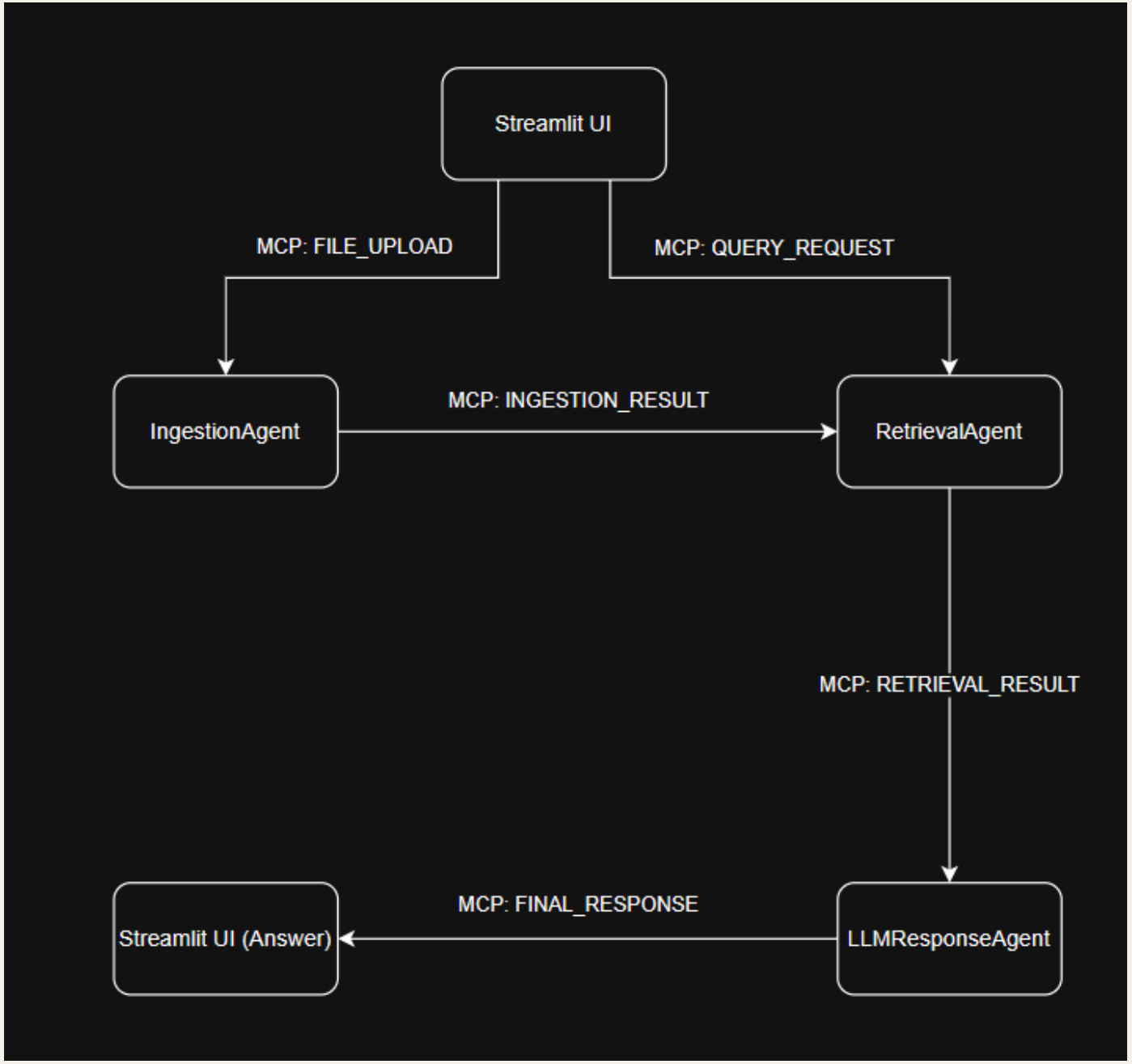
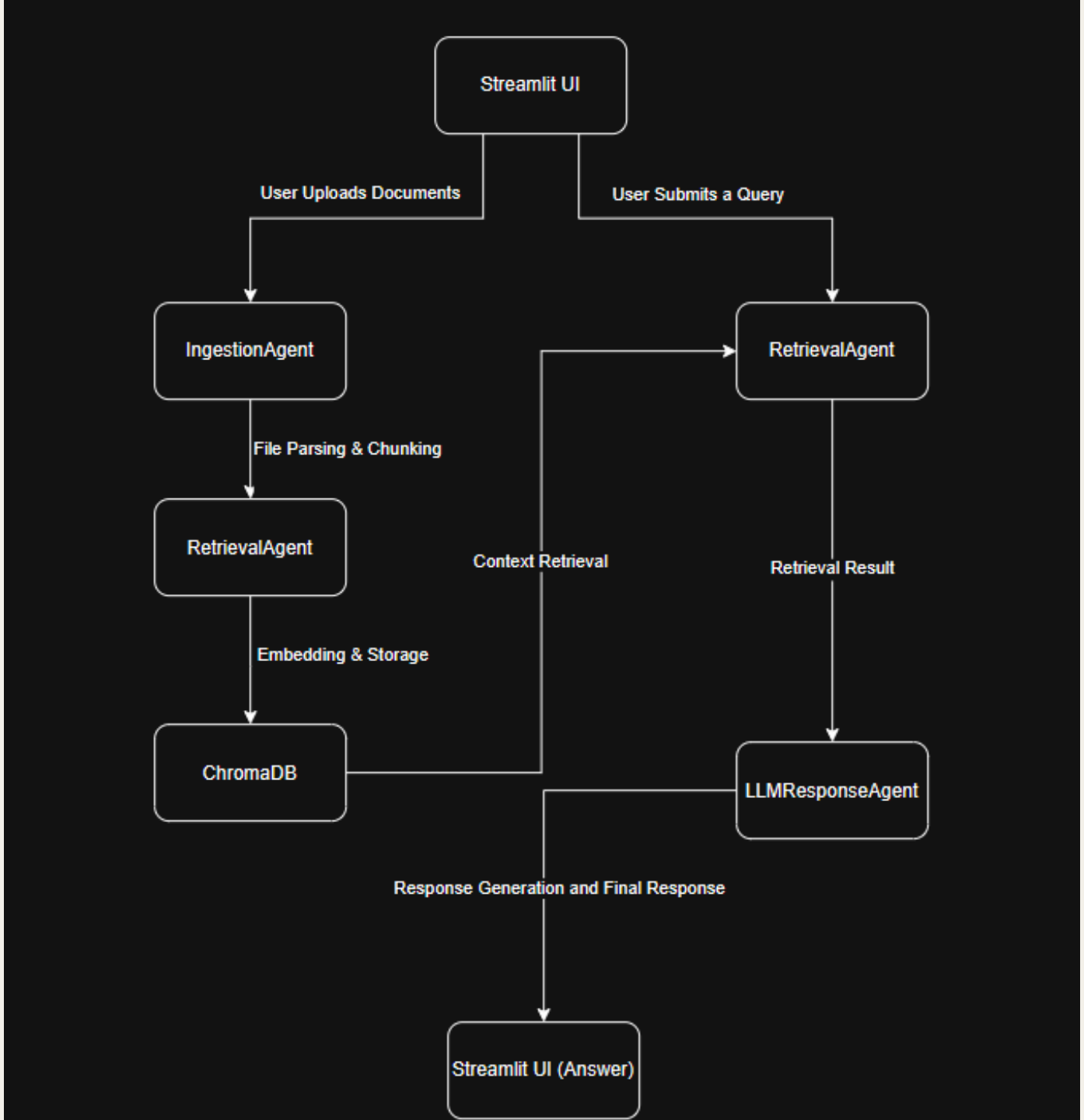
Agentic RAG Chatbot for Multi-Format Document QA using Model Context Protocol (MCP)

Siddartha Kommu

Agent-based architecture with MCP integration



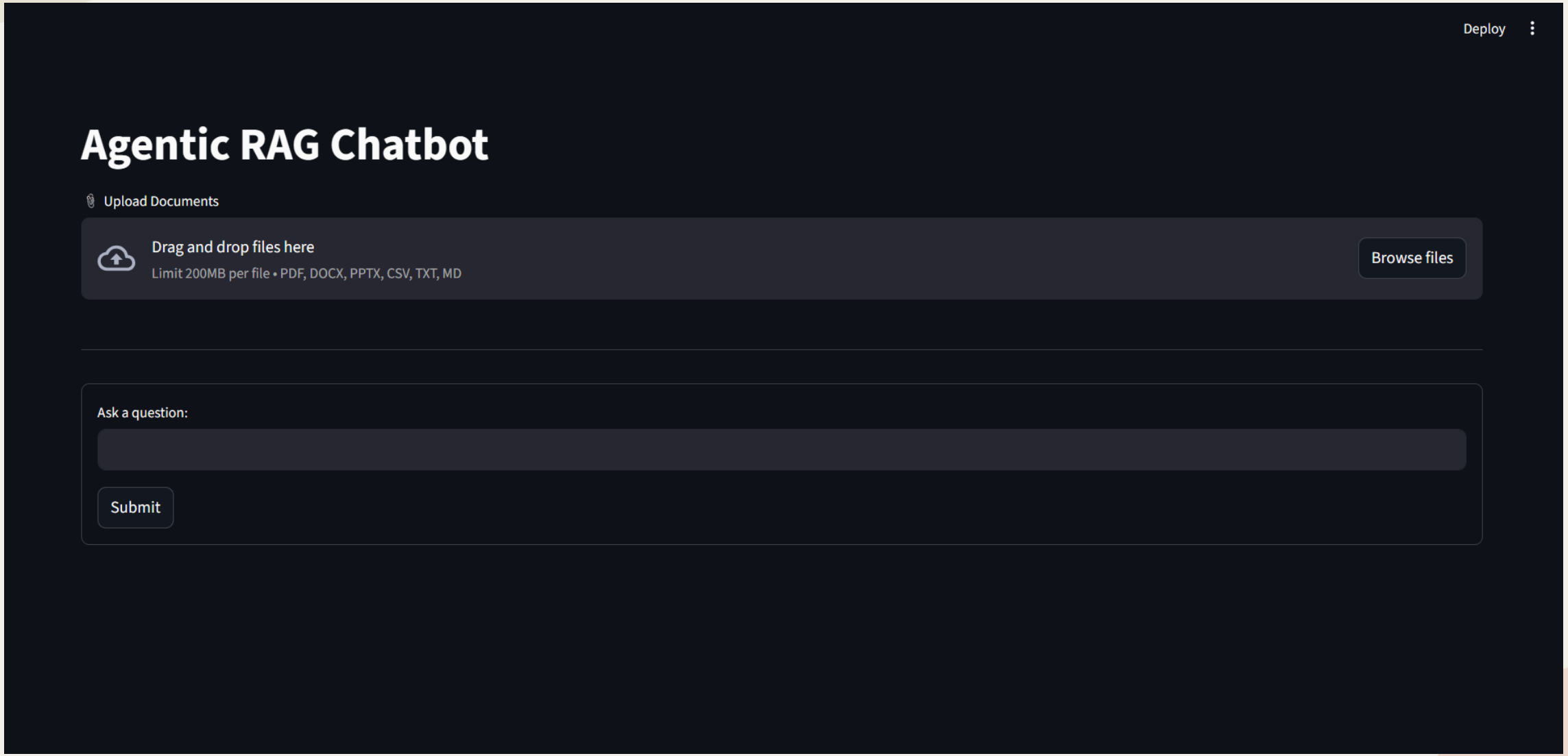
System flow diagram



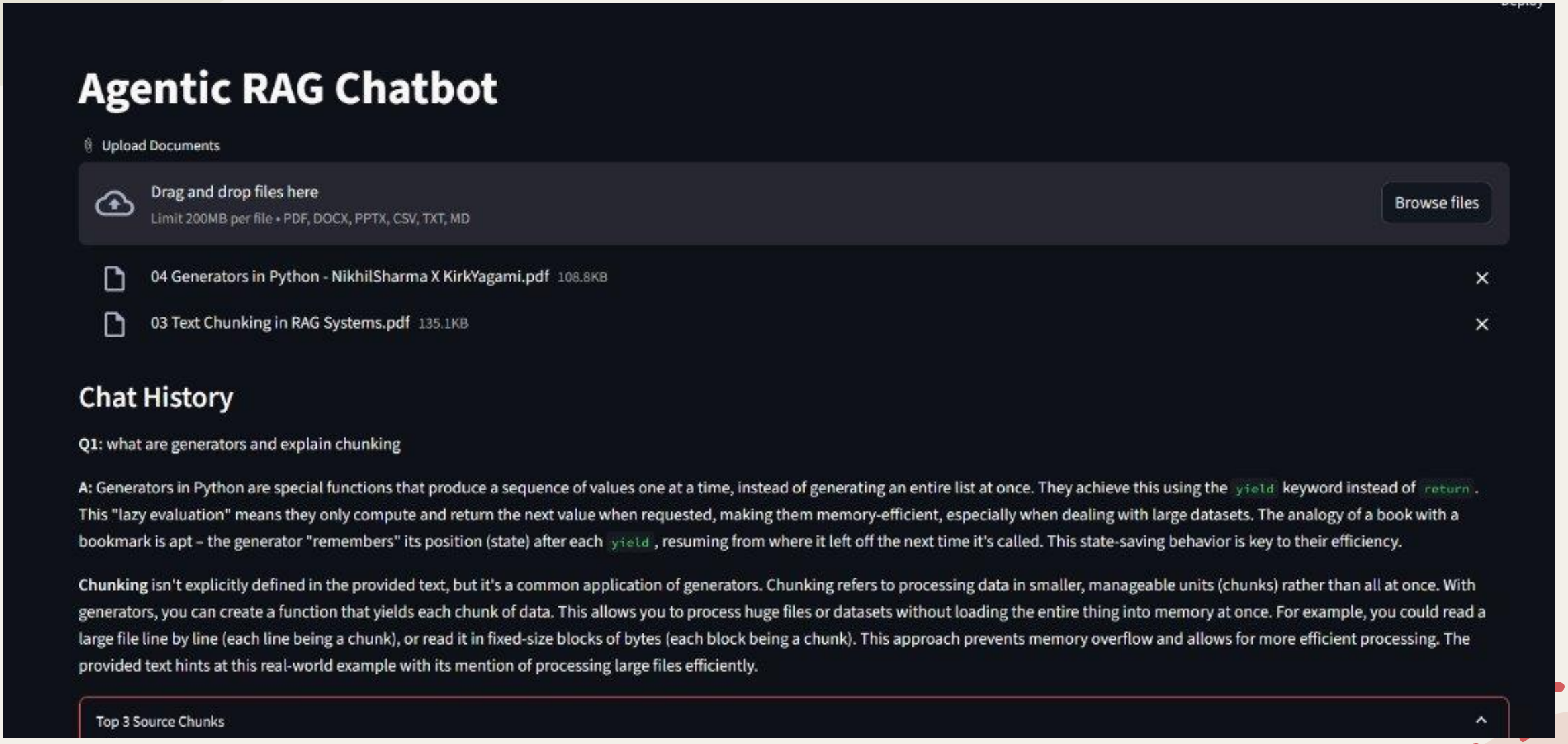
Tech Stack

- Python,
- Streamlit
- Sentence-Transformers (all-MiniLM-L6-v2)
- ChromaDB
- Google Generative AI (Gemini 1.5 Flash)

UI SCREENSHOTS



CHATBOT INTERFACE



Uploaded 2 documents and Asked questions from them and Got the answer back with the source chunks

Challenges Faced

- **Chunking Logic**
- Splitting long docs into meaningful ~400-char chunks
- **Streamlit State Management**
- Avoiding double-submit / rerun issues
- **ChromaDB Migration**
- Upgrading to new PersistentClient API
- **MCP Coordination**
- Ensuring correct traceable message passing