

DocuQuery

Introduction

DocuQuery is a flask web application that simplifies document management and information retrieval. It allows you to upload PDF documents, generate summaries, and ask questions related to the content of the documents, all through a user-friendly interface.

Features

- **Document Upload:** Easily upload PDF documents for analysis.
- **Automatic Summarization:** Generate concise summaries of uploaded documents.
- **Question & Answer:** Ask questions and receive answers based on the document content. Automatically jump to location in document where answer was found.
- **Caching:** Cache documents and summaries for improved performance.
- **User-Friendly Interface:** Intuitive web-based interface for all tasks.

LangChain Integration

DocuQuery leverages LangChain, a framework designed for developing applications powered by language models. It accomplishes efficient document analysis using chains, which are organized sets of components built to complete specific tasks.

Installation

1. Clone the project and navigate to project directory
2. Create and source the virtual environment:

```
python -m venv venv
```

```
source venv/bin/activate
```

3. Install dependencies with pip:

```
pip install -r requirements.txt
```

3. Specify OPENAI_API_KEY and SUMMARIZE_DOCUMENTS in config.py
4. Run the application

```
flask run
```