

Introduction to Testing PDF Q&A Bot

This is a test PDF created for the Intelligent PDF Q&A Bot proof-of-concept.

Use this document to verify text extraction, chunking, and retrieval.

Key points:

- The bot should find phrases like 'main conclusion', 'installation steps', and 'contact'.
- Page numbers are included in the footer so the bot can cite pages.

Sample sentence for retrieval: The main conclusion of this document is that small PoCs help validate architecture.

Installation Steps

1. Create a virtual environment.
2. Install dependencies with `pip install -r requirements.txt`.
3. Place PDFs in the `documents/` folder and run the script.

Sample command: `python pdf_qa.py "documents/myfile.pdf" "What is the main conclusion?"`

Sample sentence for retrieval: To install dependencies run `pip install pdfplumber openai scikit-learn numpy`

Example Content and Use Cases

This page contains an example paragraph about RAG: Retrieval-Augmented Generation (RAG) combines a

Sample sentence for retrieval: RAG improves factuality by providing the model with document context.

Conclusion & Contact

Conclusion: The main conclusion of this test document is that simple local RAG-lites are useful for rapid prototyping.

Contact: test@example.com

Sample sentence for retrieval: If the answer is not found in the provided context, respond 'I don't know based on the provided context.'