



Making smart queries on world history

Modern AI Pro: Level 1



Balaji Viswanathan



Key goals.



Load an external dataset

We will use a world history dataset.



Convert to SQL

The format most of the world's data is in.



Visualize data

With Word Cloud and Geopandas.



Generate Text Embeddings

Generate vectors from text.




Relational data to vector DB

How do we make SQL DB highly usable.



RAG with LLM

Make LLM use multiple data chunks.





How will you use this?



Your company might want to utilize the massive chunks of relational data sitting in its databases and tap its full value. Using visualization techniques and LLM based RAGs you can dazzle your management and customers.

Traditional Search

```
SELECT question
FROM JeopardyQuestions
WHERE (question LIKE '%dog%' OR
       question LIKE '%cat%' OR
       question LIKE '%wolf%' OR
       (... and so on))
```

VS

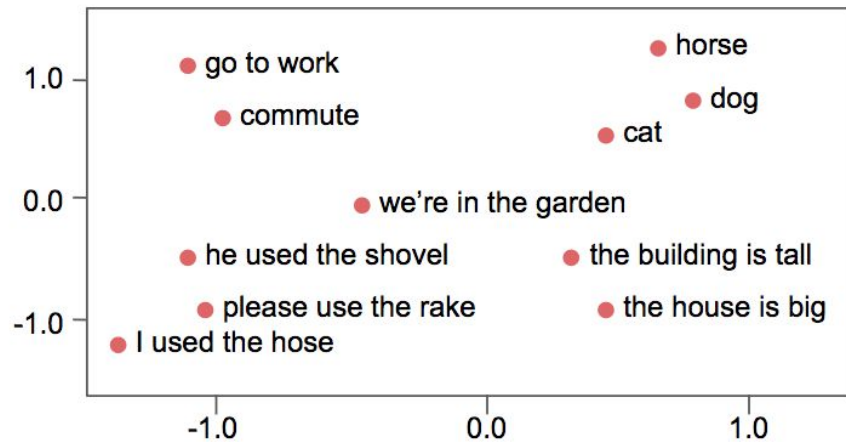
Semantic Search

```
{
  Get {
    JeopardyQuestions (
      nearText: { concepts: ["animals"] }
    ) { question }
  }
}
```



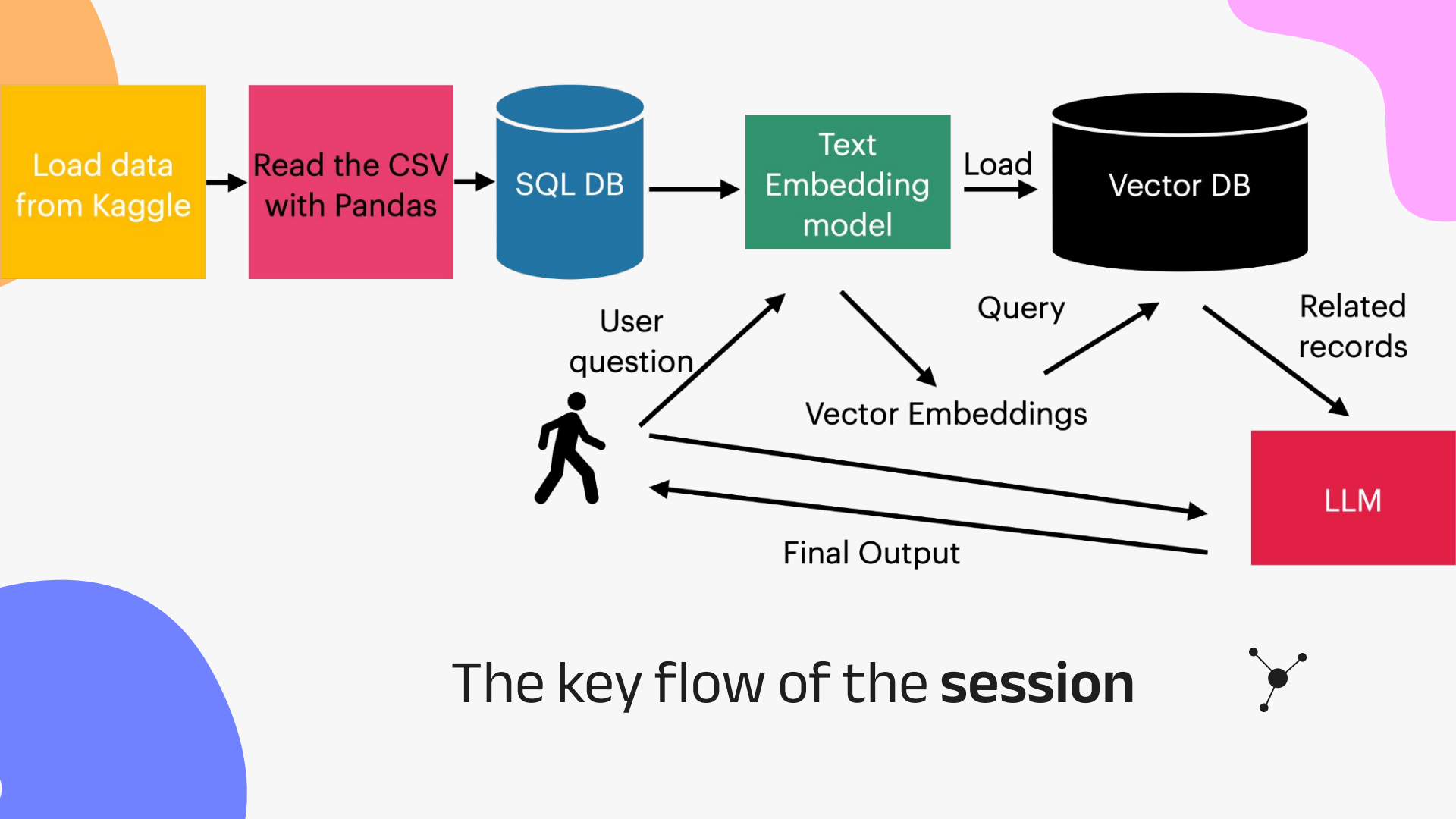


that rug really tied the room together

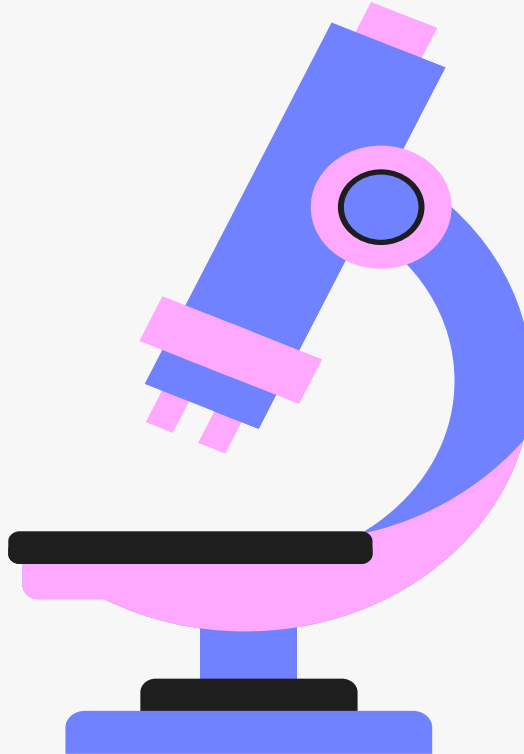
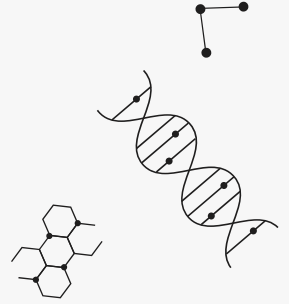





A new way of thinking about data.





Let's get **started**



→  <https://www.kaggle.com/settings>  

Create

Home

Competitions

Datasets

Models

Code

Discussions

Learn

More

Your Work

VIEWED

Real / Fake Job Postin...

News Category Dataset

Amazon Reviews for S...

COVID-19 Open Resea...

Search

Settings

Control over your Kaggle account and all communications

Account Notifications

Your email address

balaji@mitrarobot.com

Change email

Phone verification

Verified

API

Using Kaggle's beta API, you can interact with Competitions and Datasets to download data, make submissions, and more via the command line. [Read the docs](#)

Create New Token Expire Token

Quotas

Your usernar
balajiviswanath

Your account
10130253

```
! pip -q install kaggle
! mkdir ~/.kaggle
! cp kaggle.json ~/.kaggle
! chmod 600 ~/.kaggle/kaggle.json

! kaggle datasets download -d
'saketk511/world-important-events-a
ncient-to-modern'

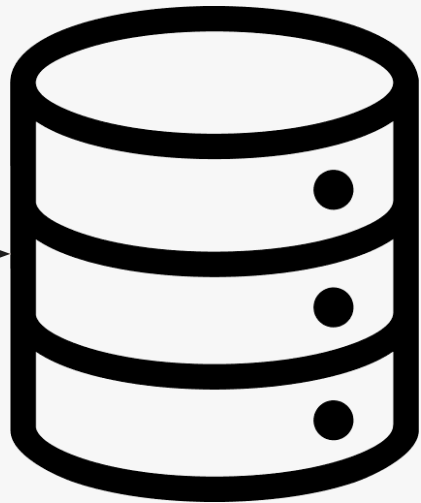
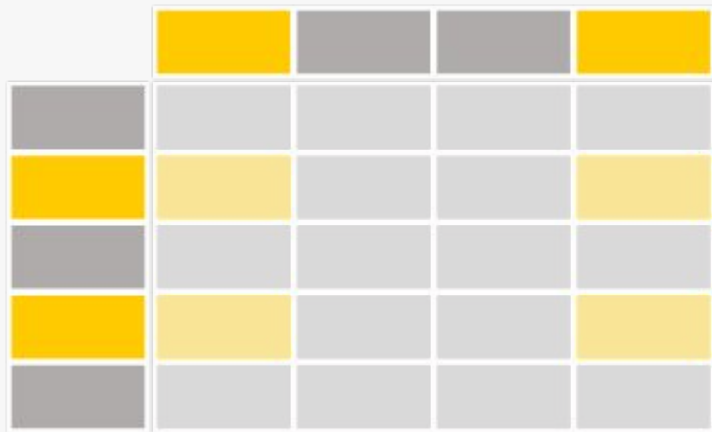
! unzip
/content/world-important-events-anc
ient-to-modern.zip
```

1. Get the dataset from Kaggle

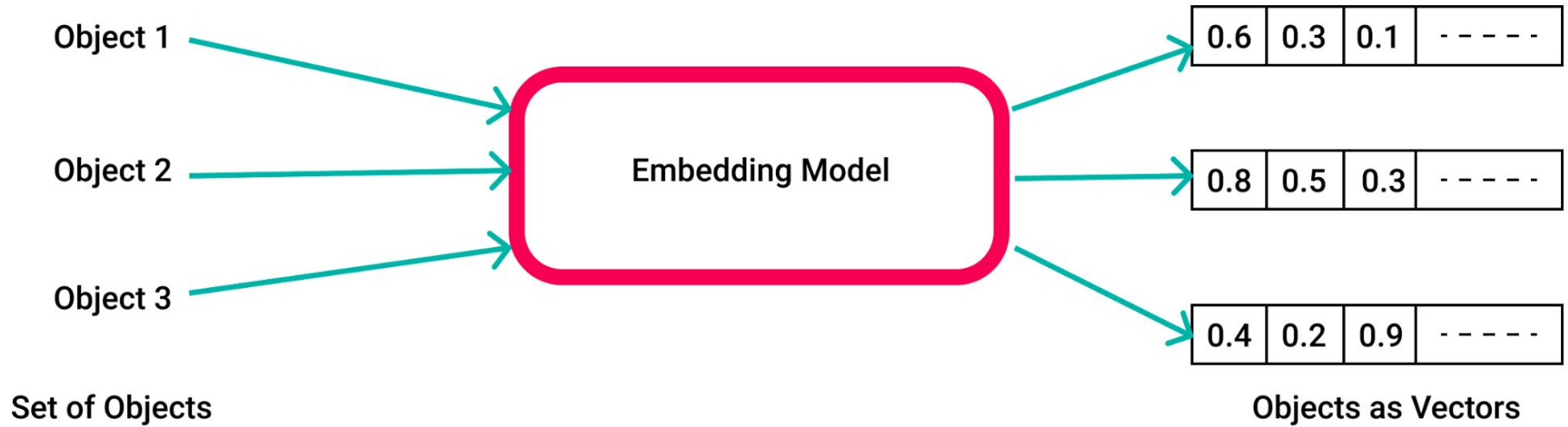

```
import pandas as pd
```



2. Load the data with Pandas



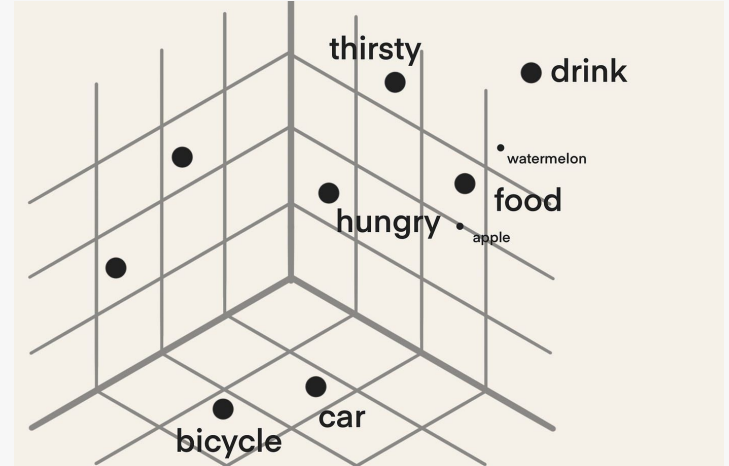
4. From Dataframe to SQL DB

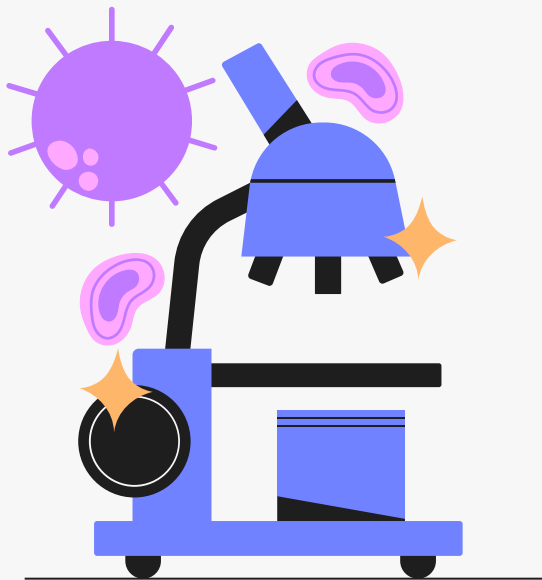


5. Use an embedding model to generate vectors

6. Load the embeddings into the Vector DB and query.

We are grouping data by concepts rather than mere text.

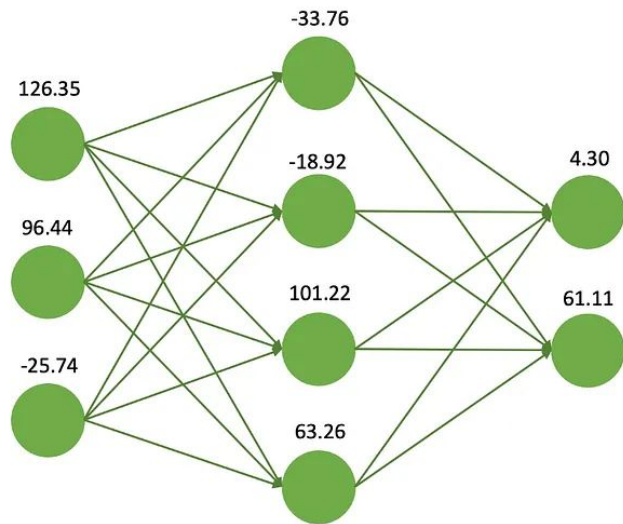




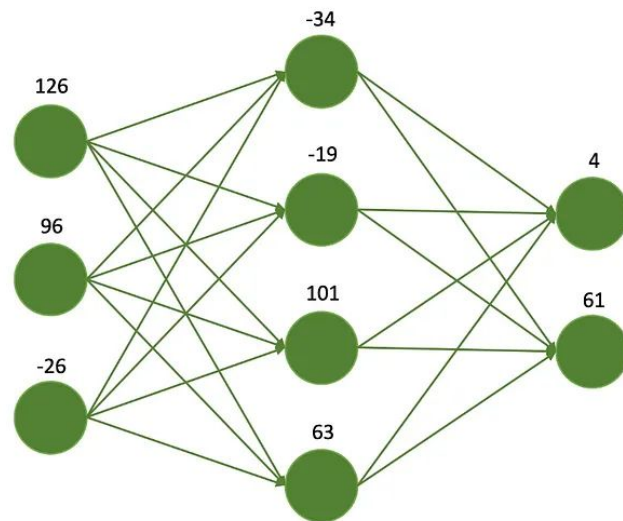
Get the LLMs work on your data!

Can you combine your internal data in
incredible new ways?





Before



After

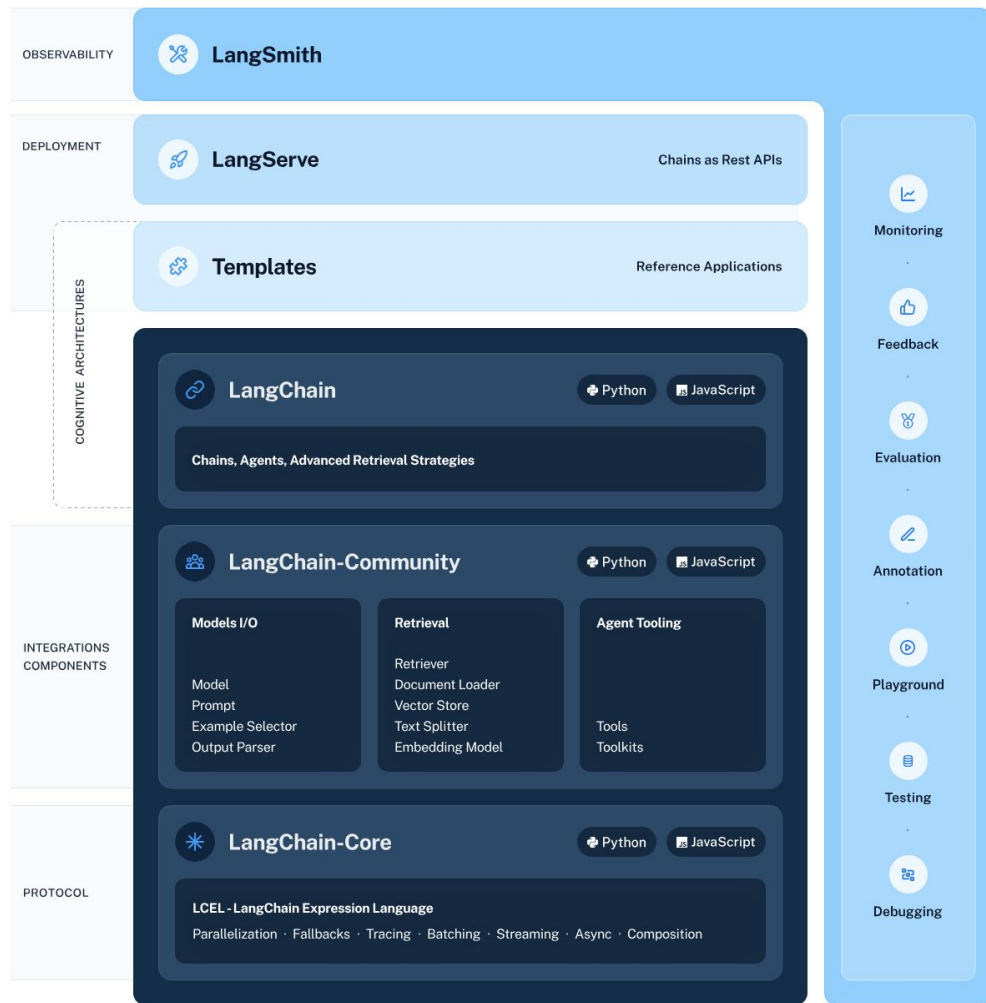
7. Quantize the chosen model [Mistral]

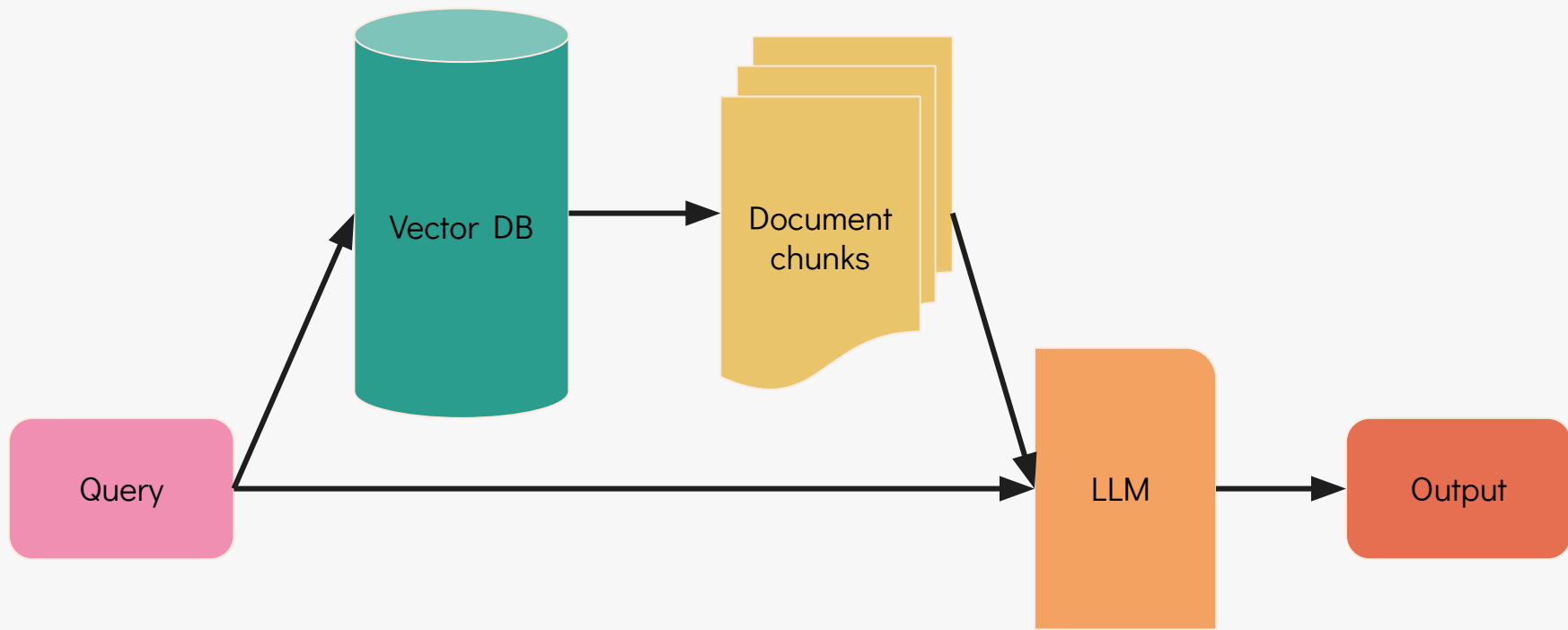


8. Load the **LLM** in memory.



9. Using the Langchain Framework





10. Setting up a basic **RAG** pipeline.
RAG: Retriever Augmented Generation.

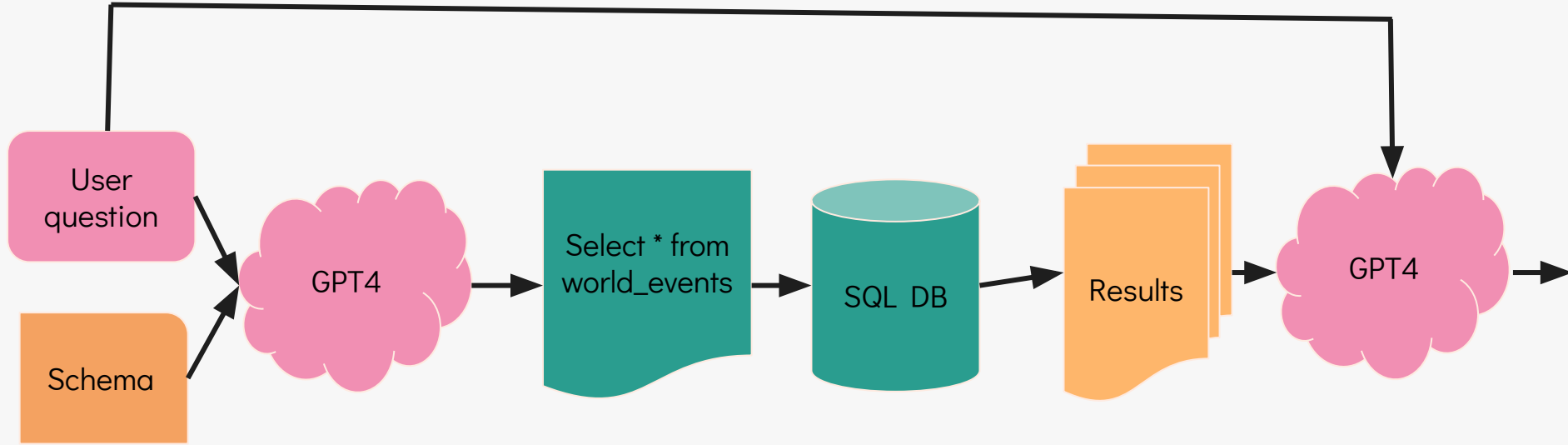
Answer:

* Partition of India: Impact of the event:India/Pakistan
* Sino-Indian War: Impact of the event:Border regions of India
* Indian Independence: Impact of the event:India
* Partition of Bengal: Impact of the event:Bengal

These events are connected in that they all involve conflicts that have occurred in India or its neighboring regions. The Partition of India and Indian Independence were major events that led to the formation of new countries and changed the political landscape of the region. The Sino-Indian War and the Partition of Bengal were conflicts that had a significant impact on the border regions of India.

1. **Partition of India (1947)**
 - **Impact of the event:** The partition of India led to the creation of two separate nations, India and Pakistan. This event resulted in one of the largest mass migrations in human history and significant communal violence.
 - **Place Name:** The partition affected multiple regions but primarily Punjab and Bengal, which were split between the two new nations.
 - **Type of Event:** It was a political and social upheaval event.

11. Calling GPT4 after setting up Google Colab Secrets.



12. Generating and executing SQL queries with GPT 4.



Welcome to our world of AI!