

Slidecaster: Human-AI Collaborative Authoring of Pedagogically Grounded Micro-Learning Videos

*Junbo Koh¹, *Hoetaek Rah², Hyeseong Kwon³, Gyuri Byun³, and Cheolil Lim³

¹ Learning Sciences Research Institute, SNU, Seoul, South Korea
gtkobo92@snu.ac.kr

² Indischool, Seoul, South Korea
hoetaek@indischool.com

³ Seoul National University, Seoul, South Korea
{comet, gl3013, chlim}@snu.ac.kr

Abstract. Producing pedagogically grounded micro-learning videos demands both deep domain expertise and instructional design competency. We present Slidecaster, a human-AI collaborative authoring system that enables subject-matter experts (SMEs) to co-create pedagogically grounded micro-learning videos from domain documents, without requiring instructional design expertise. Grounded in Mayer's Cognitive Theory of Multimedia Learning and Sweller's Cognitive Load Theory, and scaffolded by the ADDIE instructional design model, Slidecaster structures lecture production into three phases—Content Acquisition, Pedagogical Intent Negotiation, and Instructional Design Execution—each with explicit SME review touchpoints. A negotiation loop ensures AI recommendations remain editable, preserving SME agency throughout the process.

Keywords: Micro-learning, Human-AI Collaboration, Instructional Design, Cognitive Theory of Multimedia Learning, Cognitive Load Theory.

1 Introduction

Micro-learning delivers complex information in bite-sized modules through short videos, reducing cognitive overload and improving retention (Sweller, 1988). However, producing high-quality content requires both deep subject-matter expertise and instructional-design competencies, making content quality heavily dependent on individual capabilities (Dick et al., 2015).

While LLMs have been shown to support rapid content development (Leiker et al., 2023) and reduce preparation time (Lan & Chen, 2024), most existing tools rely on one-shot automation: the same prompt yields inconsistent emphasis and flow across iterations, and pedagogical intentions are difficult to reflect. Prior work argues that instructors should review and revise AI-generated recommendations to maintain their agency (Nazaretsky et al., 2022) and that AI should support rather than replace teachers' decision-making (Kozan, 2025).

This paper presents Slidecaster, a human-AI collaborative authoring system where the Subject Matter Expert (SME) and AI act as a team. Its contributions are: (i) a

*These authors contributed equally to this work.

human-in-the-loop collaboration model preserving SME agency with instructional-design support, (ii) a scalable content-transformation pipeline applicable across disciplines, and (iii) a development environment where AI supports instructional-design decision-making while preserving pedagogical quality.

2 Technical Overview

A video demo of Slidecaster is available at: https://youtu.be/cJIX0uN5N_A. The system decomposes lecture production into three sequential phases with explicit SME review touchpoints (Figure 1), grounded in Mayer's (2009) Cognitive Theory of Multimedia Learning (CTML) and Sweller's (1988) Cognitive Load Theory (CLT), and scaffolded by the ADDIE instructional design model.

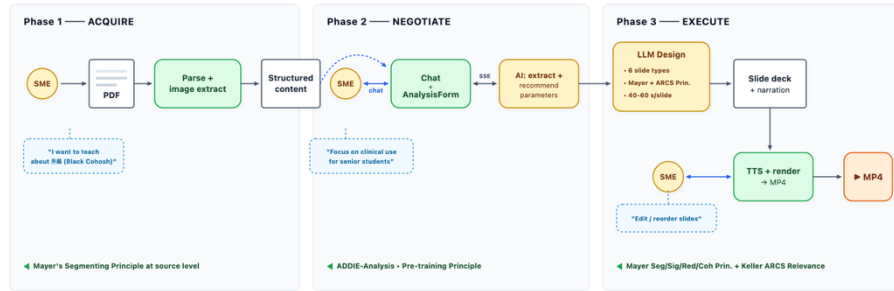


Fig. 1. Overview of Slidecaster's human-AI co-creation pipeline.

2.1 Content Acquisition

The SME uploads a domain-specific PDF, and the system performs structural extraction. The Upstage Document Parse API converts the PDF into HTML and Markdown formats, while the Docling module extracts embedded images. The resulting hierarchical tree—segmented into titles, paragraphs, tables, and images—provides granular units for subsequent instructional decisions, moving beyond flat text extraction to create pedagogically addressable chunks amenable to Mayer's Segmentation Principle.

2.2 Pedagogical Intent Negotiation

The SME articulates instructional parameters through a chat interface paired with the AnalysisForm. Each utterance triggers a Server-Sent Events call to an LLM that extracts design variables (learning objectives, target audience, prerequisite knowledge, keywords, duration) and populates the form with editable recommendations in real time. The stage concludes with formal confirmation of negotiated objectives and learner profile, with AI serving an advisory role while final decisions remain with the SME (Nazaretsky et al., 2022; Kozan, 2025). This phase operationalizes ADDIE's Analysis stage and instantiates Mayer's Pre-training Principle.

2.3 Instructional Design Execution

The final phase positions the SME as reviewer while the AI implements negotiated parameters, shaped by three design commitments.

Pedagogically grounded slide vocabulary. Modules are constructed from six slide types aligned to cognitive operations (e.g., discrimination, sequencing, hierarchy, conceptual mapping). The SME retains full control to add, modify, delete, or reorder slides.

Cognitive load management. A "one slide, one core concept" rule pairs with narration of 40–60 seconds per slide, ensuring total narration never exceeds the negotiated duration. Slide text and narration are deliberately disjoint—slides display key phrases, narration delivers context—instantiating Mayer's Redundancy Principle. Narrations are synthesized with Qwen3-TTS and aligned with transitions for dual-channel processing.

Visual coherence. Slides are rendered via HTML5 Canvas, streamed as JPEG frames over WebSocket, and encoded as MP4 by FFmpeg. Rendering reuses the frontend design system (Tailwind CSS), enforcing visual consistency across a content series without a separate style guide.

3 Originality and Strengths

(1) Human–AI co-creation. Unlike one-shot tools such as GenSpark or NotebookLM, Slidecaster preserves SME agency throughout the entire process. AI extractions are presented as editable recommendations in the AnalysisForm, and the negotiation loop ensures the SME—not the model's stochastic tendencies—retains final authority.

(2) Scalability and generalization. Domain-specific elements (e.g., Hangul–Hanja notation, efficacy-centered composition) are modularized at the prompt layer rather than hard-coded. The core workflow, slide types, chunking strategy, and review structure remain domain-independent, allowing transfer to adjacent disciplines without architectural redesign.

References

1. Dick, W., Carey, L., & Carey, J. O. (2015). *The systematic design of instruction* (8th ed.). Pearson.
2. Kozan, K. (2025). Instructional designers' integration of generative AI. *Education Sciences*, 15(9), 1133.
3. Lan, Y. J., & Chen, N. S. (2024). Teachers' agency in the era of LLM and generative AI. *Educational Technology & Society*, 27(1), I–XVIII.
4. Leiker, D., Finnigan, S., Gyllen, A. R., & Cukurova, M. (2023). Prototyping the use of LLMs for adult learning content creation at scale. arXiv:2306.01815.
5. Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press.
6. Nazaretsky, T., Ariely, M., Cukurova, M., & Alexandron, G. (2022). Teachers' trust in AI-powered educational technology. *British Journal of Educational Technology*, 53(4), 914–931.
7. Sweller, J. (1988). Cognitive load during problem solving. *Cognitive Science*, 12(2), 257–285.