Open Science
BRAZIL
DOI: https://doi.org/10.37497/opsbrazil.28

PREPRINT

DESIGNING AGENTIC AI SYSTEMS FOR EDUCATION, FROM STATIC TOOLS TO AUTONOMOUS LEARNING AGENTS

Aeromedical Operational Methodology During the COVID-19 Pandemic: Evidence from High-Complexity Missions in Remote Regions of Brazil

JULIANE URBANO
Vínculo institucional: Independent Research
Pesquisadora independente com foco em inteligência artificial, arquitetura de sistemas e transformação digital aplicada à educação. Seus estudos concentram-se no desenvolvimento de sistemas baseados em agentes autônomos, governança de IA e integração entre tecnologia e ambientes educacionais..
E-mail: juliane.urbano@gmail.com

ABSTRACT | The evolution of artificial intelligence from static models to Agentic AI systems introduces a fundamental shift in how educational technologies are designed, deployed, and utilized. Unlike traditional AI tools that respond to user inputs in a linear fashion, Agentic AI systems are capable of autonomous reasoning, planning, and execution of multi step tasks, enabling new paradigms in personalized and adaptive learning. Despite rapid technological progress, most educational implementations of AI remain limited to assistive or reactive systems, failing to leverage the full potential of autonomous agents. This article explores the architectural and operational foundations required to design effective Agentic AI systems for education. It proposes a layered framework consisting of cognitive orchestration, pedagogical alignment, feedback driven adaptation, and ethical control layers. Drawing on recent advancements in multi agent systems, large language model orchestration, and reinforcement learning, the paper demonstrates how educational environments can transition from static digital tools to dynamic, self improving learning ecosystems. The study further examines real world constraints including data quality, model hallucination risks, latency in feedback loops, and the challenge of aligning autonomous agents with pedagogical objectives. A reference architecture is proposed for integrating Agentic AI into learning management systems, enabling continuous personalization, automated tutoring, and adaptive curriculum generation. The findings suggest that the primary limitation in educational AI adoption is not model capability, but system design and orchestration. The article concludes that the future of AI in education will depend on the ability to design ecosystems of cooperating agents rather than isolated tools, shifting the focus from AI usage to AI system engineering.

Keywords | Agentic AI, Educational Systems, Multi Agent Architecture, Adaptive Learning, AI Orchestration, Autonomous Systems, Learning Technology

Submetido: 07/05/2026
Postado: 07/05/2026
e-ISSN: 2764-9822

Endereço de correspondência:
Juliane Urbano – E-mail:
juliane.urbano@gmail.com

1 INTRODUCTION

Artificial intelligence in education has historically evolved through incremental stages, beginning with rule based systems, progressing through machine learning models, and more recently adopting large language models for content generation and tutoring support. However, the emergence of Agentic AI introduces a discontinuity rather than an incremental improvement, shifting the paradigm from passive tools to autonomous systems capable of goal oriented behavior.

Traditional educational AI systems operate on a request response architecture, where the user initiates an interaction and the system provides a bounded output. In contrast, Agentic AI systems are designed to decompose goals, plan sequences of actions, interact with external tools, and adjust behavior dynamically based on feedback. This capability fundamentally alters the design space of educational technologies.

Despite this evolution, most educational deployments remain constrained by legacy system architectures that do not support autonomy, persistence, or multi step reasoning. As a result, there is a growing gap between what Agentic AI systems are technically capable of and how they are actually implemented in educational contexts.

This article addresses this gap by proposing a systems level framework for designing Agentic AI in education, focusing on architecture, orchestration, and feedback mechanisms rather than isolated model performance.

The contributions of this work are threefold. First, it defines a conceptual distinction between assistive AI and agentic AI within educational environments. Second, it proposes a layered architecture for implementing autonomous learning systems. Third, it identifies key technical and pedagogical constraints that must be addressed for successful deployment.

2 FROM ASSISTIVE AI TO AGENTIC LEARNING SYSTEMS

Most current educational AI applications can be classified as assistive systems. These include chatbots, automated grading tools, recommendation engines, and content generators. While useful, these systems operate within a narrow input output paradigm and lack persistence across tasks.

Agentic AI systems, in contrast, exhibit four core capabilities: goal decomposition, planning, tool usage, and iterative self correction. These capabilities allow them to operate as semi autonomous actors within digital environments.

In educational contexts, this enables a transition from static learning support tools to dynamic learning companions capable of adapting to student behavior over time. For example, instead of simply answering a question, an agentic tutor can identify knowledge gaps, generate a learning plan, retrieve external resources, and evaluate student progress continuously.

However, the implementation of such systems requires a fundamental rethinking of educational infrastructure. Most learning management systems were not designed for persistent agents, stateful interaction, or multi step reasoning processes.

3 REFERENCE ARCHITECTURE FOR EDUCATIONAL AGENTIC AI

A functional Agentic AI system for education can be conceptualized as a layered architecture composed of four interdependent layers.

3.1 Cognitive Orchestration Layer

This layer is responsible for reasoning, planning, and decision making. It typically consists of large language models combined with planning algorithms that enable decomposition of educational objectives into actionable steps.

3.2 Pedagogical Alignment Layer

This layer ensures that agent behavior aligns with educational goals, curriculum standards, and learning outcomes. It acts as a constraint system that prevents optimization purely for efficiency or correctness without pedagogical relevance.

3.3 Feedback and Adaptation Layer

This layer continuously collects data from learner interactions, including performance metrics, engagement signals, and error patterns. It enables iterative improvement of both content delivery and instructional strategy.

3.4 Ethical and Control Layer

This layer enforces safety, transparency, and compliance constraints. It includes bias detection, hallucination mitigation, and governance rules aligned with educational policies and data protection regulations.

Together, these layers form a closed loop system capable of continuous adaptation and controlled autonomy.

4 KEY TECHNICAL CHALLENGES

Despite significant advances, several technical challenges limit the deployment of Agentic AI in education.

One of the primary issues is hallucination in large language models, which can lead to incorrect or misleading educational content. In agentic systems, this risk is amplified due to multi step reasoning chains.

Another challenge is latency in feedback loops. Effective learning systems require near real time adaptation, but current architectures often introduce delays that reduce responsiveness.

Data quality is also a critical limitation, as many educational datasets are incomplete, biased, or non standardized. This affects the ability of agents to generate accurate learning models.

Finally, alignment between autonomous agents and pedagogical objectives remains an open problem, particularly when optimizing for engagement may conflict with long term learning outcomes.

5 DESIGN PRINCIPLES FOR EDUCATIONAL AGENTS

Based on current limitations, several design principles emerge for effective Agentic AI systems in education.

First, autonomy should be constrained, not absolute. Educational agents must operate within clearly defined pedagogical boundaries.

Second, transparency is essential. Students and educators must understand how decisions are made by AI systems.

Third, human oversight must remain embedded in the system architecture, particularly for high impact educational decisions.

Fourth, systems must be designed for continuous learning, both at the model level and at the pedagogical level.

6 DISCUSSION

The transition from assistive AI to Agentic AI in education represents not only a technological shift but a systemic redesign of how learning environments function. Rather than replacing educators, agentic systems extend their capabilities by automating repetitive tasks and enabling personalized learning at scale.

However, the success of these systems depends less on model sophistication and more on system design, orchestration, and integration within educational ecosystems.

7 CONCLUSION

Agentic AI introduces a new paradigm for educational technology, shifting the focus from static tools to autonomous learning systems. While the technical foundations are rapidly advancing, the primary barrier to adoption lies in system architecture and educational integration.

The future of AI in education will be defined not by isolated models, but by ecosystems of coordinated agents capable of supporting adaptive, personalized, and scalable learning experiences.

REFERENCES

Acemoglu, D., & Johnson, S. (2024). Technology and labor markets. NBER.

Brynjolfsson, E., & Hitt, L. M. (1996). Management Science.

Christensen, C. M. (1997). The Innovator’s Dilemma.

Deloitte. (2023). AI Adoption and Value Creation.

Gartner. (2025). AI Hype Cycle.

McKinsey & Company. (2025). State of AI Report.

MIT Sloan. (2025). AI Agents in Collaborative Systems.

NIST. (2023). AI Risk Management Framework.

Russell, S., & Norvig, P. (2021). Artificial Intelligence: A Modern Approach.

Standish Group. (1995–2020). CHAOS Reports.

UNESCO. (2021). Ethics of Artificial Intelligence.

World Economic Forum. (2025). Future of Learning Systems.