

Knowledge Assistant (MVP)

AI-Powered Q&A; over Contracts, Policies, and Reports

This document provides a structured synopsis of the Knowledge Assistant MVP. It is a lightweight tool that allows users to upload documents and ask natural language questions, receiving answers with citations.

1. Introduction

The Knowledge Assistant (MVP) enables quick extraction of insights from documents such as contracts, HR policies, and compliance reports. Users can upload files and ask questions, with answers provided along with citations.

2. Objectives

The project aims to provide a lean, functional MVP with:

- Simple UI for non-technical users
- Support for PDF, DOCX, and TXT
- Fast, citation-backed answers
- Answer export as PDF
- Password-protected access
- Workspaces for client separation

3. Features

Key features of the Knowledge Assistant MVP include:

- Document Upload: PDF, DOCX, TXT
- Ingest & Index: Splits docs into searchable chunks
- Natural Language Q&A: Plain English queries
- Citations: Trace answers back to source pages
- Export: Download answers as PDF reports
- Workspaces: One per client/project
- Password Protection: Configurable via environment secret

4. Technical Stack

The app is built using proven open-source libraries:

- Streamlit for UI
- pdfplumber and python-docx for document ingestion
- BeautifulSoup and lxml for text cleaning

scikit-learn for TF-IDF retrieval and similarity
reportlab for PDF export
Streamlit Cloud for deployment

5. System Workflow

The system processes documents in 6 steps:

Upload → PDF/DOCX/TXT
Clean and Chunk → Prepare text blocks
Vectorize → TF-IDF embeddings
Ask → User enters a natural language question
Retrieve → Top relevant text passages returned
Answer → Concise answer generated with citations and PDF export

6. Use Cases

The tool benefits multiple industries:

Legal: Extract renewal dates, liability clauses
Finance: Search audits, compliance documents
Consulting: Accelerate client document reviews
General: Q&A; over policies, research, and reports

7. Future Enhancements

Planned improvements include:

Semantic search using embeddings (OpenAI or HuggingFace)
Multi-user accounts and roles
Database-backed storage for persistence
Analytics dashboard for usage insights
Polished UI with highlights and themes

8. Conclusion

This MVP validates the workflow of a Knowledge Assistant product. It takes users from document upload to natural language Q&A; with citations. The system demonstrates strong potential for SaaS deployment across legal, finance, and consulting sectors.