

The background features a dark grey field. On the right side, there is a large, colorful gradient shape that tapers to a point, transitioning from green at the top to yellow in the middle and red at the bottom. In the lower right corner, there is a faint, light grey grid pattern of squares.

EDUMuse

EduMUSE is a comprehensive AI tutoring system that transforms study materials into personalized learning experiences. The platform combines document processing, knowledge retrieval, and interactive tutoring through a modular multi-agent architecture.

Meet the team



Soren Larsen



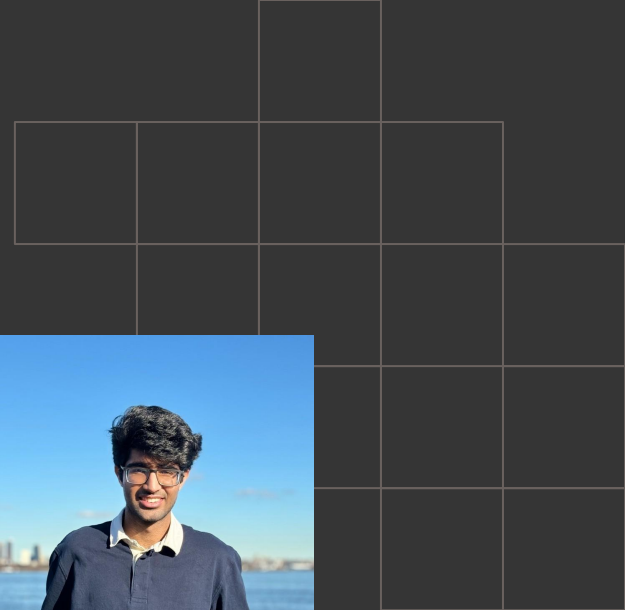
Ishika Kulkarni



Yousuf Golding



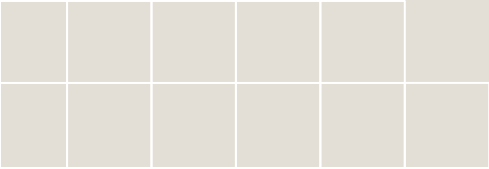
Mudit Arora



Contents

1. Aim
2. Problem Statement
3. Existing Gaps
4. How We Tried to Bridge Gaps
5. Why Use Agents for This
6. Why Use Multi-Agent Framework

7. Frameworks Explored
8. Framework Chosen
9. Current Pipeline
10. Frontend Pipeline
11. Existing Features
12. Putting it all together
7. Frameworks Explored
8. Framework Chosen



Aim

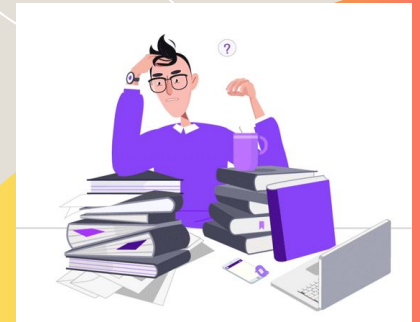
EDUMuse: Multimodal AI Tutor

An educational engine that supports personalized, adaptive, and interactive learning through AI-driven agents.



Problem Statement

- AI tools like Chat GPT offer static answers, **not structured learning**
- Students struggle with **self-pacing and focus**
- No system bridges summarization, quiz generation, grading, and feedback into **one pipeline**



Existing Gaps

- No support for **personalized** study pace
- **Minimal** multimodal interaction (e.g. audio, diagrams)
- Lack of **automatic assessments** and feedback
- **Fragmented** tools, not integrated workflows



How We Tried to Bridge Gaps

We created a modular pipeline with:

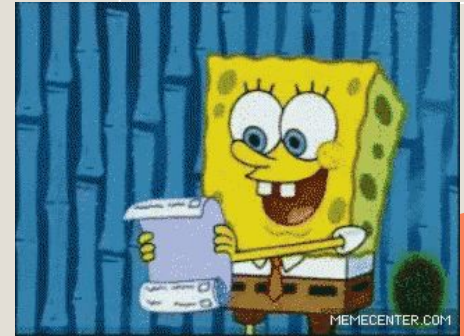
- Knowledge Retrieval
- Note summarization
- Podcast-style discussions
- Practice assessments
- Question Answering

suddenly im learning



How we went about it

1. Feature Identification & Workflow Design
2. Agent Pipeline Development
3. Agent Orchestration Logic
4. Evaluation and Testing
5. Modularity, Iteration & Scalability

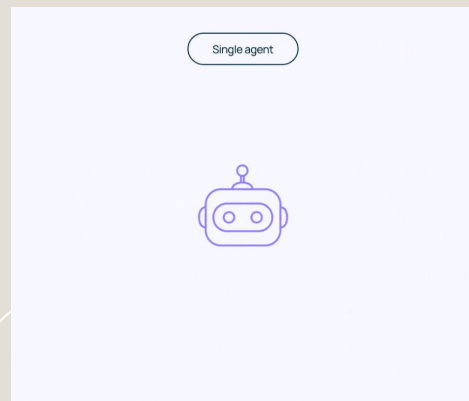


Why Use Agents for This




- Each task is a **distinct cognitive function**
- Summarization \neq grading \neq audio synthesis
- Agents allow modular **specialization** per task
- Easier to evaluate and debug
- Aligns with human roles: tutor, evaluator, explainer
- Reduces complexity within each module

Why Use Multi-Agent Framework

- Loose coupling
- Reusability
- **Parallel** execution
- **Independence** across workflows
- Easier to **maintain** and extend
- If we improve one agent (e.g. assessment), others remain untouched
- Enables future **plug-and-play** upgrades (e.g., swap summarizer model)



Frameworks Explored

	 LangGraph	 CrewAI	 Autogen
Developer Experience	High flexibility, steep learning curve, limited pre-built components	Easy to use, good documentation, fast setup	Moderate ease of use, well-structured documentation
Agent Capabilities	Highly customizable , strong task execution and tool usage, supports complex agent hierarchies	Stable task planning and execution, good collaboration through shared memory and messaging	Strong multi-agent collaboration , excellent task planning and execution
Enterprise Features	Requires additional development for enterprise features	Limited, with advanced security features only in paid version	Good enterprise support, including security and logging features
Strengths	Flexible and powerful, supports custom behaviors at any node	Great team-based agent collaboration, easy configuration by non-technical users	High extensibility, supports complex workflows, strong tooling support
Weaknesses	Hard to learn, requires building many components from scratch	Limited extensibility, poor logging capabilities	Requires familiarity with the ecosystem, initial setup complexity
Best Use Cases	Complex agent systems with custom interaction and tool usage patterns	Rapid development of simple agent applications or apps configured by non-technical users	Advanced multi-agent systems, enterprise-grade applications with robust security and logging

Why we chose CrewAI

Modular Agent Design

Flexible Orchestration

Plug-and-Play Architecture

Loose Coupling Between
Modules

Parallel and Scalable Execution

Mimics Human Roles

Better Debugging and
Evaluation

Existing Features

User Interaction

- Upload PDFs via web interface
- Highlight text to trigger AI modules
- On-demand, user-controlled workflows

Summarization

- Concise summaries from highlighted text
- Powered by GPT-4 / Gemini via CrewAI
- Shared base for all learning flows

Quiz Generation

- Auto-generates:
 - QA pairs, MCQs
 - Higher-order & situational questions
- Structured output using CrewAI prompts
- Downloadable quiz preview

Existing Features

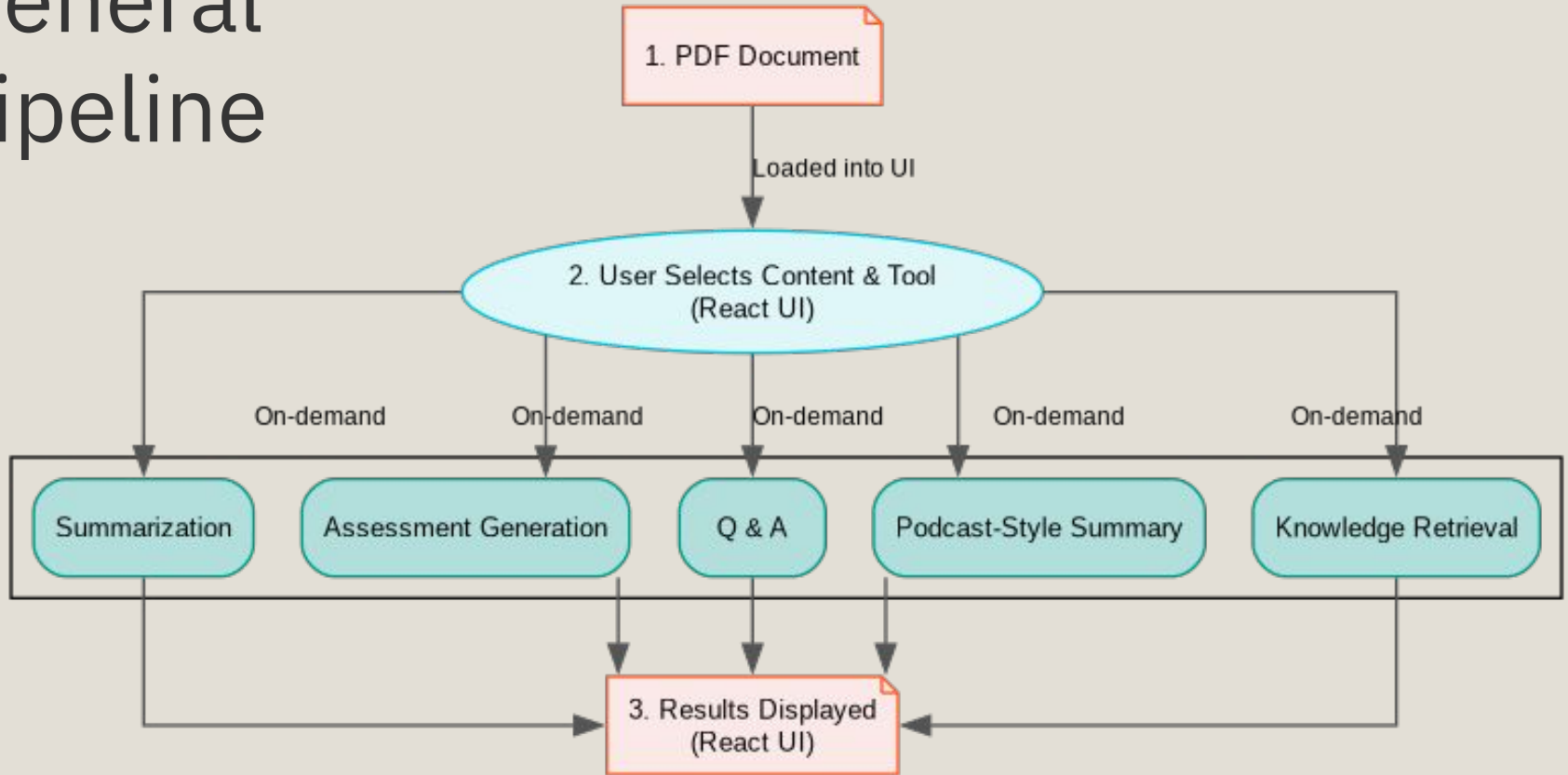
Knowledge Retrieval

- Finds academic sources and papers to support learning on user-selected topics
- Uses 3 retrieval modes:
 - Web Search — retrieves live academic content and credible sources via search APIs
 - LLM Knowledge — synthesizes foundational papers and key concepts from model training data
 - Hybrid Retrieval — combines both for broad + current coverage

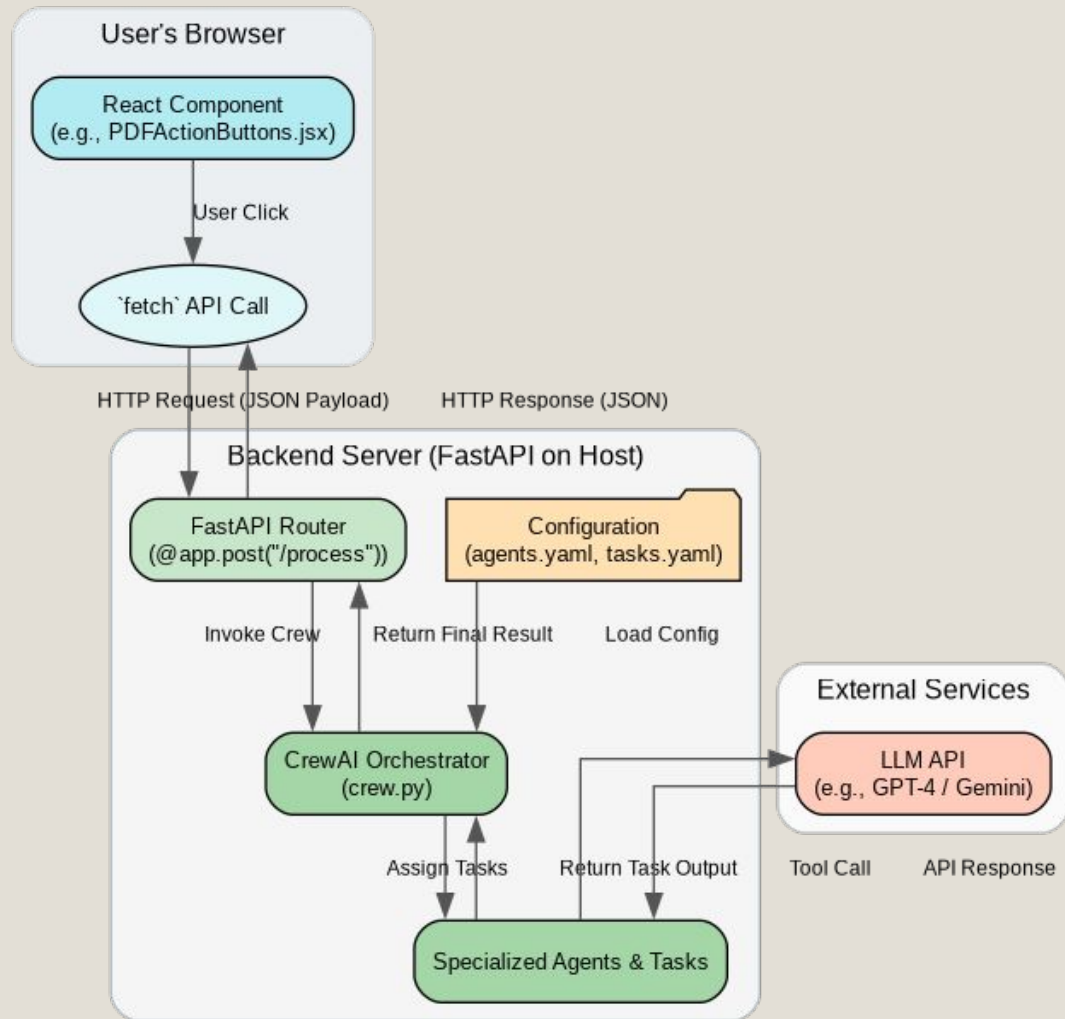
Podcast-Style TTS

- Converts summaries into podcasts
- Uses OpenAI and Eleven Labs API
- Supports auditory + mobile learning

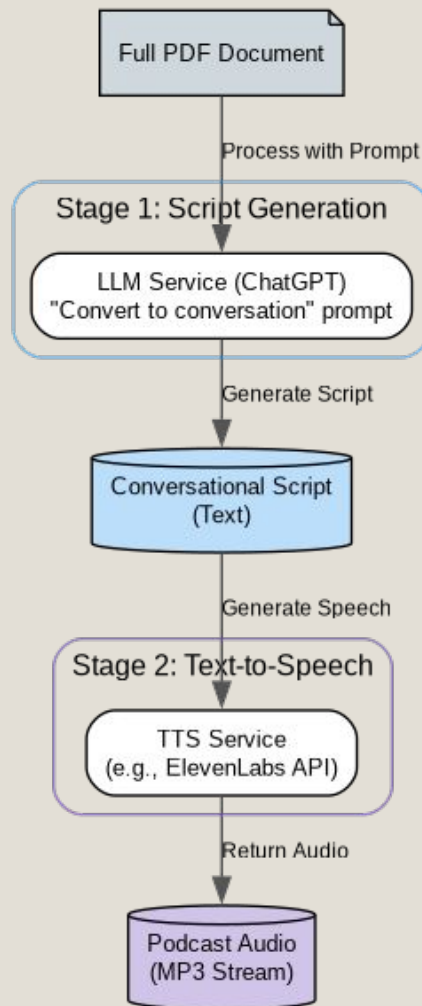
General Pipeline



Front-End Pipeline



TTS Pipeline



Putting it all together!
Demo time!



Question Answer PDF

1. Question-Answer Pairs:

1. What is photosynthesis?
 - Photosynthesis is the process by which green plants, algae, and some bacteria convert sunlight
2. What is the primary source of energy for photosynthesis?
 - Sunlight is the primary source of energy for photosynthesis.
3. Which pigment in plants absorbs sunlight for photosynthesis?
 - Chlorophyll is the pigment in plants that absorbs sunlight for photosynthesis.
4. What gas is taken in by plants during photosynthesis?
 - Carbon dioxide is taken in by plants during photosynthesis.
5. What is the main product of photosynthesis?
 - Glucose is the main product of photosynthesis.
6. Where does photosynthesis mainly occur in a plant cell?
 - Photosynthesis mainly occurs in the chloroplasts of plant cells.
7. What gas is released as a byproduct of photosynthesis?
 - Oxygen is released as a byproduct of photosynthesis.
8. How does water reach the leaves for photosynthesis?
 - Water is absorbed by the plant's roots and transported to the leaves through the xylem.
9. What role does sunlight play in the process of photosynthesis?
 - Sunlight provides the energy needed to drive the chemical reactions of photosynthesis.
10. How is glucose used by plants after it is produced in photosynthesis?
 - Glucose is used by plants as a source of energy for growth and other metabolic processes.

Podcast Demo Link



Knowledge Retrieval Agent Evaluation

Setup:

We evaluated 3 retrieval approaches:

- Web Search: live web API search (SerperDevTool)
- LLM Knowledge: pure LLM synthesis from training data
- Hybrid Retrieval: combines Web Search + LLM synthesis

Each agent was run via CrewAI flows with consistent prompts

Test topics:

- Machine Learning Transformers
- Quantum Computing Algorithms

Evaluation Method:

Manual scoring on 5 criteria:

- Relevance (30%)
- Credibility (25%)
- Coverage (20%)
- Recency (15%)
- Accessibility (10%)

Additional metrics:

- Processing time
- Number of sources returned
- Content length (tokens)

Knowledge Retrieval Results

Flow	Average Time	Total Score	Sources	Content Length
Web Search	11.64	7.5	4.0	1744.5
LLM Knowledge	17.16	8.05	4.0	3722.0
Hybrid Retrieval	19.83	9.15	5.5	3206.5

Insights:

- Hybrid Retrieval delivered the best overall performance, balancing current and foundational knowledge.
- LLM Knowledge was strong on depth and clarity but lacked recency.
- Web Search was fast and current but provided shallower coverage.

Summary Evaluation and Results

We utilized the Rouge and Bert score metrics to evaluate summary agent. They are measuring the lexical overlap and semantic quality of the ai-generated summaries compared with pre-defined summaries.

We utilized the same test topics from the knowledge retrieval evaluations.

Level	Rouge-1	Rouge-2	Rouge-L	Bert
Beginner	0.463	0.163	0.257	0.826
Intermediate	0.213	0.025	0.107	0.812
Advanced	0.190	0.035	0.129	0.805

Human Evaluation

- Since we wanted actual feedback to make our agents better, we asked volunteers to rate on specific criteria.
- 4 criteria for each agent (scale of 1-5):
 - **Note Summarization Agent:** Usefulness, Clarity, Coverage, Conciseness
 - **Podcast-style Discussion Agent:** Engagement, Clarity, Naturalness, Relevance
 - **Practice Assessment Agent:** Helpfulness, Correctness, Difficulty, Coverage

Result

Thanks to all the volunteers, these are the avg scores that we received:

- **Note Summarization Agent:**

Usefulness (4), Clarity (3.5), Coverage (3), Conciseness (4)

- **Podcast-style Discussion Agent:**

Engagement (4.1), Clarity (4.5), Naturalness (3.2), Relevance (4.4)

- **Practice Assessment Agent:**

Helpfulness (4), Correctness (3.5), Difficulty (4), Coverage (4)

Interesting Findings

- Our main focus for this project consisted of research-papers (in pdf format), which had a lot of texts, we tried on multiple different variety of content.

At 0:40 she says 't' and 'd' are interdental sounds, but they're alveolar. Seems like it might not have processed the spatial organization on this slide? 🤖

10:19 AM

• **place, manner, voicing:**

	stop	fricative
bilabial	[b], [p]	
labiodental		[v], [f]
interdental		[ð], [θ]
<u>alveolar</u>	[d], [t]	[z], [s]
postalveolar		[ʒ], [ʃ]
palatal		
velar	[g], [k]	
glottal	[ʔ]	[h]

10:20 AM

0:53 he says some gibberish lol

10:22 AM

the summarization is decent, the question assessment is okay in terms of quality

1:45 PM

Future Work

- Adding different difficulty levels
- Handwritten Answer Grading
- Grading Assistant Agent
- Session Packaging
- Enhanced UI & Accessibility
- Multimodal Content Understanding
- Longer podcast with more natural voice
- Evaluation for Question Answering
- Evaluation and User study of Frontend UI



Any questions?
Ask away!





Thank you