





DocChatAI: Intelligent Document Assistant for Efficient Information Retrieval

By Maira Rubab 01/10/24

Team: Deepanand Saha Roy, Sreekanth Palagiri, Krishna Hariramani, Ersi Ni

HOST INSTITUTIONS











Objective: Provide Solutions





Demand for User-Friendly, Al-Powered Tools

Less Productivity & Frustration



Technology Stack Used

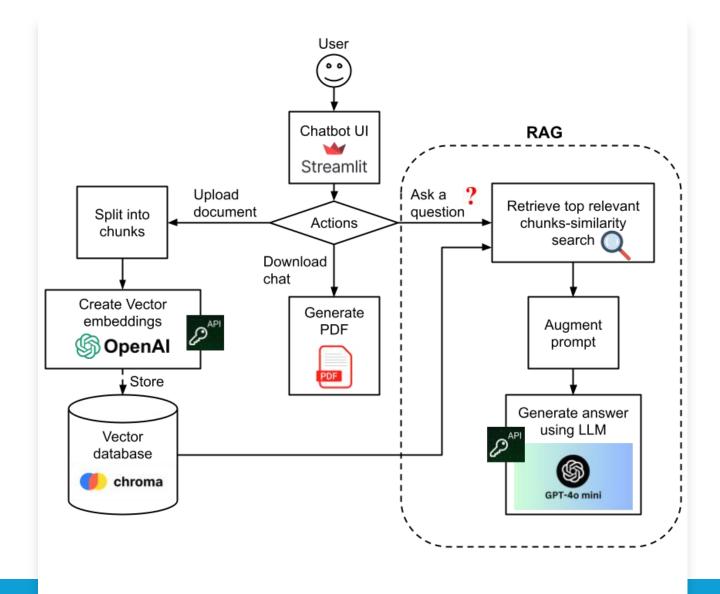
- ➤ Backend Technologies — OpenAl Embeddings, Chroma Database, RAG, & GPT Model

Retrieval Augmented Generation (RAG)

It is an AI development technique where a large language model (LLM) is connected to an external knowledge base to improve the accuracy & quality of its responses.



Application Architecture





Demo...



Deploy :

Upload the Documents

Drag and drop files here

Limit 200MB per file • PDF

Browse files



Please upload the documents to start the conversation!

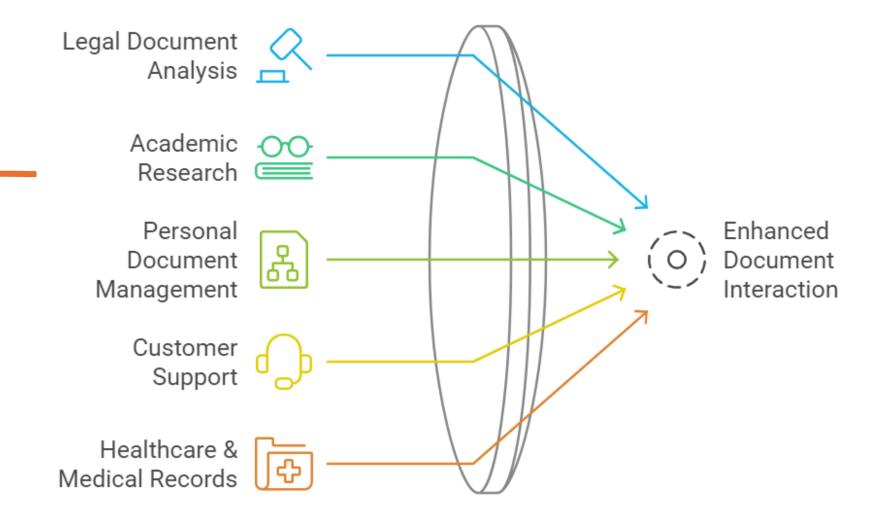












Use cases









Challenges & Limitations

- □ Prompt Engineering
- ☐ Response Completeness
- ☐ Scalability with document volume
- ☐ Dependency on external API's

Future Considerations



Improved and specialised prompts for better responses



Scaled Support for Larger Files (in GBs)



Voice input and output for advanced & natural conversational interaction



Session handling to allow users to store and access conversations and files seamlessly across devices



Accessible technical settings like text splitting configuration, embedding model, LLM model, & size of the stored files



Capability to modify and retry a query within the conversation for improved interaction

TCS Internship Experience (Overall Learnings)

- Learnt Generative AI, LLMs, & Prompt Engineering
- Learned different technologies including docker, REST API's, OpenAI API, AWS services like EC2 instance, ECR, S3, Amazon Sagemaker, & IAM, and setting virtual environments, directories, and configurations
- Improved project management in GitHub, & in Agile frameworks like Miroboard
- Developed product-oriented mindset
- Gained a sense of accomplishment
- Enhanced multi-tasking capability
- Improved work organization & communication
- Developed more consistency & efficiency
- Trained to meet timelines/deadlines for given tasks
- Boosted up confidence

Thank You For Your Attention!

Any Questions

