



“Free-Summarizer”- Clarity & Collab at No Cost.

THEME-AI/ML



SOLUTION OVERVIEW

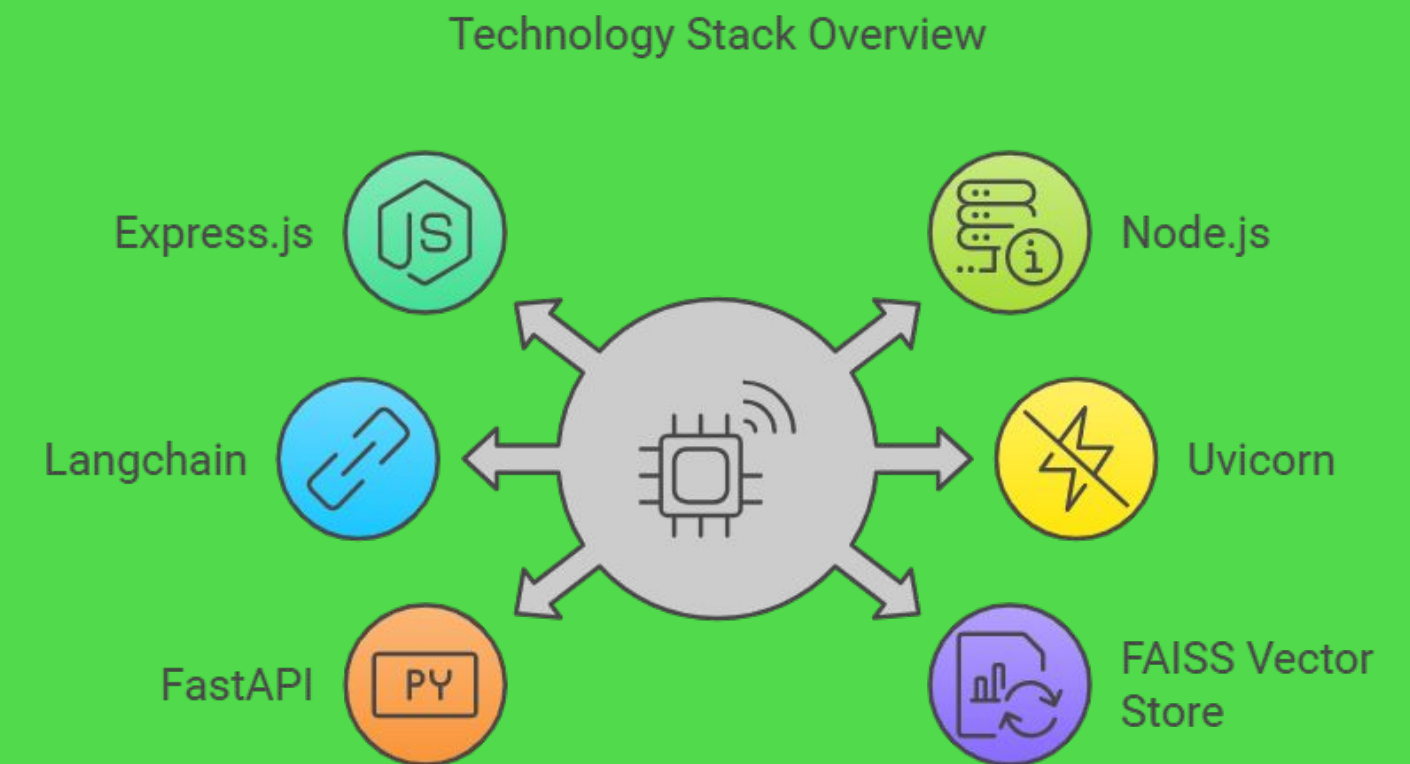
We are developing a web application that allows users to upload and share PDFs, and offers real-time collaboration with powerful summarization features. The app leverages LLM's (Large Language Models) to:

- **Summarize the content within PDFs.**
- **Answer specific questions based on the document's content via a built-in chatbot.**
- **Allow multiple users to share a link and collaborate in a shared space, enabling them to ask questions, access summaries, and interact with the document in real time.**
- **This solution not only enhances collaboration but also provides an intelligent assistant that saves time by extracting key points and answering queries on the fly.**

TECHNICAL ARCHITECTURE

Tech Stack

- Frontend Technologies: React.js
- Backend Technologies: Express.js, Node.js
- Database: FAISS Vector Store, MongoDB
- Other Tools/Services: Ollama, Langchain, Uvicorn, FastAPI



SCALABILITY AND FUTURE SCOPE

- **Increased User Load:** The application can be scaled to support a growing number of users by leveraging cloud infrastructure for load balancing, distributed databases, and microservices architecture.
- **Model Upgrades:** Future integration of more advanced LLMs like GPT-4 or custom fine-tuned models will enhance summarization accuracy and response relevance.
- **Collaboration Features:** Expanding to support features like version control, real-time annotations, and live editing can deepen the collaboration experience.
- **Integration with Third-Party Tools:** The app could integrate with cloud storage (Google Drive, Dropbox), CRMs, or project management tools to enhance document workflow automation.
- **Mobile & Offline Access:** Developing mobile apps and offering offline mode would increase accessibility and expand the user base across various platforms.
- **Enterprise Solutions:** Scaling to offer enterprise solutions with customizable models, advanced analytics, and security features for teams in legal, finance, and education sectors.

FEASIBILITY

Challenges and Risks

1. **Hosting and Deployment:** High costs and complexity in scaling real-time LLM services.
2. **Collaboration Integration:** Real-time sync, version control, and conflict management.
3. **Extending Chat Memory:** Maintaining long interaction histories impacts performance.
4. **Handling Long Context Tasks:** LLMs struggle with token limits for lengthy content.
5. **API Limitations:** Rate limits and high costs for complex or long-running tasks.

Strategies to Overcome

1. **Efficient Hosting:** Use containerization and hybrid cloud to scale and manage costs.
2. **Collaborative Tools:** Add real-time sync, document locking, and version control.
3. **Selective Memory Retention:** Store key summaries, use chunking for long chats.
4. **Context Management:** Break large content into chunks.
5. **Optimize API Usage:** Use batching, caching, and local processing to reduce API load.



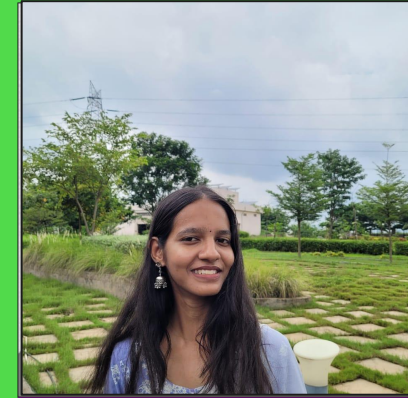
> Team Details





Shrinil Geddam

Student
VIT-AP University



Vimedha Chaturvedi

Student
VIT-AP University



Bhargav Gembali

Student
VIT-AP University



Aryan Rangapur

Student
VIT-AP University

11

**HACKTOBER
FEST**



Thanks for Joining

