



## Project: Company Knowledge Assistant (RAG + Agent)



### Description:

A smart assistant that helps employees query internal documents (project specs, HR policies, meeting notes, etc.) using natural language. The system uses Retrieval-Augmented Generation and LangChain agents to reason, search, and respond based on company documents.



### Features:

- Upload company documents (PDF, DOCX, TXT, Markdown, etc.)
- Semantic search using vector database (RAG)
- Conversational UI with memory and context
- LangChain agent to answer complex queries with reasoning
- User roles: Admin (manage documents), Employee (chat assistant)
- Real-time web UI with **Chainlit**



### Backend (Python, LangChain, FastAPI):

#### Document ingestion pipeline:

- Extract text from PDFs, DOCX, images (OCR)
- Chunking & metadata tagging
- Embedding with OpenAI, Cohere, or Gemini
- Store in vector DB (Chroma / FAISS / Weaviate / Qdrant)

#### RAG setup:

- LangChain
- Retriever
- +
- ConversationalRetrievalChain
- Optionally ensemble retriever for multimodal

## Frontend (Chainlit):

Company-branded chatbot interface

Role-based interaction: show admin upload panel,  
employee chat

Message history, citations, source document highlights