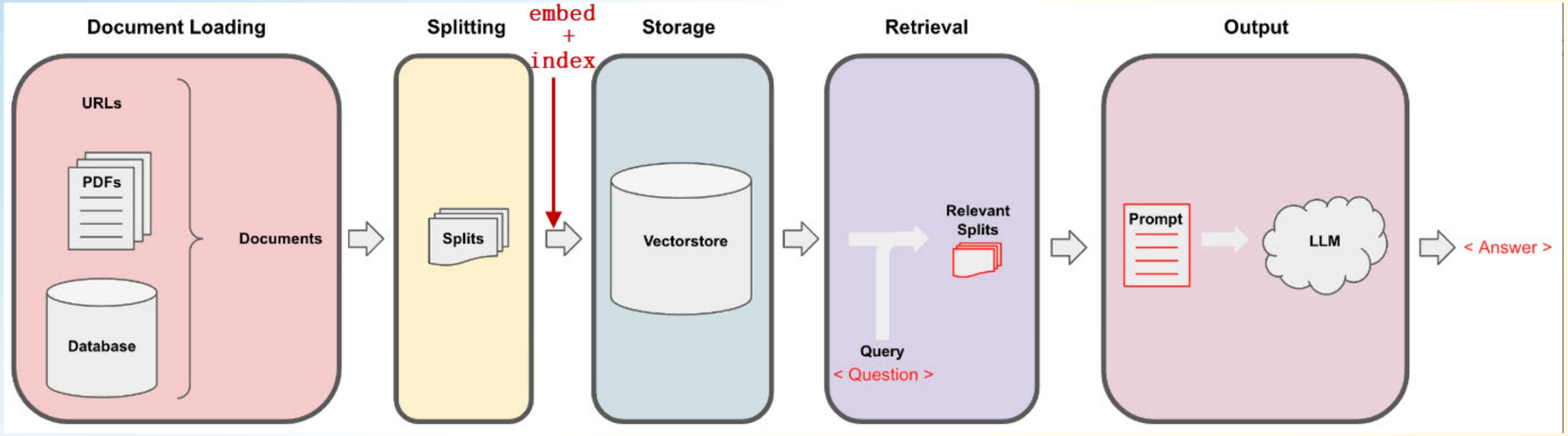
Advisor: Dr. Chang, Henry

ANSWER GENERATING USING COHERE

Presented by: Shin Cao

San Francisco Bay University

HOW DOES RAG WORK?



PROCESS EXPLAINED
STEP 1: PREPARE TEXTS

```
[1]: question = "Are side projects important when you are starting to learn about AI?"

[2]: text = """
The rapid rise of AI has led to a rapid rise in AI jobs, and many people are building excit

Three key steps of career growth are learning (to gain technical and other skills), working

Initially, you focus on gaining foundational technical skills.
After having gained foundational skills, you lean into project work. During this period, yo
Later, you might occasionally carry out a job search. Throughout this process, you'll proba
These phases apply in a wide range of professions, but AI involves unique elements. For exa

AI is nascent, and many technologies are still evolving. While the foundations of machine l
Project work often means working with stakeholders who lack expertise in AI. This can make
While searching for a job in AI can be similar to searching for a job in other sectors, the
Throughout these steps, a supportive community is a big help. Having a group of friends and

I'm excited to work with all of you to grow the global AI community, and that includes help
```

PROCESS EXPLAINED
STEP 2: CHUNKING

```
# Split into a list of paragraphs
texts = text.split('\n\n')

# Clean up to remove empty spaces and new lines
texts = np.array([t.strip(' \n') for t in texts if t])
```

```
texts[:3]
```

PROCESS EXPLAINED
STEP 3: EMBEDDING

[7]:
```
co = cohere.Client(os.environ['COHERE_API_KEY'])

# Get the embeddings
response = co.embed(
    texts=texts.tolist(),
).embeddings
```

PROCESS EXPLAINED
STEP 4: BUILD A SEARCH IND

[8]:
```
# !pip install annoy
```

[9]:
```
from annoy import AnnoyIndex
import numpy as np
import pandas as pd
```

10]:
```
# Check the dimensions of the embeddings
embeds = np.array(response)

# Create the search index, pass the size of embedding
search_index = AnnoyIndex(embeds.shape[1], 'angular')
# Add all the vectors to the search index
for i in range(len(embeds)):
    search_index.add_item(i, embeds[i])

search_index.build(10) # 10 trees
search_index.save('test.ann')
```

10]:
```
True
```

INSTALLATION & SETUP

View README on my GitHub:
https://github.com/Shining-in-galaxies/genrative_search_cohere

Thanks For Watching

Shin Cao

shin.ccx@outlook.com

https://github.com/Shining-in-galaxies

San Francisco Bay University