

LangChain: Q&A over Documents

An example might be a tool that would allow you to query a product catalog for items of interest.

```
In [1]: #pip install --upgrade langchain
```

```
In [2]: import os

from dotenv import load_dotenv, find_dotenv
_ = load_dotenv(find_dotenv()) # read local .env file
```

```
In [3]: from langchain.chains import RetrievalQA
from langchain.chat_models import ChatOpenAI
from langchain.document_loaders import CSVLoader
from langchain.vectorstores import DocArrayInMemorySearch
from IPython.display import display, Markdown
```

```
In [4]: file = 'OutdoorClothingCatalog_1000.csv'
loader = CSVLoader(file_path=file)
```

```
In [5]: from langchain.indexes import VectorstoreIndexCreator
```

```
In [6]: #pip install docarray
```

```
In [7]: index = VectorstoreIndexCreator(
    vectorstore_cls=DocArrayInMemorySearch
).from_loaders([loader])
```

```
In [8]: query = "Please list all your shirts with sun protection \
in a table in markdown and summarize each one."
```

```
In [9]: response = index.query(query)
```

```
In [10]: display(Markdown(response))
```

Name	Description
Men's Tropical Plaid Short-Sleeve Shirt	UPF 50+ rated, 100% polyester, wrinkle-resistant, front and back cape venting, two front bellows pockets
Men's Plaid Tropic Shirt, Short-Sleeve	UPF 50+ rated, 52% polyester and 48% nylon, machine washable and dryable, front and back cape venting, two front bellows pockets
Men's TropicVibe Shirt, Short-Sleeve	UPF 50+ rated, 71% Nylon, 29% Polyester, 100% Polyester knit mesh, wrinkle resistant, front and back cape venting, two front bellows pockets
Sun Shield Shirt by	UPF 50+ rated, 78% nylon, 22% Lycra Xtra Life fiber, wicks moisture, fits comfortably over swimsuit, abrasion resistant

All of the shirts provide UPF 50+ sun protection, blocking 98% of the sun's harmful rays. The fabrics are wrinkle-resistant and quick-drying, and all feature front and back cape venting and two front bellows pockets. The

```
In [11]: loader = CSVLoader(file_path=file)
```

```
In [12]: docs = loader.load()
```

```
In [13]: docs[0]
```

```
Document(page_content=": 0\nname: Women's Campside Oxfords\ndescription: This ultracomfortable lace-to-toe Oxford boasts a super-soft canvas, thick cushioning, and quality construction for a broken-in feel from the first time you put them on. \n\nSize & Fit: Order regular shoe size. For half sizes not offered, order up to next whole size. \n\nSpecs: Approx. weight: 1 lb.1 oz. per pair. \n\nConstruction: Soft canvas material for a broken-in feel and look. Comfortable EVA innersole with Cleansport NXT® antimicrobial odor control. Vintage hunt, fish and camping motif on innersole. Moderate arch contour of innersole. EVA foam midsole for cushioning and support. Chain-tread-inspired molded rubber outsole with modified chain-tread pattern. Imported. \n\nQuestions? Please contact us for any inquiries.", metadata={'source': 'OutdoorClothingCatalog_1000.csv', 'row': 0})
```

```
In [14]: from langchain.embeddings import OpenAIEmbeddings
         embeddings = OpenAIEmbeddings()
```

```
In [15]: embed = embeddings.embed_query("Hi my name is Harrison")
```

```
In [16]: print(len(embed))
```

1536

```
In [17]: print(embed[:5])
```

```
[-0.021913960576057434, 0.006774206645786762, -0.018190348520874977, -0.039148248732089996, -0.014089343138039112]
```

```
In [18]: db = DocArrayInMemorySearch.from_documents(
        docs,
        embeddings
    )
```

```
In [19]: query = "Please suggest a shirt with sunblocking"
```

```
In [20]: docs = db.similarity_search(query)
```

```
In [21]: len(docs)
```

4

```
In [22]: docs[0]
```

```
Document(page_content=': 255\nname: Sun Shield Shirt by\ndescription: "Block the sun, not the fun - our high-performance sun shirt is guaranteed to protect from harmful UV rays. \n\nSize & Fit: Slightly Fitted: Softly shapes the body. Falls at hip.\n\nFabric & Care: 78% nylon, 22% Lycra Xtra Life fiber. UPF 50+ rated - the highest rated sun protection possible. Handwash, line dry.\n\nAdditional Features: Wicks moisture for quick-drying comfort. Fits comfortably over your favorite swimsuit. Abrasion resistant for season after season of wear. Imported.\n\nSun Protection That Won\'t Wear Off\nOur high-performance fabric provides SPF 50+ sun protection, blocking 98% of the sun\'s harmful rays. This fabric is recommended by The Skin Cancer Foundation as an effective UV protectant.', metadata={'source': 'OutdoorClothingCatalog_1000.csv', 'row': 255})
```

```
In [23]: retriever = db.as_retriever()
```

```
In [24]: llm = ChatOpenAI(temperature = 0.0)
```

```
In [25]: qdocs = "".join([docs[i].page_content for i in range(len(docs))])
```

```
In [26]: response = llm.call_as_llm(f"{qdocs} Question: Please list all your \n\nshirts with sun protection in a table in markdown and summarize each one.")
```

```
In [27]: display(Markdown(response))
```

Name	Description
Sun Shield Shirt	High-performance sun shirt with UPF 50+ sun protection, moisture-wicking, and abrasion-resistant fabric. Fits comfortably over swimsuits. Recommended by The Skin Cancer Foundation.
Men's Plaid Tropic Shirt	Ultracomfortable shirt with UPF 50+ sun protection, wrinkle-free fabric, and front/back cape venting. Made with 52% polyester and 48% nylon.
Men's TropicVibe Shirt	Men's sun-protection shirt with built-in UPF 50+ and front/back cape venting. Made with 71% nylon and 29% polyester.
Men's Tropical Plaid Short-Sleeve Shirt	Lightest hot-weather shirt with UPF 50+ sun protection, front/back cape venting, and two front bellows pockets. Made with 100% polyester and is wrinkle-resistant.

All of these shirts provide UPF 50+ sun protection, blocking 98% of the sun's harmful rays. They are made with high-performance fabrics that are moisture-wicking, abrasion-resistant, and/or wrinkle-free. Some have front/back cape venting for added comfort in hot weather. The Sun Shield Shirt is recommended by The Skin Cancer Foundation.

```
In [28]: qa_stuff = RetrievalQA.from_chain_type(
          llm=llm,
          chain_type="stuff",
          retriever=retriever,
          verbose=True
        )
```

```
In [29]: query = "Please list all your shirts with sun protection in a table \
in markdown and summarize each one."
```

```
In [30]: response = qa_stuff.run(query)
```

> Entering new RetrievalQA chain...

> Finished chain.

```
In [31]: display(Markdown(response))
```

Shirt Number	Name	Description
618	Men's Tropical Plaid Short-Sleeve Shirt	This shirt is made of 100% polyester and is wrinkle-resistant. It has front and back cape venting that lets in cool breezes and two front bellows pockets. It is rated UPF 50+ for superior protection from the sun's UV rays.
374	Men's Plaid Tropic Shirt, Short-Sleeve	This shirt is made with 52% polyester and 48% nylon. It is machine washable and dryable. It has front and back cape venting, two front bellows pockets, and is rated to UPF 50+.
535	Men's TropicVibe Shirt, Short-Sleeve	This shirt is made of 71% Nylon and 29% Polyester. It has front and back cape venting that lets in cool breezes and two front bellows pockets. It is rated UPF 50+ for superior protection from the sun's UV rays.
255	Sun Shield Shirt	This shirt is made of 78% nylon and 22% Lycra Xtra Life fiber. It is handwashable and line dry. It is rated UPF 50+ for superior protection from the sun's UV rays. It is abrasion-resistant and wicks moisture for quick-drying comfort.

All of these shirts provide UPF 50+ sun protection, blocking 98% of the sun's harmful rays. They are all designed to be lightweight and comfortable in hot weather. They all have front and back cape venting that lets in cool breezes and two front bellows pockets. The Men's Tropical Plaid Short-Sleeve Shirt is made of 100% polyester and is wrinkle-resistant. The Men's Plaid Tropic Shirt, Short-Sleeve is made with 52% polyester and 48% nylon. The Men's TropicVibe Shirt, Short-Sleeve is made of 71% Nylon and 29% Polyester. The Sun Shield Shirt is made of 78% nylon and 22% Lycra Xtra Life fiber and is abrasion-resistant and wicks moisture for quick-drying comfort.

```
In [32]: response = index.query(query, llm=llm)
```

```
In [33]: index = VectorstoreIndexCreator(  
        vectorstore_cls=DocArrayInMemorySearch,  
        embedding=embeddings,  
    ).from_loaders([loader])
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

