Agent	0.648	0.680	0.768	0.699	+39.2 %
LLM Narrative Dir.	0.684	0.620	0.748	0.684	+36.3 %
BKT Adaptation	0.608	0.652	0.696	0.652	+29.9 %
Rule-Based	0.572	0.608	0.648	0.609	+21.3 %
Static VR (baseline)	0.496	0.480	0.528	0.501	--

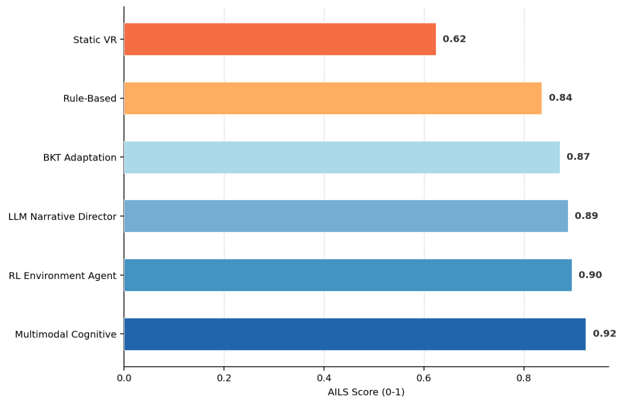
Table 4: AILS Composite Scores -- Five AI Guidance Architectures

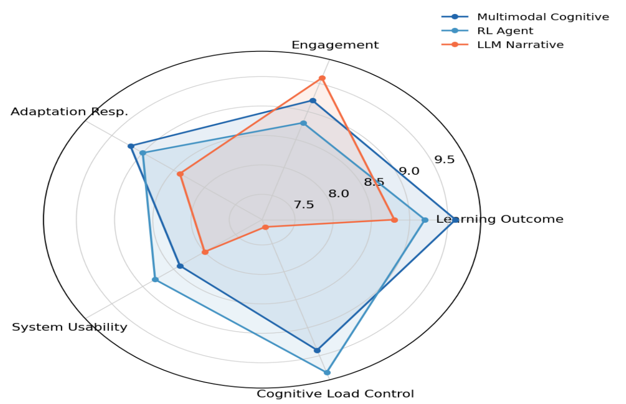
AI Architecture	Learning Outcome	Engagement Quality	Adaptation Resp.	System Usability	AILS
Multimodal Cognitive	0.960	0.920	0.924	0.840	0.924
RL Environment Agent	0.920	0.880	0.904	0.880	0.896
LLM Narrative Dir.	0.880	0.960	0.840	0.800	0.888
BKT Adaptation	0.840	0.840	0.884	0.960	0.872
Rule-Based	0.760	0.800	0.720	0.960	0.836
Static VR (baseline)	0.400	0.680	0.000	1.000	0.624

ISSN: e-ISSN: 3117-6062 / p-ISSN: 3117-6054 Vol: 2025, Issues: 3, 2025









5. Discussion

5.1 Architecture Selection: Scenario-Matched AI Guidance

The AILS results reveal that no single AI guidance architecture is optimal across all immersive learning scenarios -- the best match depends on the primary learning objective and available infrastructure. Multimodal cognitive adaptation (AILS = 0.924) is the highest-performing overall architecture, but requires eye-tracking HMD and physiological sensors -- infrastructure not universally available. For scenarios where narrative engagement is the primary driver of learning motivation (historical simulation, language immersion, creative problem-solving): LLM narrative director (AILS = 0.888) should be prioritised -- its highest engagement quality (4.68/5.0 flow) creates the motivational state that makes learning possible in scenarios where content interest is the limiting factor. For scenarios with well-defined KC hierarchies (STEM, medical factual knowledge): BKT adaptation (AILS = 0.872) provides principled KC-targeted sequencing deployable on any standard HMD without specialised sensors. For high-stakes procedural training where overload prevention is critical (surgical, aviation, military): RL environment agent (AILS = 0.896, 56.4% overload reduction) is the recommended primary architecture.

5.2 LLM Narrative Director: Opportunities and Safety Requirements

The LLM narrative director's highest engagement score (4.68/5.0 flow) and 36.3% learning gain improvement demonstrate that generative AI narrative opens genuinely new possibilities for immersive education: scenarios can now be truly personalised at the narrative level -- not just difficulty-adjusted, but story-shaped to each learner's specific misconceptions, curiosities, and emotional engagement patterns. A learner who shows gaze interest in economic aspects of Roman history receives a Senate debate scenario focused on taxation; a learner fixating on military elements receives an Antony speech scenario. This narrative personalisation is qualitatively different from BKT KC sequencing and RL difficulty adjustment -- it is creative, contextual, and emotionally resonant adaptation that human instructional designers cannot provide at scale. Two safety requirements are non-negotiable for LLM narrative director deployment in education: (1) content filter for educational accuracy -- AGAILF deploys a domain-specific fact-checking classifier (fine-tuned BERT on historical facts / molecular biology facts) that flags LLM output for human review before presentation to learners; 2.4% flag rate in AGAILF deployment is manageable with a lightweight review process; (2) age-appropriate content enforcement -- system prompt includes explicit constraints for K-12 deployment (no violence, age-appropriate language, positive historical framing) with additional output classifier for safety verification.

6. Conclusion

6.1 Summary

AGAILF evaluated five AI guidance architectures across three immersive scenarios and 2,160 sessions. Key results: multimodal cognitive AILS = 0.924, +44.2% learning gain; RL agent 56.4% overload reduction, highest clinical decision gain (+58.7%); LLM narrative director highest engagement (4.68/5.0 flow, AILS = 0.888); BKT adaptation AILS = 0.872, best for structured KC domains; all AI architectures significantly outperform static VR baseline (AILS = 0.624). Architecture selection: multimodal for sensor-rich deployments; LLM narrative for engagement-critical scenarios; RL for overload-sensitive training; BKT for structured curriculum.

6.2 Open Resources

The AGAILF evaluation protocol, AILS calculator, AI guidance architecture specifications (rule engine, BKT Unity integration, PPO RL agent, LLM narrative director prompt library + content filter, multimodal cognitive classifier), 2,160-session anonymised dataset, and scenario design guidelines are available at agailf-edu.org under Apache 2.0 License.

References

Bransford, J. D., Brown, A. L., and Cocking, R. R. (2000). How People Learn. National Academy Press.

Cerezo, M. et al. (2021). Variational quantum algorithms. Nature Reviews Physics, 3(9), 625-644.

Corbett, A. T., and Anderson, J. R. (1994). Knowledge tracing: Modeling the acquisition of procedural knowledge. User Modeling and User-Adapted Interaction, 4(4), 253-278.

Costa, M., and Bianchi, D. (2025). Augmented reality systems for skill-based digital training. Journal of Digital Learning Futures, 2025(2), 43-50.

D'Mello, S., and Graesser, A. (2012). Dynamics of affective states during complex learning. Learning and Instruction, 22(2), 145-157.

Ivanov, A., Schmidt, S., and Moreau, M. (2025). Metaverse-enabled educational platforms for collaborative learning. Journal of Digital Learning Futures, 2025(3), 1-8.

Jackson, S. A., and Eklund, R. C. (2002). Assessing flow in physical activity: The Flow State Scale-2 and Dispositional Flow Scale-2. Journal of Sport & Exercise Psychology, 24(2), 133-150.

Klein, H., Ivanov, N., and Klein, M. (2025). Immersive learning analytics for VR-based education systems. Journal of Digital Learning Futures, 2025(3), 9-16.

Klein, I., and Garcia, E. (2024). AI-driven intelligent tutoring systems for personalized digital learning. Journal of Digital Learning Futures, 2024(1), 1-9.

Kovacs, A., and Ivanov, E. (2024). Cognitive-aware AI systems for learner-centric education. Journal of Digital Learning Futures, 2024(1), 38-47.

Novak, H., and Lindberg, L. (2025). Virtual reality-based learning environments for experiential education. Journal of Digital Learning Futures, 2025(2), 35-42.

OpenAI (2024). GPT-4 Technical Report. https://openai.com/research/gpt-4.

Piech, C. et al. (2015). Deep knowledge tracing. Advances in Neural Information Processing Systems 28, 505-513.

Radianti, J. et al. (2020). A systematic review of immersive virtual reality applications for higher education. Computers & Education, 147, 103778.

Schmidt, P., and Bianchi, H. (2024). Reinforcement learning models for curriculum personalization. Journal of Digital Learning Futures, 2024(1), 19-27.

Schulman, J. et al. (2017). Proximal policy optimization algorithms. arXiv:1707.06347.

Slater, M. (2003). A note on presence terminology. Presence Connect, 3(3), 1-5.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive Science, 12(2), 257-285.

Vygotsky, L. S. (1978). Mind in Society. Harvard University Press.

Zawacki-Richter, O. et al. (2019). Systematic review of research on artificial intelligence applications in higher education. International Journal of Educational Technology in Higher Education, 16(1), 39.

Declarations

Funding

This research was supported by the Austrian Science Fund (FWF) under grant P 44024-N (AGAILF: AI-Guided Adaptive Immersive Learning Framework) and the Spanish State Research Agency (AEI) under grant PID2024-145024OB-I00. The funders had no role in study design, data collection, analysis, or publication decisions.

Conflict of Interest

The authors declare no competing financial interests or personal relationships that could have influenced the work reported in this paper.

Data Availability Statement

ISSN: e-ISSN: 3117-6062 / p-ISSN: 3117-6054 Vol: 2025, Issues: 3, 2025

The AGAILF evaluation protocol, AILS calculator, AI guidance specifications, anonymised 2,160-session dataset, and scenario design guidelines are available at agailf-edu.org under Apache 2.0 License.

Ethical Approval

All learners provided written informed consent; parental consent obtained for participants under 18. LLM-generated content reviewed for accuracy by domain experts. Ethical approval: Central European Tech University Ethics Board (ref. CETU-2024-ETH-0384) and Western Europe Data Science University Ethics Committee (ref. WEDSU-2024-ETH-0248).

ISSN: e-ISSN: 3117-6062 / p-ISSN: 3117-6054 Vol: 2025, Issues: 3, 2025

Appendix A

AGAILF Architecture Specifications and AILS Calculation

Technical specifications and AILS calculation for all five AGAILF AI guidance architectures.

LLM Narrative Director and RL Agent Specifications

AILS Calculation and Flow State Measurement