Democratizing Pedagogical Content Creation: A Multi-Agent YouTube AI Framework for Empowering Indian Educators in Higher Education

Dhiyanesh Sidhaiyan
ACV Auctions Inc, Chennai, India

Divyansh Bhatia
eBay Inc, Bangalore, India

Abstract

India's higher education system serves over 43 million students, yet a critical gap persists between NEP 2020's mandate for technology-integrated instruction and ground-level pedagogical practice. We present the YouTube AI Agent Framework—an open-source, locally-deployable multi-agent pipeline that converts YouTube videos, academic abstracts, and transcripts into animated instructional videos aligned with Mayer's Cognitive Theory of Multimedia Learning (CTML), with multilingual narration. The architecture comprises seven interoperating agents: Transcript, Script, Animation (Manim CE), TTS, Video, Evaluator, and Fixer. Key contributions include: (1) a three-tier animation generation hierarchy—parameterized templates → dynamic LLM generation → pre-built fallback—achieving 100% render success; (2) a five-pass auto-layout safety post-processor eliminating off-screen rendering failures; (3) a Self-Refine quality loop (Generate → Evaluate → Feedback → Refine, averaging 2.1 iterations); (4) a 15-type parameterized scene template library; and (5) a three-tier TTS stack supporting 9 scheduled Indian languages with fully offline fallback. Evaluated with N=18 faculty across three institutional tiers, Pathway A achieves 98.9% production time reduction (14.3 days → 3.2 h), 89.6% CTML compliance (vs. 35.0% manual baseline, +157%), SUS 82.4, and self-efficacy gain d=2.1 ($p<0.001$).

Keywords

Multi-agent AI systems, educational video generation, Mayer's CTML, Manim animation, Self-Refine loop, multilingual TTS, Indian higher education, NEP 2020, local LLM inference, open-source EdTech.