

[Open in app ↗](#)

You have **1 free member-only story left** this month. [Upgrade](#) for unlimited access.

◆ Member-only story

How to use Private LLM GPT4All with LangChain



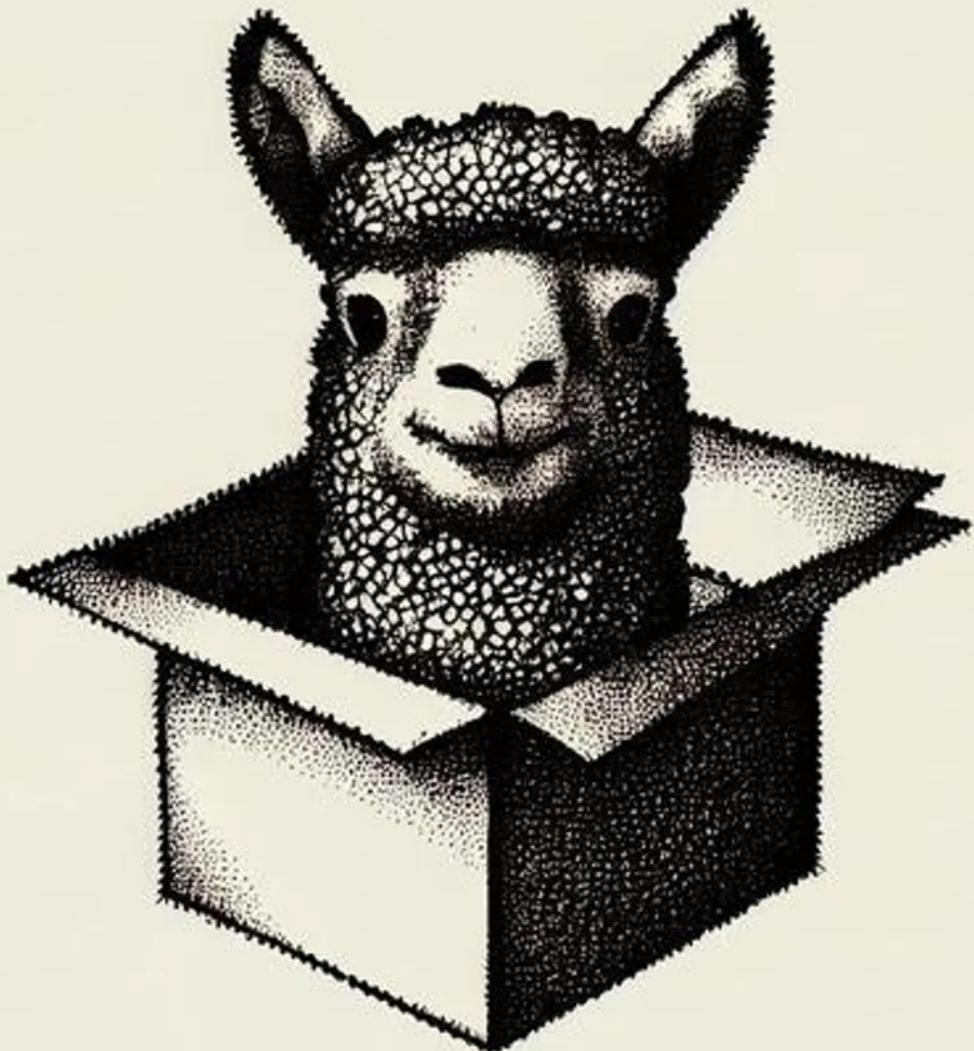
Meta Heuristic · [Follow](#)

3 min read · May 29

Listen

Share

More



<https://metaheuristic.co>

LangChain, a language model processing library, provides an interface to work with various AI models including OpenAI's gpt-3.5-turbo and Private LLM gpt4all. It enables users to embed documents, retrieve similar documents, and use document retrieval to augment Language Model conversations. This tutorial walks you through the process of using Private LLM gpt4all with LangChain to perform information extraction from PDF documents.

Prerequisites

Ensure that you have the following installed:

- langchain==0.0.173
- chromadb==0.3.23

- pypdf==3.8.1
- pygpt4all==1.1.0
- pdf2image==1.16.3
- poppler-utils

These packages are essential for processing PDFs, generating document embeddings, and using the gpt4all model. Poppler-utils is particularly important for converting PDF pages to images.

Step 1: Load the PDF Document

First, we need to load the PDF document. We use LangChain's `PyPDFLoader` to load the document and split it into individual pages.

```
from langchain.document_loaders import PyPDFLoader
loader = PyPDFLoader("ms-financial-statement.pdf")
documents = loader.load_and_split()
```

Step 2: Text Splitting

Next, we split the text into manageable chunks for the AI model. We use LangChain's `RecursiveCharacterTextSplitter` for this task.

```
from langchain.text_splitter import RecursiveCharacterTextSplitter
text_splitter = RecursiveCharacterTextSplitter(chunk_size=1024, chunk_overlap=64)
texts = text_splitter.split_documents(documents)
```

Step 3: Creating Embeddings

We then create embeddings of the split text using HuggingFaceEmbeddings. This step creates a vector representation of each text chunk.

```
from langchain.embeddings import HuggingFaceEmbeddings
embeddings = HuggingFaceEmbeddings(model_name="sentence-transformers/all-MiniLM-L6-v1")
```

Step 4: Vector Store

We use the embeddings to create a vector store using Chroma. Chroma is a high-performance in-memory vector database designed for machine learning workloads.

```
from langchain.vectorstores import Chroma
db = Chroma.from_documents(texts, embeddings, persist_directory="db")
```

Step 5: Load the gpt4all Model

We load the gpt4all model using LangChain's `GPT4All` class.

```
from langchain.llms import GPT4All
model_path = "./ggml-gpt4all-j-v1.3-groovy.bin"
llm = GPT4All(model=model_path, n_ctx=1000, backend="gptj", verbose=False)
```

Step 6: Create the Question-Answering Chain

We create a question-answering chain using LangChain's `RetrievalQA` class. The retriever uses the Chroma vector store we created earlier.

```
from langchain.chains import RetrievalQA
qa = RetrievalQA.from_chain_type(
    llm=llm,
    chain_type="stuff",
    retriever=db.as_retriever(search_kwargs={"k": 3}),
    return_source_documents=True,
    verbose=False,
)
```

Step 7: Ask Questions

Finally, we can now ask questions. The question-answering chain will retrieve similar documents from the vector store, use the gpt4all model to generate responses, and return the responses.

```
res = qa("How much is the dividend per share during 2022? Extract it from the text.")
```

This would generate a response like:

```
{'query': 'How much is the dividend per share during 2022? Extract it from the text.',  
 'result': 'The dividend per share during 2022 is $0.62.',  
 'source_documents': [Document(page_content='...', metadata={'source': 'ms-financials'})]}
```

You can print the result as follows:

```
print(res["result"])
```

This will output:

```
The dividend per share during 2022 is $0.62.
```

Conclusion

This tutorial walked you through using Private LLM gpt4all with LangChain. It included a step-by-step guide to loading and processing PDFs, generating embeddings, creating a vector store, and creating a question-answering chain. LangChain is a versatile and robust tool that allows you to leverage AI models, like Private LLM gpt4all, in various ways. Whether you're trying to extract specific information from a document or aiming to augment your applications with AI, LangChain provides a streamlined and efficient approach.

...

Disclaimer: The code and models presented in this blog post are illustrative and not intended for production use without additional measures for handling edge cases, errors, and security.

Checkout ETL for Vector databases: <https://metaheuristic.co>

Machine Learning

Data Science

Langchain

Large Language Models

ChatGPT

[Follow](#)

Written by Meta Heuristic

454 Followers

ML design, best practices and optimizations. ETL for LLM <https://metaheuristic.co>

More from Meta Heuristic

```
from fastapi import FastAPI
from langcorn import create_service

app: FastAPI = create_service(
    "examples.ex1:chain",
    "examples.ex2:chain",
    "examples.ex3:chain",
    "examples.ex4:sequential_chain",
    "examples.ex5:conversation",
    "examples.ex6:conversation_with_summary",
)
```

 Meta Heuristic

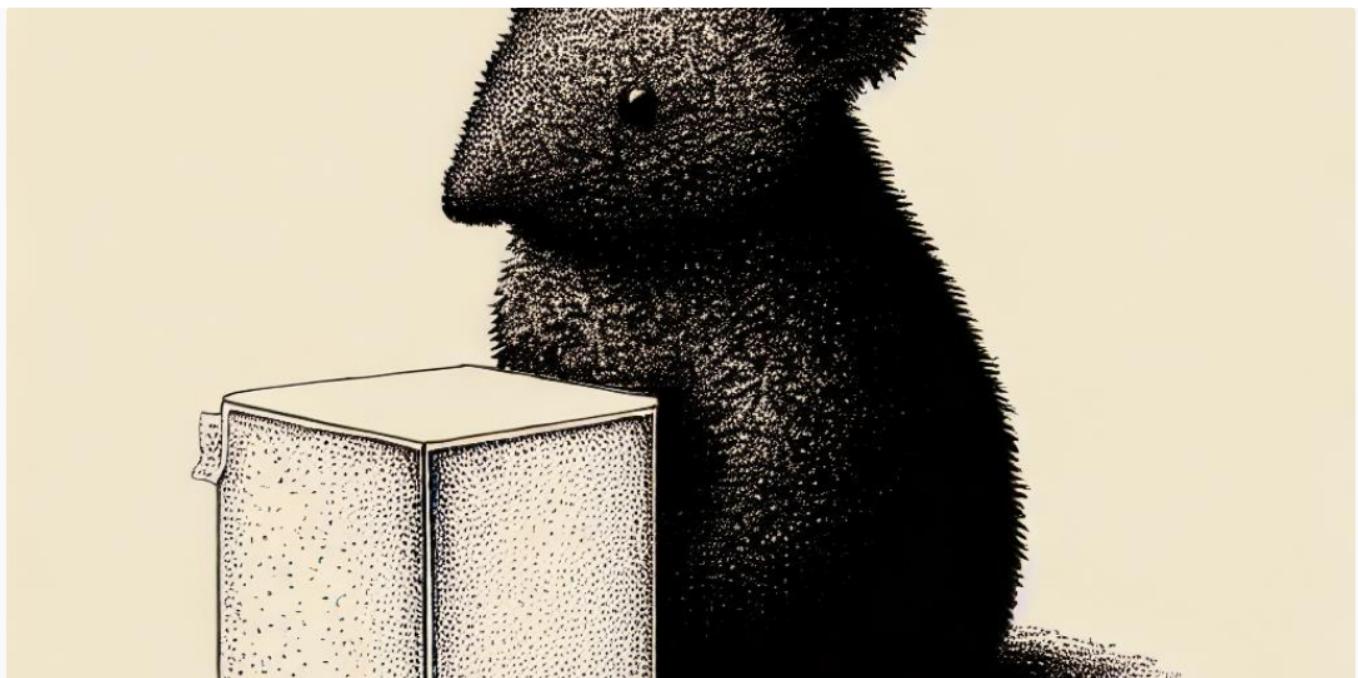
Deploying LangChain with FastAPI to Vercel: Your Step-By-Step Guide

Deploying models and pipelines has never been this easy with LangCorn, an API server designed to leverage the power of the FastAPI...

3 min read · May 23



...

 Meta Heuristic

How to use StructuredOutputParser with LangChain

Langchain's StructuredOutputParser is a powerful tool for working with AI language models. By allowing you to define the structure of the...

◆ · 4 min read · Jun 12

 2

...

 Meta Heuristic

How to cache LLM calls in LangChain

As the demand for machine learning models and data processing pipelines increases, there is a growing need to optimize these systems for...

◆ · 3 min read · Jun 15

👏 2



Bookmark +

...

 Meta Heuristic

How to setup Token Usage Tracking in LangChain

LangChain, a powerful framework for designing language models, allows developers to orchestrate complex Natural Language Processing (NLP)...

◆ · 2 min read · Jun 16

👏 17

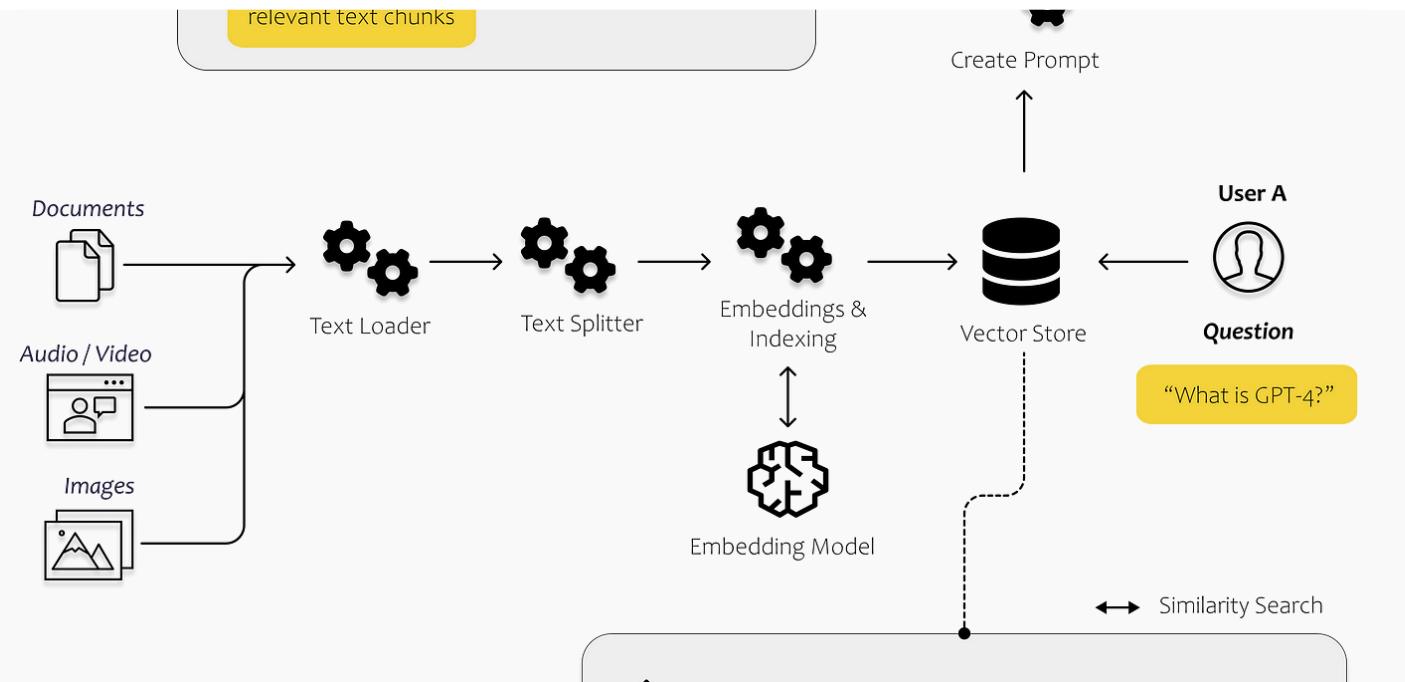


Bookmark +

...

See all from Meta Heuristic

Recommended from Medium



Dominik Polzer in Towards Data Science

All You Need to Know to Build Your First LLM App

A step-by-step tutorial to document loaders, embeddings, vector stores and prompt templates

★ · 26 min read · Jun 22

3.3K 28



...



Wei-Meng Lee in Level Up Coding

Training Your Own LLM using privateGPT

Learn how to train your own language model without exposing your private data to the provider

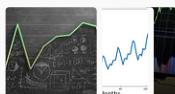
★ · 8 min read · May 19

1.2K 10



...

Lists



Predictive Modeling w/ Python

18 stories · 169 saves



Practical Guides to Machine Learning

10 stories · 187 saves



The New Chatbots: ChatGPT, Bard, and Beyond

13 stories · 57 saves



Natural Language Processing

438 stories · 80 saves



Q Leonie Monigatti in Towards Data Science

Getting Started with LangChain: A Beginner's Guide to Building LLM-Powered Applications

A LangChain tutorial to build anything with large language models in Python

★ · 12 min read · Apr 25

3.6K

23



...



Mirantha Jayathilaka, PhD

Building a Chat-AI to answer about your own data—Part I

First things first

◆ · 6 min read · Feb 22

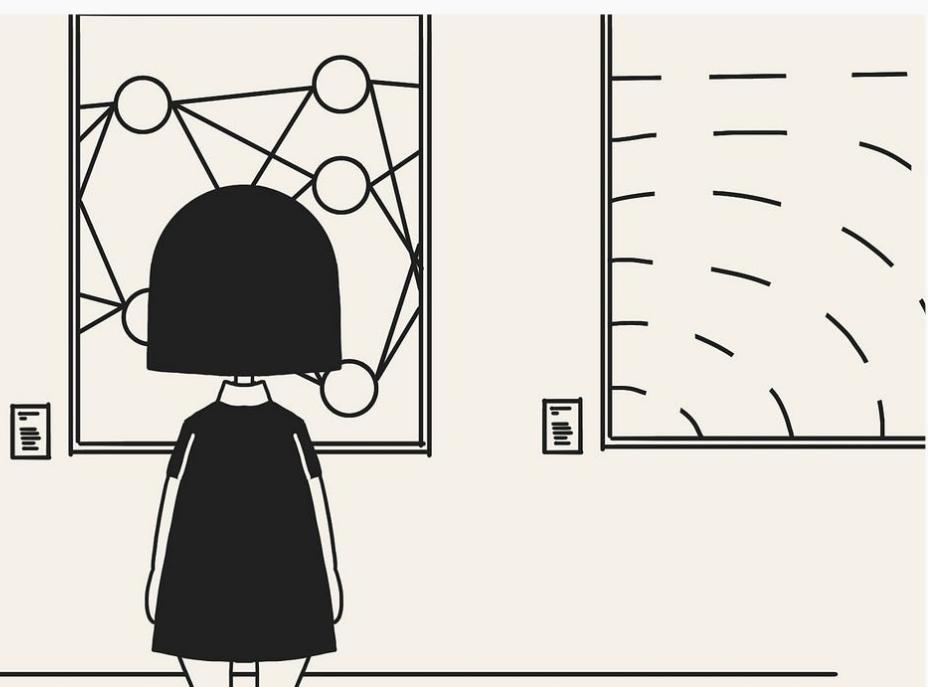
83

Q 1



...

```
000100010001000100010001111
001110011110000011111010000
011111000101101110011110000/
101001010001000100010001000
110101010111011011110101010
1010110101010111101110001
01101010100010000001000100
010002000001000100010001000
100011110011101111000001111
101000001111100010110111001
1110000/10100101000100010001
000100011010101011101101111
0101010101101010101011110
111000101101010101000100000
100010010002000001000100010
001000100011110011101111000.
```



Leonie Monigatti in Towards Data Science

10 Exciting Project Ideas Using Large Language Models (LLMs) for Your Portfolio

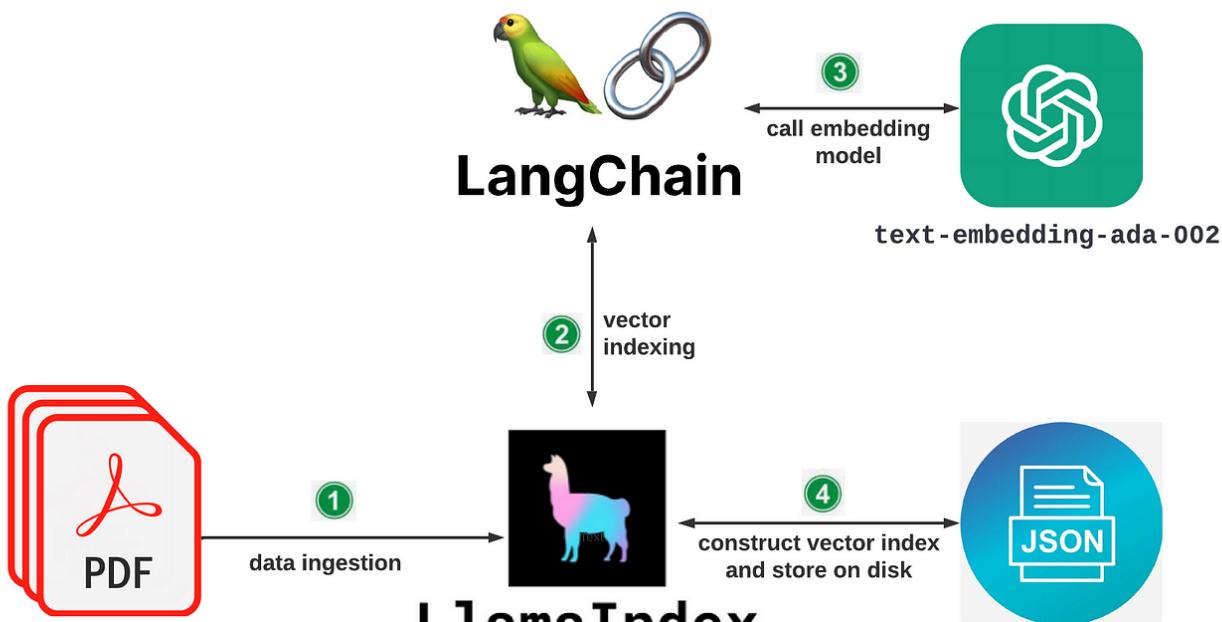
Learn how to build apps and showcase your skills with large language models (LLMs). Get started today!

◆ · 11 min read · May 15

👏 1.8K 🎧 11



...



👤 Wenqi Glantz in Better Programming

Building Your Own DevSecOps Knowledge Base with OpenAI, LangChain, and LlamaIndex

Building your custom knowledge base chatbot

◆ · 11 min read · Jun 24

👏 928 🎧 10



...

See more recommendations