

Documentation for FAISS

Retrieval Augmented Generation using FAISS

1 Environment Requirements

- This code runs on **Python 3.11**
- This code is also compatible with **Python 3.14**

2 Steps to Run the Code

```
git clone https://github.com/adityasaicareer/  
    celestial_systems_faiss.git  
cd celestial_systems_faiss  
pip install -r requirements.txt  
python rag_faiss.py
```

3 Libraries Required

```
from langchain_community.document_loaders import PyPDFLoader  
from langchain_text_splitters import  
    RecursiveCharacterTextSplitter  
from langchain_huggingface import HuggingFaceEmbeddings  
from langchain_community.vectorstores import FAISS  
import re  
import pprint
```

3.1 Library Description

- **PyPDFLoader** parses PDF documents required for context generation.
- **RecursiveCharacterTextSplitter** divides documents into overlapping chunks.
- **HuggingFaceEmbeddings** uses the `sentence-transformers/all-MiniLM-L6-v2` model to generate vector embeddings.
- **FAISS** (Facebook AI Similarity Search) enables efficient storage and fast similarity search over embeddings.

4 Loading the PDF

```
filepath = "./example.pdf"

loader = PyPDFLoader(filepath)
print(loader)

docs = loader.load()
```

The above code demonstrates loading a PDF document using the LangChain PyPDFLoader.

5 Chunking the Document

```
text_splitter = RecursiveCharacterTextSplitter(
    chunk_size=500,
    chunk_overlap=100
)

chunks = text_splitter.split_documents(docs)

for idx, chunk in enumerate(chunks):
    chunk.metadata["chunk_id"] = idx
```

Each document is split into chunks of 500 characters with an overlap of 100 characters to preserve contextual continuity.

6 Embedding and FAISS Vector Store Creation

```
embeddings = HuggingFaceEmbeddings(
    model_name="sentence-transformers/all-MiniLM-L6-v2"
)

faiss_db = FAISS.from_documents(
    documents=chunks,
    embedding=embeddings
)

faiss_db.save_local("ragdata")
```

The FAISS vector store is created using document chunks and their corresponding embeddings and saved locally for reuse.

7 Querying the FAISS Vector Database

```
query = "How does top management demonstrate leadership and  
commitment to the ISMS?"  
  
docs = faiss_db.similarity_search_with_score(  
    query,  
    k=5  
)
```

The raw query is automatically embedded using the configured embedding model, and FAISS performs similarity search to retrieve the most relevant chunks.

8 Outputs

8.1 Query 1

How does top management demonstrate leadership and commitment to the ISMS?

Top K = 5

```

Query : How does top management demonstrate leadership and commitment to the ISMS?
Top K :5

-----
Chunk :1
Chunk Content:
management system, including the processes needed and their interactions, in accordance with the
requirements of this document.
5 Leadership
5.1 Leadership and commitment
Top management shall demonstrate leadership and commitment with respect to the information
security management system by:
a) ensuring the information security policy and the information security objectives are established
and are compatible with the strategic direction of the organization;

Source:./example.pdf
Page : 7
chunk_id : 51
Score : 0.9229822754859924
-----
Chunk :2
Chunk Content:
security are assigned and communicated within the organization.
Top management shall assign the responsibility and authority for:
a) ensuring that the information security management system conforms to the requirements of this
document;
b) reporting on the performance of the information security management system to top management.
NOTE Top management can also assign responsibilities and authorities for reporting performance of the

Source:./example.pdf
Page : 8
chunk_id : 57
Score : 0.986425518989563
-----
Chunk :3
Chunk Content:
management system implementation will be scaled in accordance with the needs of the organization.
This document can be used by internal and external parties to assess the organization's ability to meet
the organization's own information security requirements.
The order in which requirements are presented in this document does not reflect their importance
or imply the order in which they are to be implemented. The list items are enumerated for reference
purpose only.

Source:./example.pdf
Page : 4
chunk_id : 38
Score : 1.204768419265747
-----
```

Figure 1: FAISS Query 1 - Output 1

```
-----  
Chunk :4  
  
Chunk Content:  
to the information security management system requirements;  
e) ensuring that the information security management system achieves its intended outcome(s);  
f) directing and supporting persons to contribute to the effectiveness of the information security  
management system;  
g) promoting continual improvement; and  
h) supporting other relevant management roles to demonstrate their leadership as it applies to their  
areas of responsibility.  
  
Source:./example.pdf  
  
Page : 7  
  
chunk_id : 53  
  
Score : 1.2773454189300537  
-----  
Chunk :5  
  
Chunk Content:  
organization's needs and objectives, security requirements, the organizational processes used and the  
size and structure of the organization. All of these influencing factors are expected to change over time.  
The information security management system preserves the confidentiality, integrity and availability  
of information by applying a risk management process and gives confidence to interested parties that  
risks are adequately managed.  
  
Source:./example.pdf  
  
Page : 4  
  
chunk_id : 36  
  
Score : 1.329432487487793
```

Figure 2: FAISS Query 1 - Output 2

8.2 Query 2

What are the requirements for establishing and communicating the information security policy?

Top K = 5

```
Query : What are the requirements for establishing and communicating the information security policy???
Top K :5
```

```
-----  
Chunk :1
```

```
Chunk Content:
```

- a) be consistent with the information security policy;
- b) be measurable (if practicable);
- c) take into account applicable information security requirements, and results from risk assessment and risk treatment;
- d) be monitored;
- e) be communicated;
- f) be updated as appropriate;
- g) be available as documented information.

```
The organization shall retain documented information on the information security objectives.
```

```
Source:./example.pdf
```

```
Page : 10
```

```
chunk_id : 69
```

```
Score : 0.4954074025154114
```

```
-----  
Chunk :2
```

```
Chunk Content:
```

```
d) includes a commitment to continual improvement of the information security management system.
```

```
The information security policy shall:
```

- e) be available as documented information;
- f) be communicated within the organization;
- g) be available to interested parties, as appropriate.

```
5.3 Organizational roles, responsibilities and authorities
```

```
Top management shall ensure that the responsibilities and authorities for roles relevant to information
```

```
Source:./example.pdf
```

```
Page : 8
```

```
chunk_id : 56
```

```
Score : 0.525383472442627
```

Figure 3: FAISS Query 2 - Output 1

```

-----
Chunk :3

Chunk Content:
requirements.
7.4 Communication
The organization shall determine the need for internal and external communications relevant to the information security management system including:
a) on what to communicate;
b) when to communicate;
c) with whom to communicate;
d) how to communicate.
7.5 Documented information
7.5.1 General
The organization's information security management system shall include:
a) documented information required by this document; and
    © ISO/IEC 2022 – All rights reserved

Source:./example.pdf

Page : 11

chunk_id : 75

Score : 0.5567135810852051
-----

```

```

Chunk :4

Chunk Content:
laws and regulations and contractual requirements.
5.35 Independent review of information security
Control
The organization's approach to managing information security and its implementation including people, processes and technologies shall be reviewed independently at planned intervals, or when significant changes occur.
5.36 Compliance with policies, rules
and standards for information
security
Control
Compliance with the organization's information security policy, top -

```

```

Source:./example.pdf

Page : 18

chunk_id : 119

Score : 0.6476861238479614

```

Figure 4: FAISS Query 2 - Output 2

```

Chunk :5

Chunk Content:
ISO/IEC 27001:2022(E)
5.2 Policy
Top management shall establish an information security policy that:
a) is appropriate to the purpose of the organization;
b) includes information security objectives (see 6.2) or provides the framework for setting information security objectives;
c) includes a commitment to satisfy applicable requirements related to information security;
d) includes a commitment to continual improvement of the information security management system.

Source:./example.pdf

Page : 8

chunk_id : 55

Score : 0.6742203831672668

```

Figure 5: FAISS Query 2 - Output 3

8.3 Query 3

What are the key steps involved in the information security risk assessment process?

Top K = 3

```
Query :What are the key steps involved in the information security risk assessment process?
Top K :3

-----
Chunk :1

Chunk Content:
materialize;
2) assess the realistic likelihood of the occurrence of the risks identified in 6.1.2 c) 1); and
3) determine the levels of risk;
e) evaluates the information security risks:
1) compare the results of risk analysis with the risk criteria established in 6.1.2 a); and
2) prioritize the analysed risks for risk treatment.
The organization shall retain documented information about the information security risk assessment
process.
6.1.3 Information security risk treatment

Source:./example.pdf

Page : 9
chunk_id : 63
Score : 0.4325544834136963
-----
Chunk :2

Chunk Content:
process.
NOTE 4 The information security risk assessment and treatment process in this document aligns with the
principles and generic guidelines provided in ISO 31000 [5].
6.2 Information security objectives and planning to achieve them
The organization shall establish information security objectives at relevant functions and levels.
The information security objectives shall:
a) be consistent with the information security policy;
b) be measurable (if practicable);

Source:./example.pdf

Page : 10
chunk_id : 68
Score : 0.477225661277771
```

Figure 6: FAISS Query 3 - Output 1

```
-----  
Chunk :3  
  
Chunk Content:  
The organization shall perform information security risk assessments at planned intervals or when  
significant changes are proposed or occur, taking account of the criteria established in 6.1.2 a).  
The organization shall retain documented information of the results of the information security risk  
assessments.  
8.3 Information security risk treatment  
The organization shall implement the information security risk treatment plan.  
  
Source:./example.pdf  
  
Page : 13  
  
chunk_id : 85  
  
Score : 0.4804421067237854
```

Figure 7: FAISS Query 3 - Output 2