

Streamlit Book Q&A System

The **Streamlit Book Q&A System** is a powerful and user-friendly web application that transforms static book files (PDF or TXT) into interactive, searchable knowledge bases. Designed with simplicity and functionality in mind, this tool allows users to upload books, generate knowledge graphs, and ask natural language questions with contextual answers attributed to source material.

Features

-  **File Upload:** Easily upload PDF or TXT files.
-  **Content Processing:** Automatically extracts and segments text, identifying chapters and sections.
-  **Vector Database:** Uses Google's embedding model to embed text chunks and stores them in a FAISS vector store for fast similarity search.
-  **Ontology Generation (Optional):** Generates a basic knowledge graph using Google Generative AI (Gemini), extracting key concepts and relationships.
-  **Knowledge Graph Viewer:** Visualizes the ontology using NetworkX and Matplotlib.
-  **Multi-Book Querying:** Query one or more books at once for comprehensive results.
-  **RAG-Based Q&A:** Utilizes a Retrieval-Augmented Generation (RAG) pipeline powered by LangChain and Google Generative AI for accurate, contextual answers.
-  **Source Attribution:** Displays book title, chapter, and page for each answer's context.
-  **Chat Interface:** Includes a conversational interface to explore book contents interactively.
-  **Data Persistence:** Saves processed data locally to avoid repeated processing.
-  **Password Protection:** Simple authentication mechanism to secure access.
-  **Streamlit Interface:** Intuitive and interactive UI for seamless user experience.

🛠 Tech Stack

- **Frontend:** [Streamlit](#)
 - **NLP & RAG:** [LangChain](#), Google Generative AI (Gemini)
 - **Embedding & Vector Store:** Google Embeddings, FAISS
 - **Knowledge Graph:** NetworkX, Matplotlib
 - **File Parsing:** PyMuPDF, pdfminer, or pdfplumber for PDFs; standard I/O for TXT
 - **Authentication:** Simple password mechanism
-

📦 Installation:

```
bash                                     ⌂ Copy ⌂ Edit

git clone https://github.com/yourusername/streamlit-book-qa.git
cd streamlit-book-qa
pip install -r requirements.txt
```

▶ Usage

1. Start the app:

```
bash                                     ⌂ Copy ⌂ Edit

streamlit run app.py
```

2. Login:

Enter the password to access the interface.

3. Upload Books:

Upload one or more PDF/TXT files via the file uploader.

4. Process & Analyze:

Books are split into sections, embedded, and indexed. Optionally generate a knowledge graph.

5. Query Content:

Ask natural language questions about any processed book(s). The system returns accurate answers with source citations.

6. Explore Knowledge Graph:

Visualize the ontology derived from the book using the Knowledge Graph viewer.

📁 Project Structure

```
bash                                ⚒ Copy    ⚒ Edit

Streamlit-book-qa/
|
|   └── processed_books/          # Processed data, vector stores, ontology graph, and metadata
|
|   └── app.py                   # Main Streamlit application
|
|   └── .env                      # Contains API Key and environment variables
|
└── README.md                  # Project documentation
```

🌐 Authentication

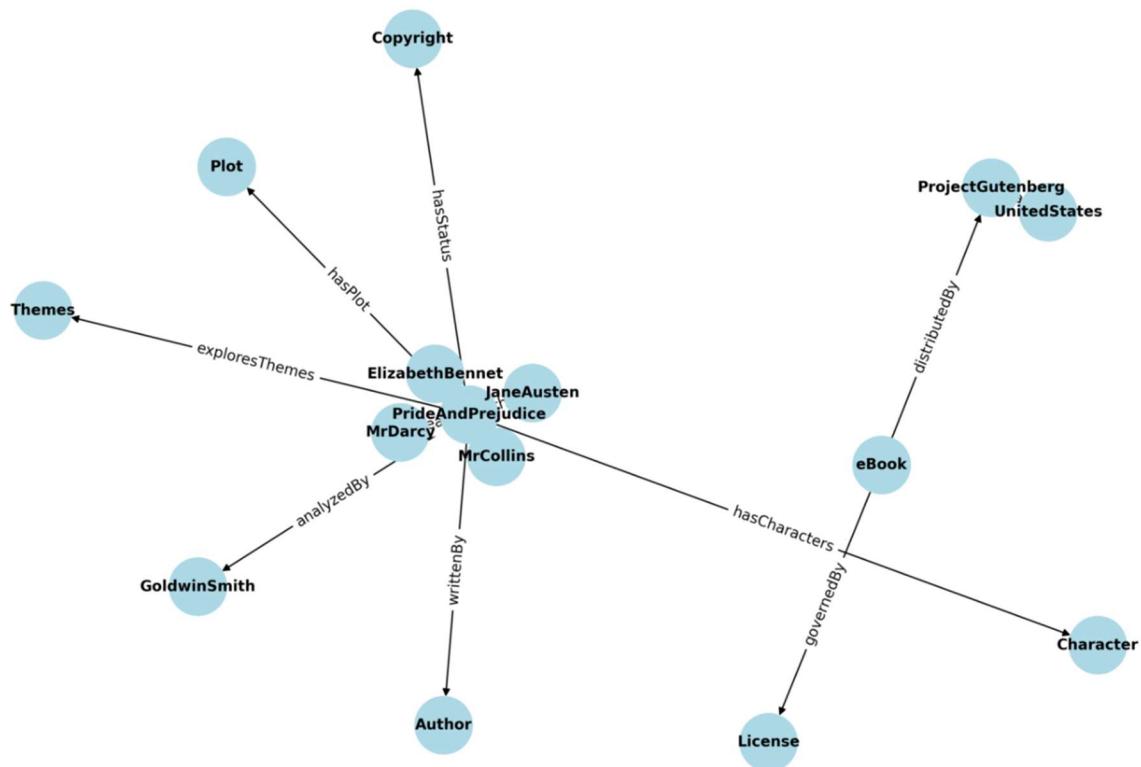
The app uses a basic password-based login system. Update the password in the config or environment variable for security.

✉️ Contact

For questions, suggestions, or collaborations:

Ajay Bhatt

Email: ajayabhatt441aj@gmail.com



Knowledge Graph (Ontology)

Relationships Between Concepts

	Source	Relationship	Target	Description
0	eBook	distributedBy	Project Gutenberg	eBooks are distributed by Project Gutenberg.
1	eBook	governedBy	License	Use of eBooks is governed by the Project Gutenberg License.
2	Pride and Prejudice	writtenBy	Author	Pride and Prejudice was written by Jane Austen.
3	Pride and Prejudice	hasStatus	Copyright	Pride and Prejudice is in the public domain in the United States.
4	Project Gutenberg	operatesIn	United States	Project Gutenberg operates primarily in the United States.
5	Pride and Prejudice	hasCharacters	Character	Pride and Prejudice contains various characters.
6	Pride and Prejudice	hasPlot	Plot	Pride and Prejudice has a plot.
7	Pride and Prejudice	exploresThemes	Themes	Pride and Prejudice explores themes such as pride and prejudice.
8	Elizabeth Bennet	appearsIn	Pride and Prejudice	Elizabeth Bennet is a character in Pride and Prejudice.
9	Mr. Darcy	appearsIn	Pride and Prejudice	Mr. Darcy is a character in Pride and Prejudice.
10	Jane Austen	isAuthorOf	Pride and Prejudice	Jane Austen is the author of Pride and Prejudice.
11	Pride and Prejudice	analyzedBy	Goldwin Smith	Pride and Prejudice was analyzed by Goldwin Smith.
12	Mr. Collins	appearsIn	Pride and Prejudice	Mr. Collins is a character in Pride and Prejudice.

Ontology Relationship Concepts

The screenshot displays the 'Book Q&A System' application interface, divided into two main sections: a sidebar on the left and a main content area on the right.

Processing Steps:

- Step 1: Upload Document
- Step 2: Extract Chapters
- Step 3: Create Text Chunks
- Step 4: Build Vector Database
- Step 5: Save Processed Book

Select Books:

Choose books to query:

- Default x ▼
- [Load Selected Books](#)

Book Q&A System:

Upload books and ask questions about their content.

About:

Main Content Area:

Book Q&A System

[Query Books](#) [Upload New Book](#) [Ontology Viewer](#)

Query Books

Loaded Books

Book: Default	Total Books	Total Chapters	Total Text Chunks
	1	87	964

Ask a Question

Enter your question about the book(s):

who is Bennett

[Ask](#)

Answer

There are multiple Bennet family members mentioned in the provided text from the book "Default". Miss Bennet is acknowledged as pretty, and Mr. Bennet enjoys leisure in his library. Mrs. Bennet is also mentioned in a conversation with Mr. Collins.

Streamlit Application UI