(Retrieval-Augmented Generation) AI assistant that reads company documents and answers questions

Tech Stack:

- Self hosted server (Digital Ocean)
- Email client
- AI workflow Node
- Json Text extraction
- Gpt 4.1o mini

Use Case

- The organization stores important internal knowledge across multiple documents such as policies, manuals, reports, and guides.

- Employees often struggle to quickly find the right information inside long documents.

- Searching manually through folders or PDFs is time-consuming and reduces productivity.

- Traditional keyword search fails when users don't know the exact wording used in the documents.

- The business needs an intelligent assistant that can understand employee questions and retrieve answers directly from company documents.

Solution Design

- Collect company documents from a central storage location for processing.

- Break each document into smaller readable sections so the system can search efficiently.

- Convert these document sections into semantic representations that allow meaning-based search rather than keyword matching.

- Store these representations in a searchable knowledge index.

- When a user asks a question, convert the question into the same semantic format.

- Retrieve the most relevant document sections related to the question.

- Provide those sections to an AI model so it can generate an accurate answer grounded in company knowledge.

- Return the final answer to the user in a conversational format.

Hosting

- Self Hosted via Server Node
- Using AI workflow tool

Ai Chatbot

- Open Ai Gpt 4.1o mini

Workflow