

OpenAI Document Search Assistant - Usage Guide

This guide will help you set up and use the OpenAI Document Search Assistant to extract relevant information from Word documents.

Setup Instructions

1. **Environment Setup:** Create a `.env` file in your project directory with your OpenAI API key:

```
OPENAI_API_KEY=your_api_key_here
```

2. **Install Dependencies:**

```
bash
```

```
pip install openai python-dotenv tqdm
```

3. **Prepare Document Directory:** Create a directory named `documents` in the same folder as the script and place your Word (.docx) files there.

Basic Usage

python

```
from openai_doc_assistant import DocumentSearchAssistant

# Initialize the assistant
assistant = DocumentSearchAssistant()

# Upload documents
file_ids = assistant.upload_files("documents")

# Create assistant
assistant.create_assistant()

# Create thread
assistant.create_thread()

# Ask questions
questions = [
    "What does the document say about project planning?",
    "What are the key recommendations in the report?",
    "Summarize the findings on budget allocation."
]

results = assistant.process_questions(questions)

# Print results
for question, answer in results.items():
    print(f"\nQ: {question}")
    print(f"A: {answer}")
    print("-" * 50)

# Cleanup when done
assistant.cleanup()
```

Advanced Usage

Custom Instructions

You can provide custom instructions to guide how the assistant responds:

python

```
assistant.create_assistant(  
    instructions="You are a document analysis expert. Extract key insights from documents, focu  
)
```

Different File Types

While the script is optimized for Word documents, you can specify other file types:

python

```
file_ids = assistant.upload_files("documents", file_pattern="*.pdf")
```

Single Question Mode

For one-off questions:

python

```
response = assistant.ask_question("What are the main conclusions in the documents?")  
print(response)
```

Troubleshooting

- **API Key Issues:** Ensure your API key is correct and has appropriate permissions
- **File Upload Failures:** Check file size limits (less than 512MB per file)
- **Timeout Errors:** For large documents, consider increasing the timeout in `ask_question()`

Logging

The script logs all actions to both console and `assistant_log.log`. Check this file for debugging information if you encounter issues.