



AGENTIC RAG CHATBOT

for Multi-Format Document QA



-BY RAKSHA BHAIRI

PROBLEM STATEMENT

- Build a Retrieval-Augmented Generation (RAG) chatbot
- Capable of answering questions from diverse document formats
- Must follow an agent-based architecture
- Use Model Context Protocol (MCP) for agent communication

PROJECT GOALS

- Develop modular, agent-based QA system
- Implement MCP messaging
- Provide a user-friendly document upload + query interface

TECH STACK

Languages & Libraries:

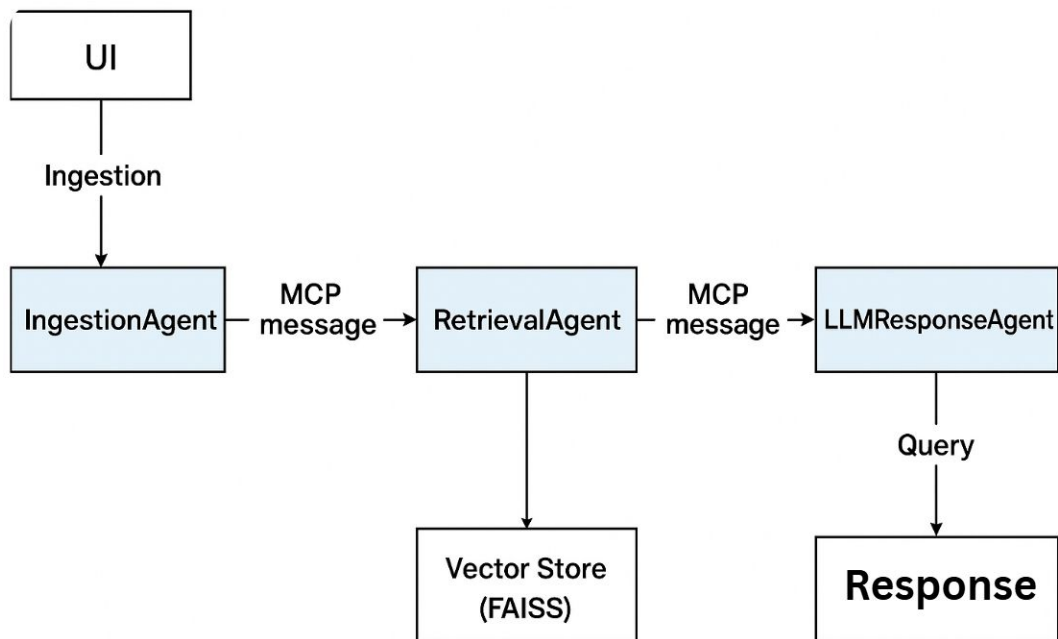
- Python 3
- Streamlit (UI)
- FAISS (Vector Store)
- Cohere API (LLM)
- Sentence Transformers (Embeddings)
- python-docx, python-pptx, PyMuPDF, pandas (Parsers)

TECH STACK

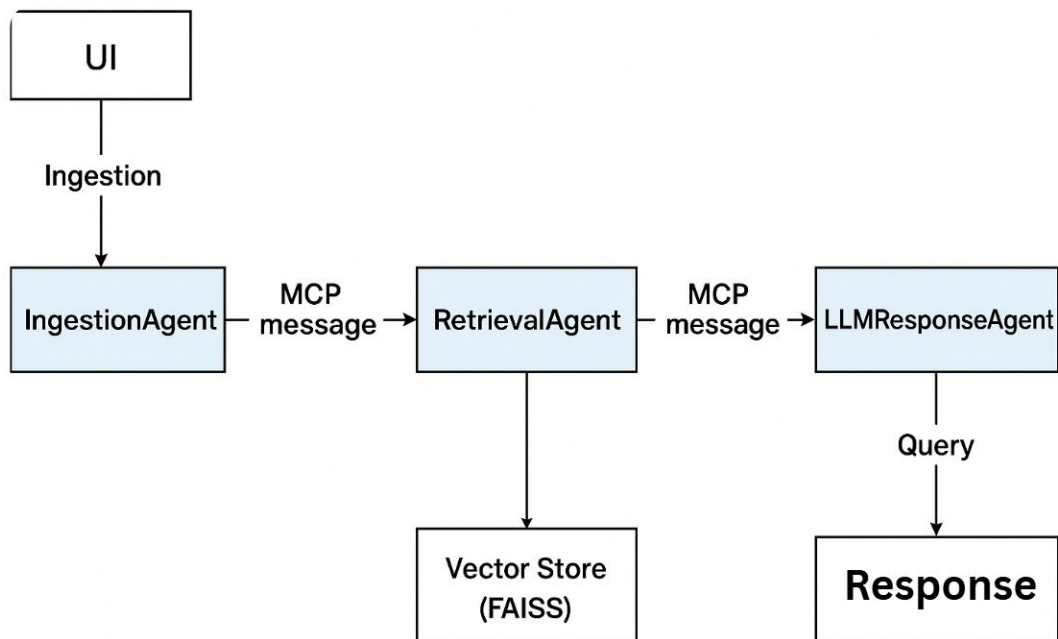
Architecture Principles:

- Agent-based modularity
- In-memory MCP messaging
- RAG + vector retrieval

AGENT-BASED ARCHITECTURE WITH MCP INTEGRATION



AGENT-BASED ARCHITECTURE WITH MCP INTEGRATION



UI SNAPSHOTS

Deploy



Agentic RAG Chatbot

Upload your file



Drag and drop file here

Limit 200MB per file • PDF, PPTX, DOCX, CSV, TXT

Browse files



eia.pdf 441.7KB



Ask a question based on the uploaded document

key steps in preparing eia report

eia.pdf successfully ingested!



Response from LLM:

Here are the key steps in preparing an Environmental Impact Assessment (EIA) report:

- Scoping:** Define the scope and objectives of the EIA, including the specific environmental issues and impacts to be assessed.
- Data Collection and Analysis:** Gather relevant data and information through field surveys, literature reviews, and consultations with experts and local communities. Analyze and interpret the data to establish baseline environmental conditions and identify potential impacts.
- Impact Prediction and Evaluation:** Predict and evaluate the potential environmental impacts of the proposed project or activity. Use appropriate methodologies, models, and impact assessment matrices to assess impacts on various environmental factors such as air quality, water resources, biodiversity, soil, human health, etc.
- Mitigation Measures and Management Plans:** Propose mitigation measures to minimize, reduce, or eliminate adverse environmental impacts. Develop management plans that outline specific actions, responsibilities, and timelines for implementing the mitigation measures.
- Consultation and Public Participation:** Engage stakeholders, including local communities, government agencies, and special interest groups, through public consultations and hearings. Consider their concerns, suggestions, and feedback in the EIA report.
- Report Preparation:** Prepare a comprehensive EIA report that documents the entire process, including the scoping, data collection, impact assessment results, proposed mitigation measures, management plans, and stakeholder engagement outcomes. Ensure the report adheres to the required format and guidelines specified by the relevant authorities.

UI SNAPSHOTS

Deploy 



Agentic RAG Chatbot

Upload your file



Drag and drop file here

Limit 200MB per file • PDF, PPTX, DOCX, CSV, TXT

Browse files



eia.pdf 441.7KB



Ask a question based on the uploaded document

what is eia report?

eia.pdf successfully ingested!



Response from LLM:

An EIA report, or Environmental Impact Assessment report, is a document that identifies, predicts, and evaluates the likely environmental impacts of a proposed project or development. It aims to ensure that decision-makers consider the environmental impacts of a project when deciding whether or not to approve it. The specific process and requirements for an EIA report can vary depending on the country and the specific legislation in place.

Would you like to know more about EIA reports or the environmental impact assessment process?

CHALLENGES

Challenges include ensuring document chunk consistency across formats, tuning FAISS indexing for smaller datasets, effectively handling multi-turn queries with context retention, and adapting response formatting to suit different contexts.

FUTURE SCOPE

Future scope includes enhancing multi-turn conversational memory for deeper, more contextual interactions; expanding support for cloud-based vector databases like Pinecone and Chroma to improve scalability and performance; evolving the system into a robust API-based microservices architecture for greater modularity; and integrating additional LLMs such as Gemini and Mistral to offer broader model flexibility and capabilities.

CONCLUSION

This project successfully demonstrates the power of an agent-based Retrieval-Augmented Generation (RAG) system built on the Model Context Protocol (MCP). By integrating modular agents for document ingestion, retrieval, and LLM response generation, the system offers a scalable, multi-format document QA solution. The architecture ensures clean message passing, efficient document parsing, and reliable context-aware answers — paving the way for enterprise-ready, intelligent chatbot systems capable of handling diverse and dynamic knowledge sources.



THANK YOU