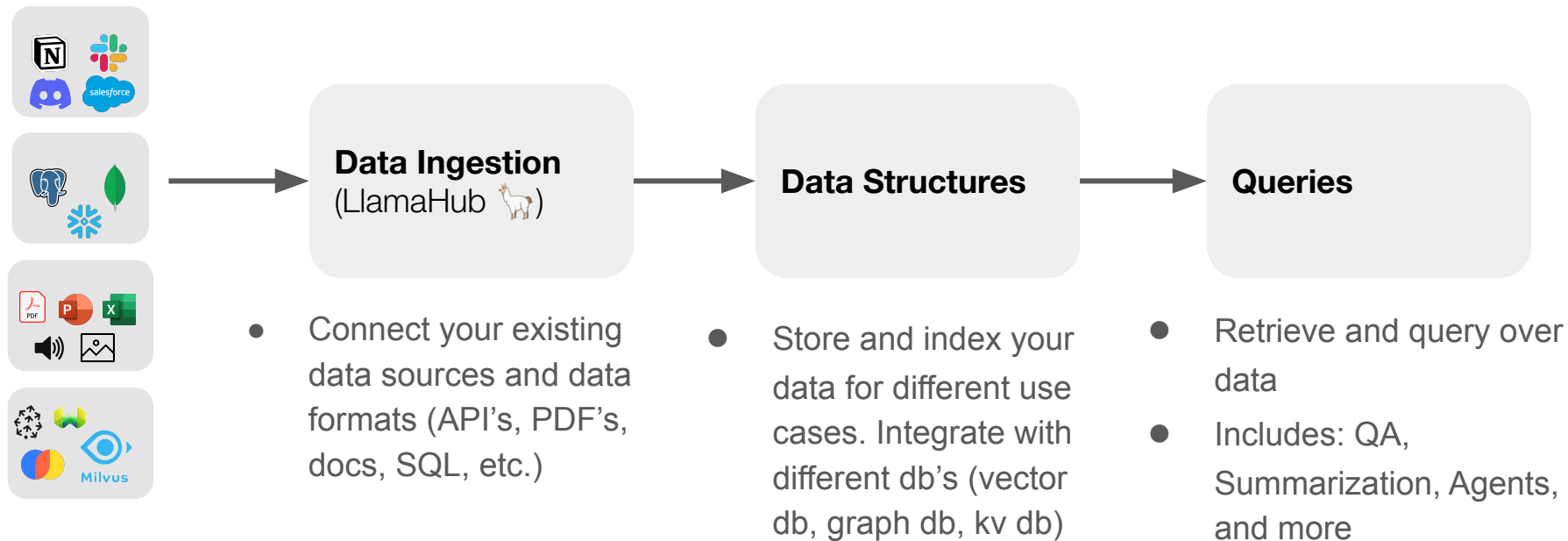# Building LLM Apps over Financial Data

Jerry Liu, LlamaIndex co-founder/CEO

https://github.com/run-llama/llama_index

LlamaIndex: A data framework for LLM applications

- Data Management and Query Engine for your LLM application
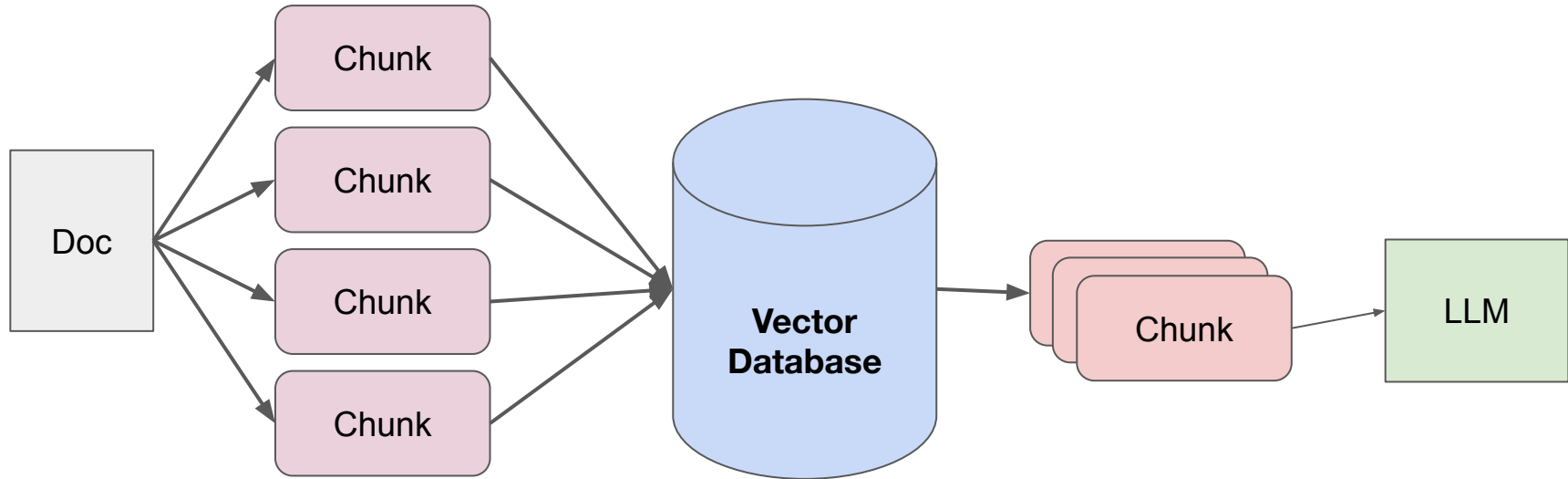- Offers components across the data lifecycle: ingest, index, and query over data

Data Ingestion
(LlamaHub 🦙)

Data Structures

Queries

- Connect your existing data sources and data formats (API's, PDF's, docs, SQL, etc.)

- Store and index your data for different use cases. Integrate with different db's (vector db, graph db, kv db)

- Retrieve and query over data

- Includes: QA, Summarization, Agents, and more

RAG Stack

Current RAG Stack for building a QA System

Data Ingestion / Parsing

Data Querying

Doc

Chunk

Chunk

Chunk

Chunk

Vector Database

Chunk

LLM

5 Lines of Code in LlamaIndex!

# Current RAG Stack (Data Ingestion/Parsing)
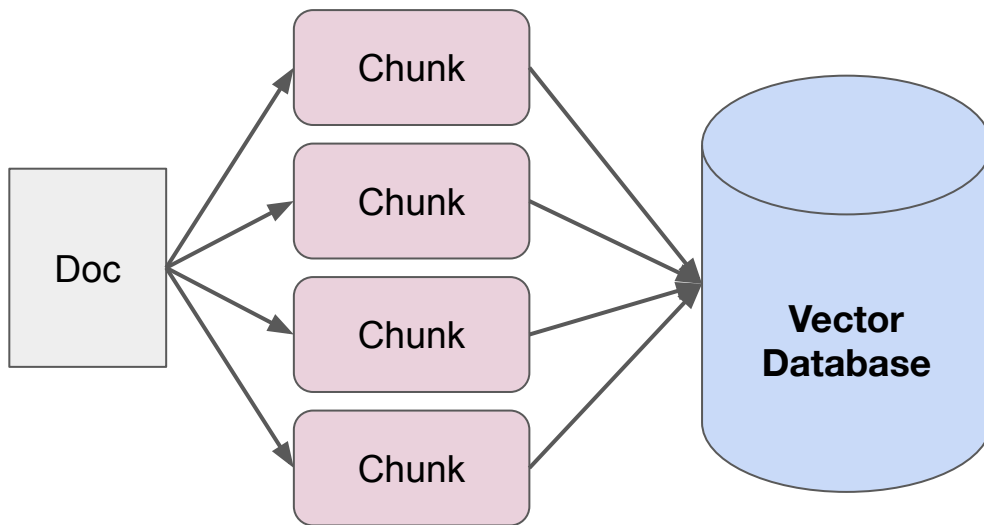
Process:
- Split up document(s) into even chunks.
- Each chunk is a piece of raw text.
- Generate embedding for each chunk (e.g. OpenAI embeddings, sentence_transformer)
- Store each chunk into a vector database
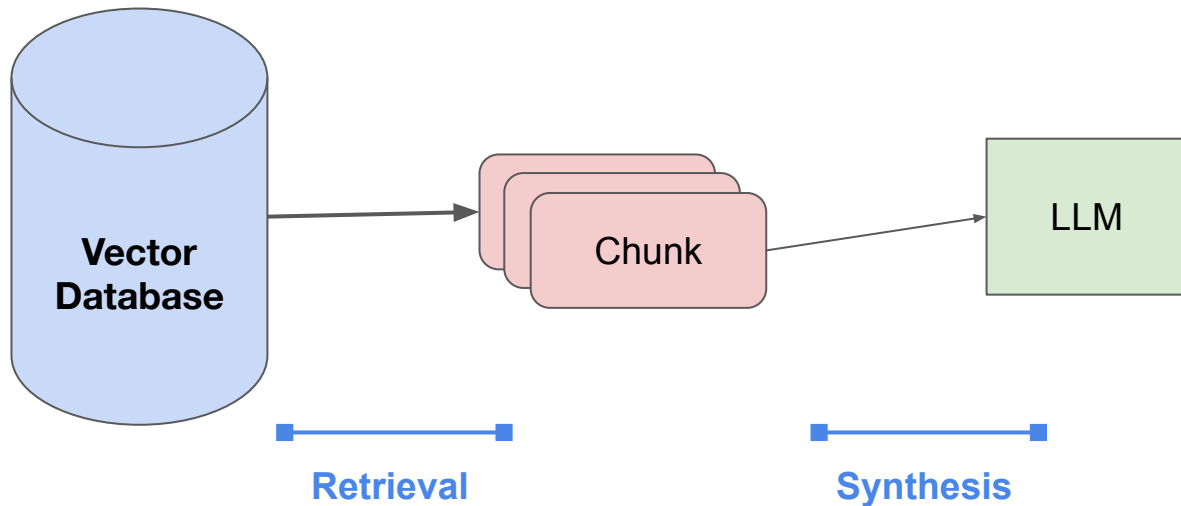
Current RAG Stack (Querying)

Process:
- Find top-k most similar chunks from vector database collection
- Plug into LLM response synthesis module
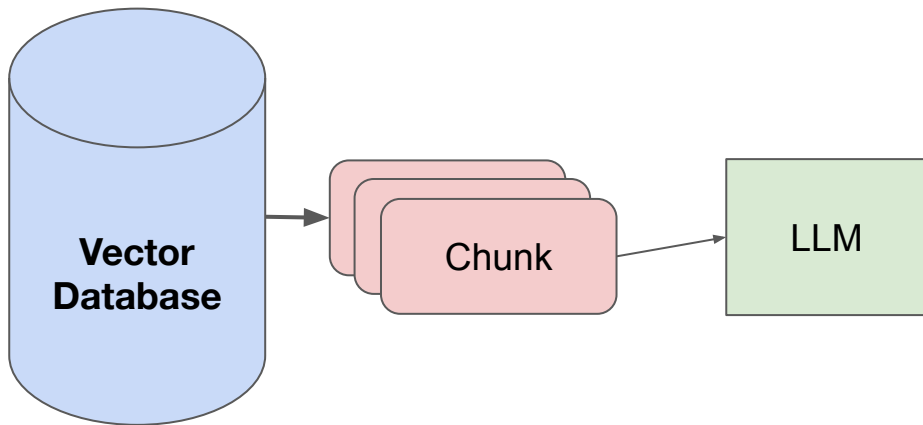
Current RAG Stack (Querying)

Process:
- Find top-k most similar chunks from vector database collection
- Plug into LLM response synthesis module

```python
from llama_index import VectorStoreIndex, SimpleDirectoryReader

documents = SimpleDirectoryReader('data').load_data()
index = VectorStoreIndex.from_documents(documents)
query_engine = index.as_query_engine()
response = query_engine.query("What did the author do growing up?")
print(str(response))
```
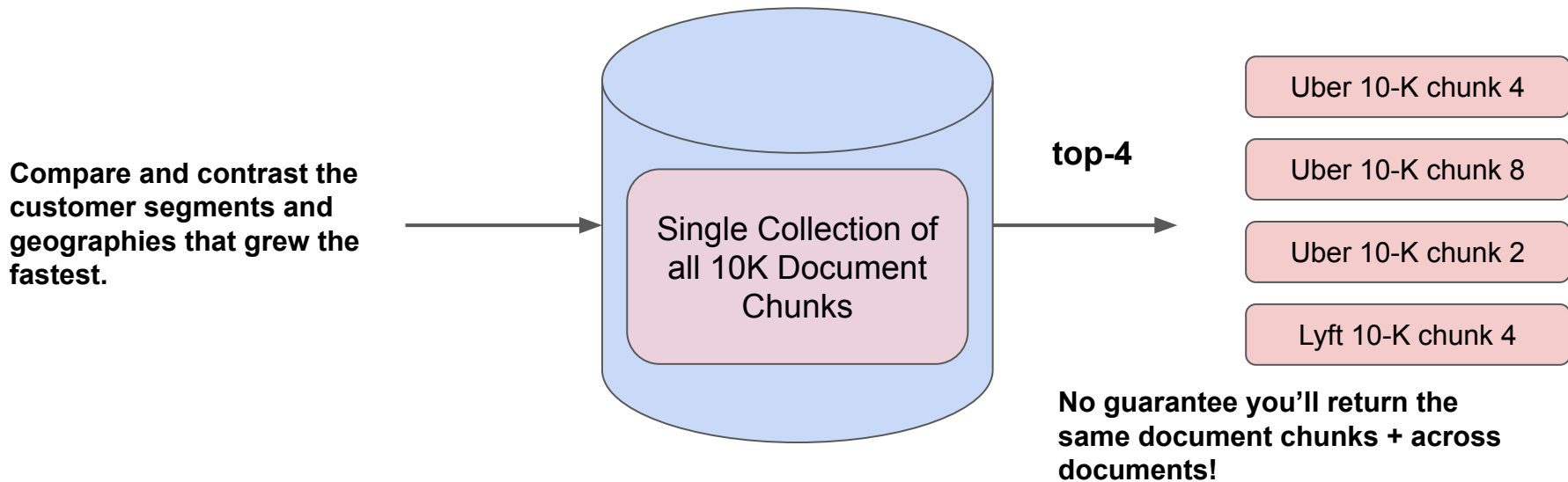
"Chat with Data" Use Cases in Finance

- **Use Case: Multi-Document Comparisons**
  - How do we ask complex questions over multiple documents?
  - Application Demo: SEC Insights
- **Use Case: Complex Document**
  - How do we deal with embedded tables within a document?
- **Use Case: Text-to-DSL**
  - OpenBB Demonstration

Use Case: Multi-Document Comparisons

Example: Compare the 2021 10-K filings for Uber and Lyft

Question: "Compare and contrast the customer segments and geographies that grew the fastest."

Doing top-k retrieval over all 10-K chunks doesn't always work.

Compare and contrast the customer segments and geographies that grew the fastest.

Single Collection of all 10K Document Chunks

top-4

Uber 10-K chunk 4

Uber 10-K chunk 8

Uber 10-K chunk 2

Lyft 10-K chunk 4

No guarantee you'll return the same document chunks + across documents!

Use Case: Multi-Document Comparisons

Instead:

- Index documents separately
- Break down question into question per document → combine results!
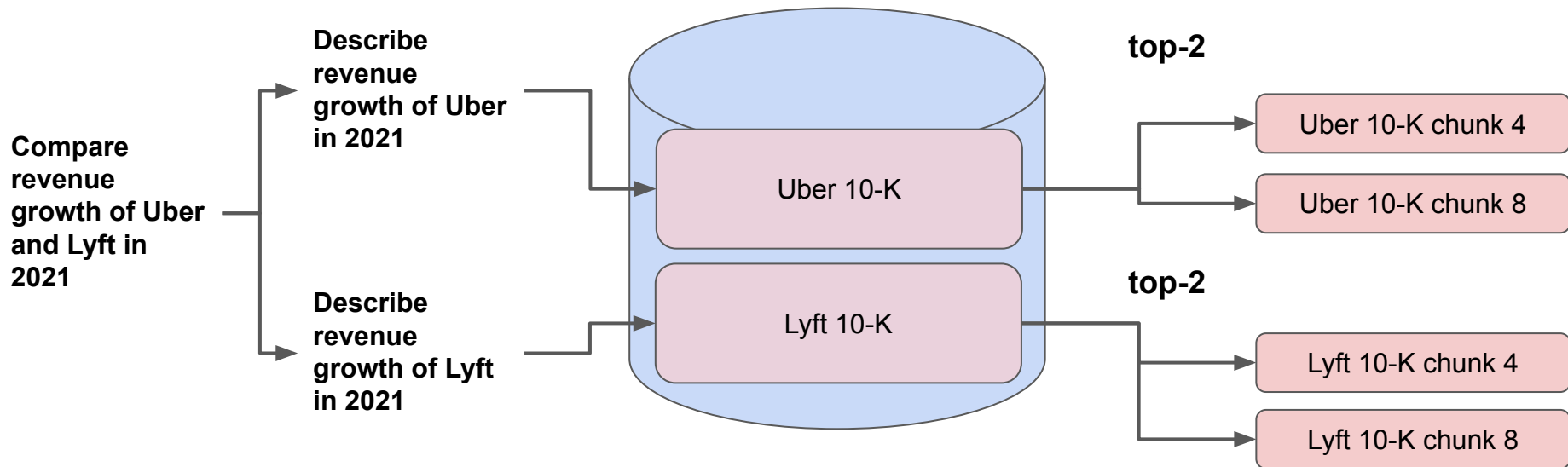
Notebook Walkthrough

https://colab.research.google.com/drive/1Y_IUUKMdC627J5EP0dK1H8NveovpYisM?usp=sharing

Application Demo: SEC Insights

SEC Insights Demo

- https://www.secinsights.ai/
- Conversational AI experience over financial documents built with LlamaIndex

Use Case: Complex Document

How do we model PDFs with embedded tables?

Splitting text into even chunks → leads to hallucinations!

Embedded Table ⟶

Annual rankings

The rankings are published annually in March, so the net worths listed are snapshots taken at that time. These lists only show the top 10 wealthiest billionaires for each year.

Legend

Icon	Description
—	Has not changed from the previous ranking.
▲	Has increased from the previous ranking.
▼	Has decreased from the previous ranking.

2023

In the 37th annual *Forbes* list of the world's billionaires, the list included 2,640 billionaires with a total net wealth of $12.2 trillion, down 28 members and $500 billion from 2022. Over half of the list is poorer than the previous year, including Elon Musk, who fell from No. 1 to No. 2.[2] The list also marks for the first time a French citizen was in the top position as well as a non-American for the first time since 2013 when the Mexican Carlos Slim Helu was the world's richest person. The list, like in 2022, counted 15 under 30 billionaires with the richest of them being Red Bull heir Mark Mateschitz with a net worth of $34.7 billion. The youngest of the lot were Clemente Del Vecchio, heir to the Luxottica fortune shared with his six siblings and stepmother, and Kim Jung-yang, whose fortune lies in Japanese-South Korean gaming giant Nexon, both under-20s.[10]
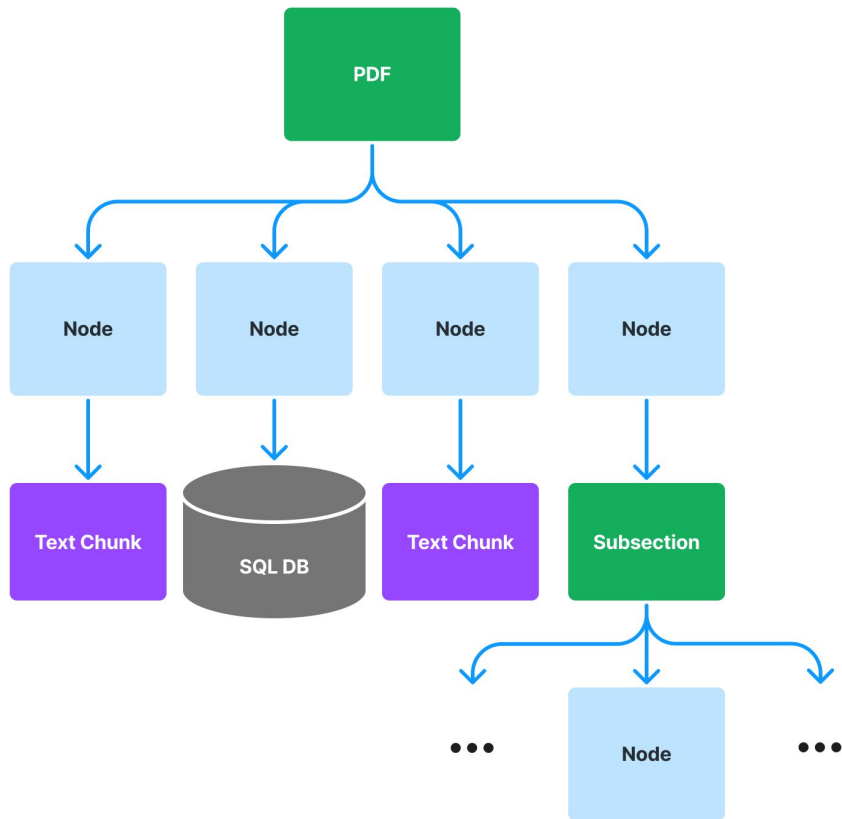
No. ⇅	Name ⇅	Net worth (USD) ⇅	Age ⇅	Nationality ⇅	Primary source(s) of wealth ⇅
1 ▲	Bernard Arnault & family	$211 billion ▲	74	France	LVMH
2 ▼	Elon Musk	$180 billion ▼	51	United States	Tesla, SpaceX
3 ▼	Jeff Bezos	$114 billion ▼	59	United States	Amazon
4 ▲	Larry Ellison	$107 billion ▲	78	United States	Oracle Corporation
5 —	Warren Buffett	$106 billion ▼	92	United States	Berkshire Hathaway
6 ▼	Bill Gates	$104 billion ▼	67	United States	Microsoft
7 ▲	Michael Bloomberg	$94.5 billion ▲	81	United States	Bloomberg L.P.
8 ▲	Carlos Slim & family	$93 billion ▲	83	Mexico	Telmex, América Móvil, Grupo Carso
9 ▲	Mukesh Ambani	$83.4 billion ▼	65	India	Reliance Industries
10 ▼	Steve Ballmer	$80.7 billion ▼	67	United States	Microsoft

Use Case: Complex Document

Instead: model data hierarchically.

Embed a text-to-pandas query engine within a text chunk.

Recursively query the chunks!

Notebook Walkthrough

https://colab.research.google.com/drive/1DldMhszgSI4KKI2UziNHHM4w8Cb5OxEL?usp=sharing

Text-to-DSL

- Instead of vector db to store the raw data, use it to store *instructions* to query your raw data


```
Code    Blame    15 lines (14 loc) · 3.13 KB                                                    Raw

  1   parent_command: stocks/
  2   usage: load [-t TICKER] [-s START] [-e END] [-i {1,5,15,30,60}] [-p]        [-f FILEPATH] [-m] [-w] [--perf
  3   Summary: Load stock ticker data for analysis, with options to specify the date range, interval, data source, and
  4
  5   Examples:
  6   - Load stock data for <SYMBOL> with default settings: stocks/load -t <SYMBOL>
  7   - Load stock data for <SYMBOL> with a custom date range: stocks/load -t <SYMBOL> -s 2022-01-01 -e 2022-12-31
  8   - Load stock data for <SYMBOL> with 15-minute intervals: stocks/load -t <SYMBOL> -i 15
  9   - Load stock data for <SYMBOL> including pre/after market hours: stocks/load -t <SYMBOL> -p
 10   - Load stock data for <SYMBOL> from a custom file: stocks/load -t <SYMBOL> -f FILEPATH
 11   - Load monthly data for <SYMBOL>: stocks/load -t <SYMBOL> -m
 12   - Load weekly data for <SYMBOL>: stocks/load -t <SYMBOL> -w
 13   - Load stock data for <SYMBOL> and show performance information: stocks/load -t <SYMBOL> --performance
 14   - Load stock data for <SYMBOL> and export in json format: stocks/load -t <SYMBOL> --export json
 15   - Load stock data for <SYMBOL> using a specific data source: stocks/load -t <SYMBOL> --source AlphaVantage
```

Text-to-DSL

- **Case Study:** OpenBB
- Translate natural language query to an OpenBB command

```
prompt_string = f"""From the cli argparse help text above, provide the terminal
    command for {query_text}. If relevant, use the examples as guidance.
    Provide the exact command along with the parent command with a "/" separation to get that information,
    and nothing else including any explanation. Don't add any other word such as 'Command to get', 'Answer'
    or the likes.  If you do not know, reply "I don't know"

    Current date: {current_date.strftime("%Y-%m-%d")}
    Current day of the week: {current_date.strftime("%A")}

    Remember:
    1. It is very important to provide the full path of the command including the parent command and loading
    the particular target before running any subsequent commands
    2. If you are asked about dates or times, load the target dates, times span during the "load" command
    before running any subsequent commands. replace all <> with the actual dates and times.  The date format should
    be YYYY-MM-DD. If there is no date included in the query, do not specify any.
    3. Country names should be snake case and lower case. example: united_states.
    4. Always use a comma to separate between countries and no spaces: example: united_states,italy,spain
    5. Always use "load" command first before running any subsequent commands. example:
    stocks/load <symbol>/ ....
    crypto/load <symbol>/ .... etc.
    6. Do not include --export unless the request asks for the data to be exported or saved to a specific file type.
    7. Do not make up any subcommands or options for the specific command.
    8. Do not provide anything that could be interpreted as investment advice.
    9. Any request asking for options refers to stocks/options.

    Only do what is asked and only provide a single command string, never more than one.
    """
```

Text-to-DSL

- **Case Study:** OpenBB
- Translate natural language query to an OpenBB command

LlamaIndex + MindsDB

We have an integration with MindsDB!

Easily integrate LlamaIndex with the MindsDB SQL interface

```
15
16   -- CREATE MODEL my_qa_model2
17   -- PREDICT answer
18   -- USING
19   --    engine = 'llama_index',
20   --    index_class = 'GPTVectorStoreIndex',
21   --    query_engine = 'as_query_engine',
22   --    reader = 'SimpleWebPageReader',
23   --    source_url_link = 'https://mindsdb.com/about',
24   --    input_column = 'question',
25   --    openai_api_key = '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'
26
27   SELECT question,answer
28   FROM mindsdb.my_qa_model2
29   WHERE question = "What problem does MindsDB solves?";
```

                                                              ▪▪ Data Insights   Export ⌄

question	answer
What problem does MindsDB solves?	MindsDB solves the problem of quickly and easily integrating machine learning into applications. It provides a platform to connect any data s place.

```
27
28  SELECT t.question, m.answer
29  FROM mindsdb.my_qa_model2 as m
30  JOIN files.question_table as t;
```

Data Insights   Export ⌄

question	answer
Who founded MindsDB?	Adam Carrigan and Jorge Torres founded MindsDB in 2017.
What problem does MindsDB solve?	MindsDB solves the problem of deploying machine learning models quickly and easily. It provides an easy-to-use platform that allows users to i databases, apply OpenAI's GPT-3 or Hugging Face to rich text in their databases, and access more than 100 data sources. It also provides resou hub, a hackathon, Slack channels, and more.
Where MindsDB team is located?	The MindsDB team is located in the US, UK, and other international locations.

Thanks!

Building LLMs in an enterprise setting? We'd love to chat!