

This is a sample document for testing the LLM RAG Chatbot.

The goal of this chatbot is to retrieve relevant information from uploaded documents using a combination of vector search (FAISS) and a local language model (like Mistral via Ollama).

This document contains some basic information about the Retrieval-Augmented Generation (RAG) process.

RAG combines the strengths of retrieval-based methods and generative models to produce more accurate and context-aware answers.

Key Points:

- RAG improves factual accuracy by grounding generation on external data.
- FAISS enables fast and efficient vector similarity search.
- LangChain helps orchestrate the entire retrieval and response pipeline.

This document is intended for demonstration purposes.