Documentation for Qdrant

Retrieval Augmented Generation using Qdrant Vector Database

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_qdrant.git
cd celestial_systems_qdrant
pip install -r requirements.txt
python rag_qdrant.py
```

3 Libraries Required

```
from langchain_community.document_loaders import PyPDFLoader
from langchain_text_splitters import
   RecursiveCharacterTextSplitter
from langchain_huggingface import HuggingFaceEmbeddings
from qdrant_client import QdrantClient
from qdrant_client.http.models import VectorParams, Distance,
   PointStruct
import re
import pprint
```

3.1 Library Description

- **PyPDFLoader** parses PDF documents to extract textual content.
- **RecursiveCharacterTextSplitter** splits documents into overlapping chunks.
- **HuggingFaceEmbeddings** uses the `sentence-transformers/all-MiniLM-L6-v2` model to generate vector embeddings.

- **QdrantClient** initializes and manages the Qdrant vector database.
- **VectorParams, Distance, PointStruct** configure vector size, similarity metric, and define the data structure used to insert points into Qdrant.

4 Loading the PDF Document

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

loader = PyPDFLoader(filepath)
print(loader)

docs = loader.load()
```

The above code demonstrates loading a PDF document using LangChain's `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 the Document Chunks

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

texts = [chunk.page_content for chunk in chunks]
metadata = [chunk.metadata for chunk in chunks]
ids = [str(chunk.metadata["chunk_id"]) for chunk in chunks]
```

```
vectors = embeddings.embed_documents(texts)
```

The embeddings are generated using an industry-standard MiniLM transformer model.

7 Creating the Qdrant Client and Preparing Data

```
client = QdrantClient(path="./vectordb/qdrant")

client.recreate_collection(
    collection_name="ragdata",
    vectors_config=VectorParams(
        size=384,
        distance=Distance.COSINE
    )
)

points = []
for chunk, vector in zip(chunks, vectors):
    points.append(
        PointStruct(
            id=chunk.metadata["chunk_id"],
            vector=vector,
            payload={
                "text": chunk.page_content,
                **chunk.metadata
            }
        )
    )
```

- A local Qdrant client is initialized
- A collection named `ragdata` is created
- Vector size is set to 384 and similarity is computed using cosine distance
- Document chunks are converted into `PointStruct` objects

8 Inserting Data into Qdrant

```
client.upsert(
    collection_name="ragdata",
    points=points
)

info = client.get_collection("ragdata")
print(info.points_count)
```

The `upsert` operation inserts the prepared vectors and metadata into the Qdrant collection.

9 Querying the Qdrant Vector Database

```
query = "How␣does␣top␣management␣demonstrate␣leadership␣and␣
   commitment␣to␣the␣ISMS?"

query_embedding = embeddings.embed_query(query)

results = client.query_points(
    collection_name="ragdata",
    query=query_embedding,
    limit=5
)

for index, i in enumerate(results.points):
    print("\n" + "-" * 40)
    print(f"Result␣:␣{index+1}")
    print(f"Chunk␣ID␣:␣{i.id}")
    print(f"Chunk␣Content␣:␣{i.payload['text']}")
    print(f"Score␣:␣{i.score}")
    print(f"Page␣:␣{i.payload['page']}")
    print(f"Source␣:␣{i.payload['source']}")

client.close()
```

The query is embedded using the same model as the document chunks and then used to retrieve the most relevant vectors from Qdrant.

10 Outputs

10.1 Query 1

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

```
The Query :  How does top management demonstrate leadership and commitment to the ISMS?



---------------------------------------

Result : 1

Chunk ID : 51

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;

Score : 0.5385089103271771

Page :7

Source : ./example.pdf



---------------------------------------

Result : 2

Chunk ID : 57

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

Score : 0.506787221224803

Page :8

Source : ./example.pdf



---------------------------------------
```

Figure 1: Qdrant Query 1 - Output 1

Figure 2: Qdrant Query 1 - Output 2

10.2 Query 2

What are the requirements for establishing and communicating the information security policy?
Top K = 5

```
The Query :  What are the requirements for establishing and communicating the information security policy?



----------------------------------------

Result : 1

Chunk ID : 69

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.

Score : 0.7544924035586212

Page :10

Source : ./example.pdf



----------------------------------------

Result : 2

Chunk ID : 56

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

Score : 0.7383052704048987

Page :8

Source : ./example.pdf



----------------------------------------
```

Figure 3: Qdrant Query 2 - Output 1

```
Result : 3

Chunk ID : 75

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

Score : 0.7065966622044351

Page :11

Source : ./example.pdf



----------------------------------------

Result : 4

Chunk ID : 119

Chunk Content :laws and regulations and contractual requirements.
5.35 Independent review of informa-
tion 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 -

Score : 0.6591124526267529

Page :18

Source : ./example.pdf



----------------------------------------
```

Figure 4: Qdrant Query 2 - Output 2

Figure 5: Qdrant Query 2 - Output 3

10.3 Query 3

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

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



---------------------------------------

Result : 1

Chunk ID : 63

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

Score : 0.7837227773917174

Page :9

Source : ./example.pdf



---------------------------------------

Result : 2

Chunk ID : 68

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);

Score : 0.76138721445966

Page :10

Source : ./example.pdf



---------------------------------------
```

Figure 6: Qdrant Query 3 - Output 1

```
---------------------------------------

Result : 3

Chunk ID : 85

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.

Score : 0.7597790140214663

Page :13

Source : ./example.pdf
```

Figure 7: Qdrant Query 3 - Output 2