# DocuChat AI

## Document-Based Question Answering Application

> DocuChat AI is an AI-powered system that allows users to query PDF or TXT documents and receive instant, context-aware answers. It transforms static documents into interactive knowledge bases for research, analysis, and learning.

## USER GUIDE

*Version 1.0*

🤖 **AI-Powered** **Document Analysis** 💬 **Natural Q&A** 🔍 **Instant Answers** ⚙ **Smart Processing**

👤 **Author:** K.K. Ravindu
♡ Built with Python, Streamlit & AI

# Contents

# 1   Introduction

**DocuChat AI** is an intelligent document analysis platform that enables you to upload PDF or TXT documents and ask questions about their content using advanced AI technology. Powered by Large Language Models (LLMs) through OpenRouter API, it transforms static documents into interactive knowledge bases.

> **ⓘ Key Information**
>
> Instead of manually searching through pages, simply ask questions in plain English and receive accurate, context-aware answers instantly. Perfect for students, researchers, professionals, and anyone working with documents regularly.

## 1.1   Key Features

- ✅ Document upload (PDF/TXT)
- ✅ AI-powered Q&A
- ✅ Document analytics
- ✅ Chat history tracking
- ✅ Modern responsive UI
- ✅ Real-time processing

# 2   System Requirements

| Component | Requirement |
|---|---|
| Operating System | Windows 10+, macOS 10.14+, or Linux (Ubuntu 18.04+) |
| Python Version | 3.8 or higher |
| RAM | 4 GB minimum (8 GB recommended) |
| Storage | 500 MB free space |
| Internet | Broadband connection required |
| Browser | Chrome 90+, Firefox 88+, Edge 90+, or Safari 14+ |
| API Key | Valid OpenRouter API key (https://openrouter.ai/) |

# 3   Installation Guide

## 3.1   Step 1: Clone the Repository

```
git clone https://github.com/KKRavindu/-Document-Based-Question-Answering-
    Application
cd Document-Based-Question-Answering-Application
```

## 3.2   Step 2: Create Virtual Environment

**Windows:**

```
python -m venv venv
venv\Scripts\activate
```

**macOS/Linux:**

```
python3 -m venv venv
source venv/bin/activate
```

## 3.3   Step 3: Install Dependencies

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

## 3.4   Step 4: Configure API Key

**Option 1 - Direct Configuration:** Open `app.py` and edit line 18:

```
OPENROUTER_API_KEY = "your-api-key-here"
```

**Option 2 - Environment Variable (Recommended):**

Windows: `set OPENROUTER_API_KEY=your-api-key-here`

macOS/Linux: `export OPENROUTER_API_KEY="your-api-key-here"`

> **Pro Tip**
>
> Never commit API keys to version control. Use environment variables for better security.

# 4   Running the Application

## 4.1   Launch Command

```
streamlit run app.py
```

The application will open automatically at `http://localhost:8501`

**To use a different port:**

```
streamlit run app.py --server.port 8502
```

**To stop the application:** Press `Ctrl+C` in the terminal

# 5   User Interface Overview

## 5.1   Interface Components

**Sidebar Panel:**
- ⬆ File uploader for documents
- ❶ Document statistics
- 🚀 Process Document button

**Main Content Area:**
- ⌨ Question input field
- 🤖 AI-generated answers
- ↻ Chat history display
- 🗑 Clear history options

Figure 1: DocuChat AI Main Interface

# 6    Step-by-Step Usage Guide

## 6.1    Step 1: Upload Document



Figure 2: Upload Interface

1. Click **Browse files** in the sidebar

2. Select a PDF or TXT file (max 200MB)

3. Wait for upload confirmation

**Supported Formats:** PDF (.pdf), Text (.txt - UTF-8/Latin-1)

## 6.2    Step 2: Process Document
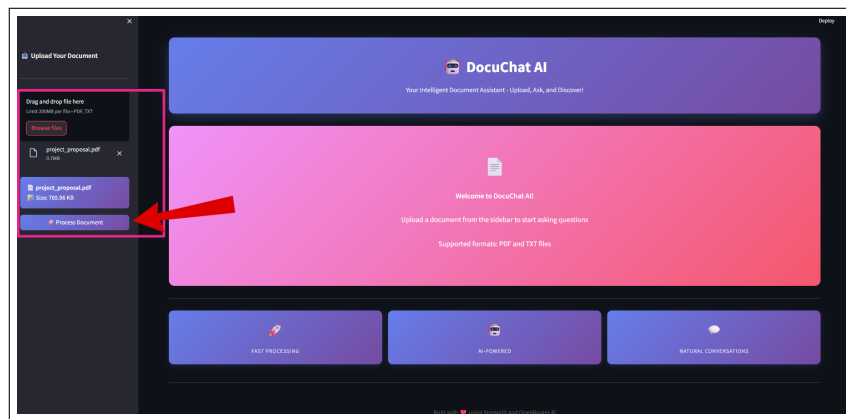


Figure 3: Document Processing

1. Click 🚀 `Process Document`
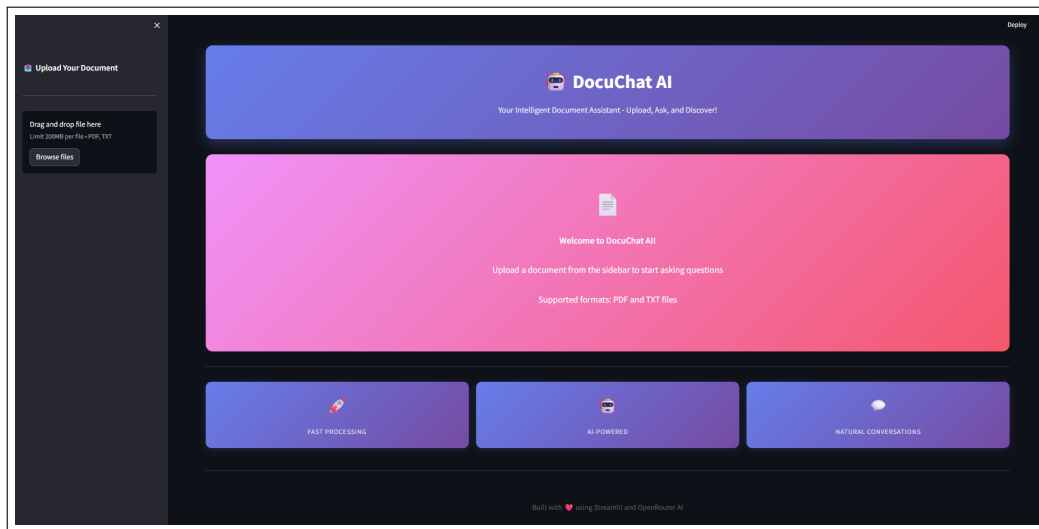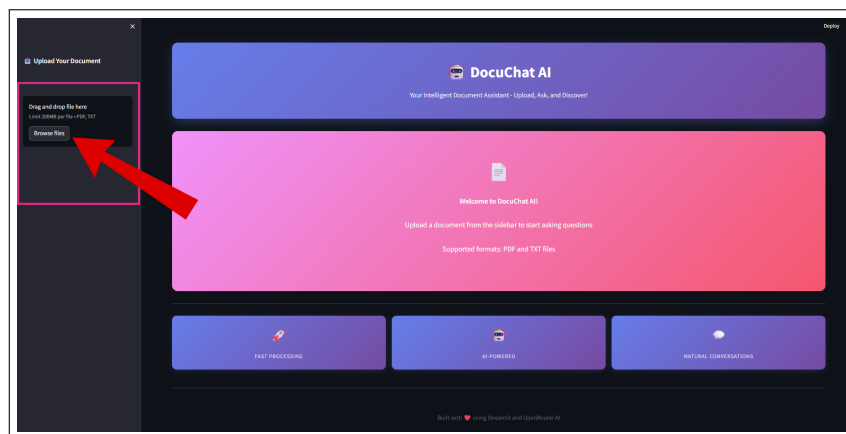
2. Wait for text extraction (2-30 seconds depending on size)

3. Review document statistics (characters, words, estimated pages)

## 6.3    Step 3: Ask Questions



Figure 4: Q&A Interaction

1. Type your question in the input field

2. Click 🚀 `Ask Question`

3. Wait 5-10 seconds for AI response

4. Review the answer and ask follow-up questions

## 6.4    Managing Your Session

**Clear Chat History:** Remove Q&A history while keeping document loaded
**Clear All:** Reset entire application (removes document and history)

# 7    Sample Questions

## 7.1    General Questions

❷ What is the main topic of this document?

❷ Summarize the key points

❷ What are the main conclusions?

❷ List important dates or numbers mentioned
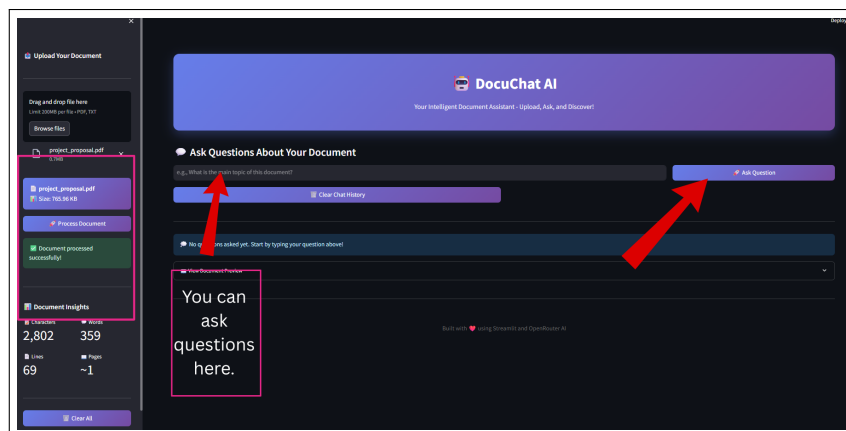
## 7.2    Domain-Specific Examples

**Research Papers:**
- What is the research hypothesis?
- What methods were used?
- What are the study limitations?

**Business Reports:**
- What is the revenue for this quarter?
- What are the main growth drivers?
- What risks are identified?

**Legal Documents:**
- What are the key terms of this agreement?
- What are the obligations of each party?
- What are the termination conditions?

> 💡 **Pro Tip**
>
> **Tips for Better Questions:**
> - Be specific and clear
> - Use complete sentences
> - Ask one question at a time
> - Include context if needed (e.g., "In section 3...")

# 8    Troubleshooting

## 8.1    Common Issues

**API Error - Invalid Key**
- Verify API key is correctly entered
- Check OpenRouter account for available credits
- Test internet connection

**PDF Extraction Fails**
- Remove password protection from PDF
- Ensure PDF is not corrupted
- Use OCR for image-based PDFs

**Import/Dependency Errors**
- Ensure virtual environment is activated
- Run: `pip install --upgrade -r requirements.txt`
- Check Python version: `python --version`

# 9   Best Practices

## 9.1   Security

Never share API keys publicly

Use environment variables for sensitive data

Be cautious with confidential documents

Review OpenRouter's privacy policy

## 9.2   Performance Optimization

Keep documents under 50MB for optimal speed

Use text-based PDFs (not image-only)

Split large documents into sections

Close unnecessary browser tabs

## 9.3   Question Formulation

💬 Start with broad questions, then get specific

💬 Reference specific sections when possible

💬 Use keywords from the document

💬 Rephrase if answers are unclear

### Thank You for Using DocuChat AI!

*"Transforming Documents into Interactive Knowledge"*

♡ Built with passion using Python, Streamlit & AI