

Voice AI Assistant Documentation

This is a sample document for testing the Voice AI Assistant.

About the Project:

The Voice AI Assistant is a production-grade application that uses RAG (Retrieval-Augmented Generation) to answer questions from PDF documents. It uses Google Gemini for generating responses and ChromaDB for storing document embeddings.

Key Features:

1. Voice Input - The assistant listens to your microphone and transcribes speech to text.
2. RAG Pipeline - Documents are chunked, embedded, and stored in ChromaDB for semantic search.
3. Gemini LLM - Questions are answered using the gemini-1.5-flash model for fast responses.
4. Voice Output - Responses are converted to speech using gTTS and played back.

How to Use:

1. Place your PDF documents in the data folder
2. Run the ingest.py script to build the knowledge base
3. Start the assistant with python app.py or use the web interface
4. Ask questions about your documents and get voice responses

Technical Stack:

LangChain for the RAG pipeline, ChromaDB for vector storage, Google Gemini API for language generation, SpeechRecognition for voice input, gTTS for voice output, FastAPI for the REST backend, and React for the web frontend.