ParliaGPT

An FAQ Bot for Parliamentary Minutes

Date: 28 Apr ’24

Yong Ting Rui | E1339749

Introduction

- ParliaGPT is an FAQ Bot for **general members of the public** that have a keen interest in parliamentary debates

- Objective Statement: How can the process of finding information from parliamentary debate documents be streamlined, so that users can quickly obtain a response, and more information for future deep dive?

Project Background

- Current approach of obtaining information from parliamentary debates are:
  - Watching the recorded parliamentary debate meeting video
  - Reading the parliamentary meeting minutes
- For either approach, the process to obtain a snippet of information from the entire process is **tedious** and **time-consuming**

Market Research

- Existing Competitors
  - ChatGPT-
    - Strengths: Allow users to make **natural language queries**, and provide **natural language responses** back to its' users
    - Weakness: Unable to provide information on events/data it has not been trained on
  - Search Engines (Google, Duckduckgo, Bing, etc)-
    - Strengths: Allow users to **search and obtain extensive results** from the public domain
    - Weakness: Users have to deep dive into search results for more information, and information may come alternative sources (ie: news articles) instead from the root source (ie: parliamentary debate documents)

Project Scope

- ParliaGPT shall be built on two main features:
  - A **Retriever** that retrieves relevant parliamentary debate documents and pages
  - A **Text Generator** that provides a summarized response from the relevant parliamentary debate pages retrieved

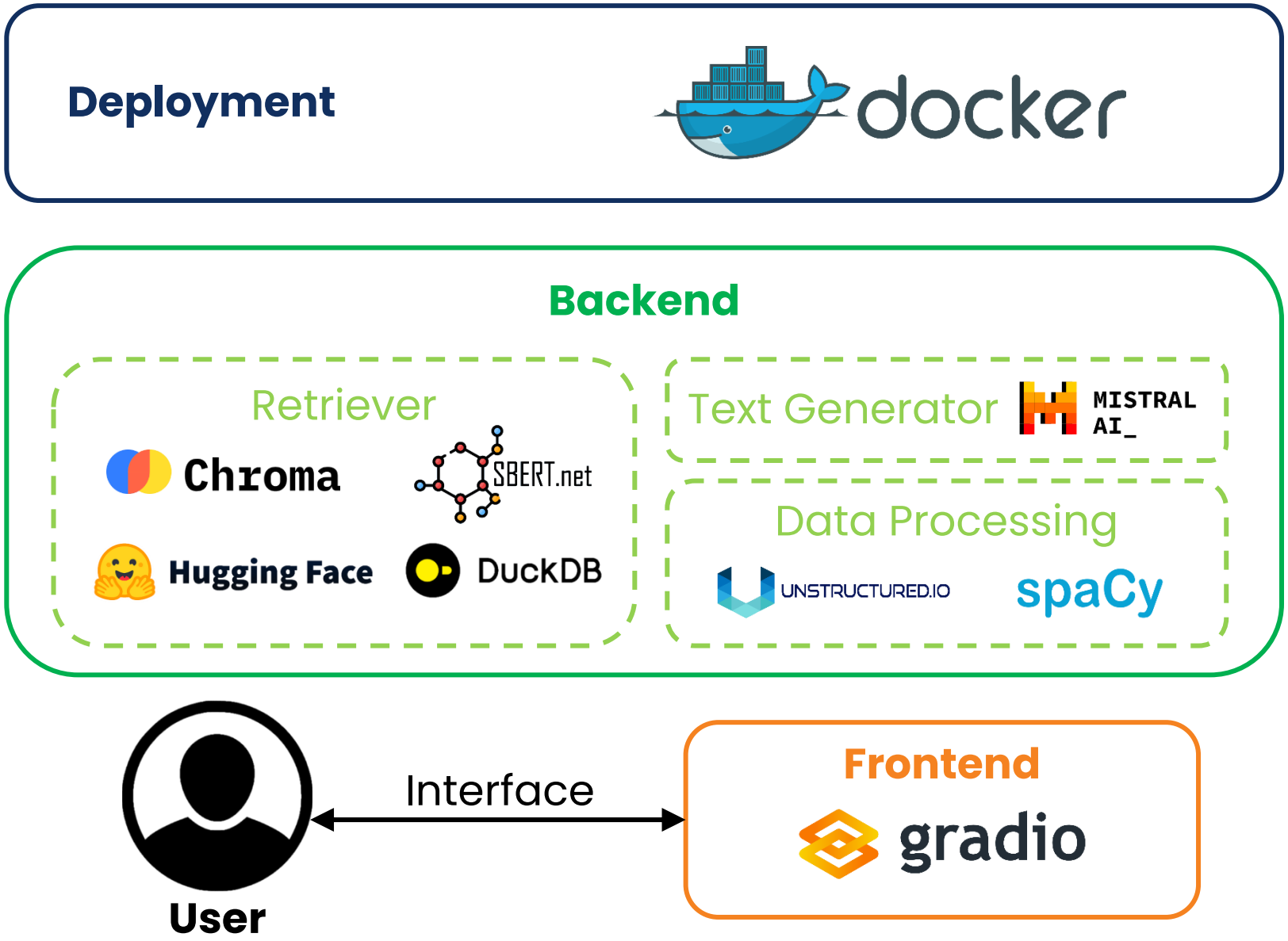
- ParliaGPT will be deployed as a frontend user web-interface without the need for account setup or creation

Data Collection and Preparation

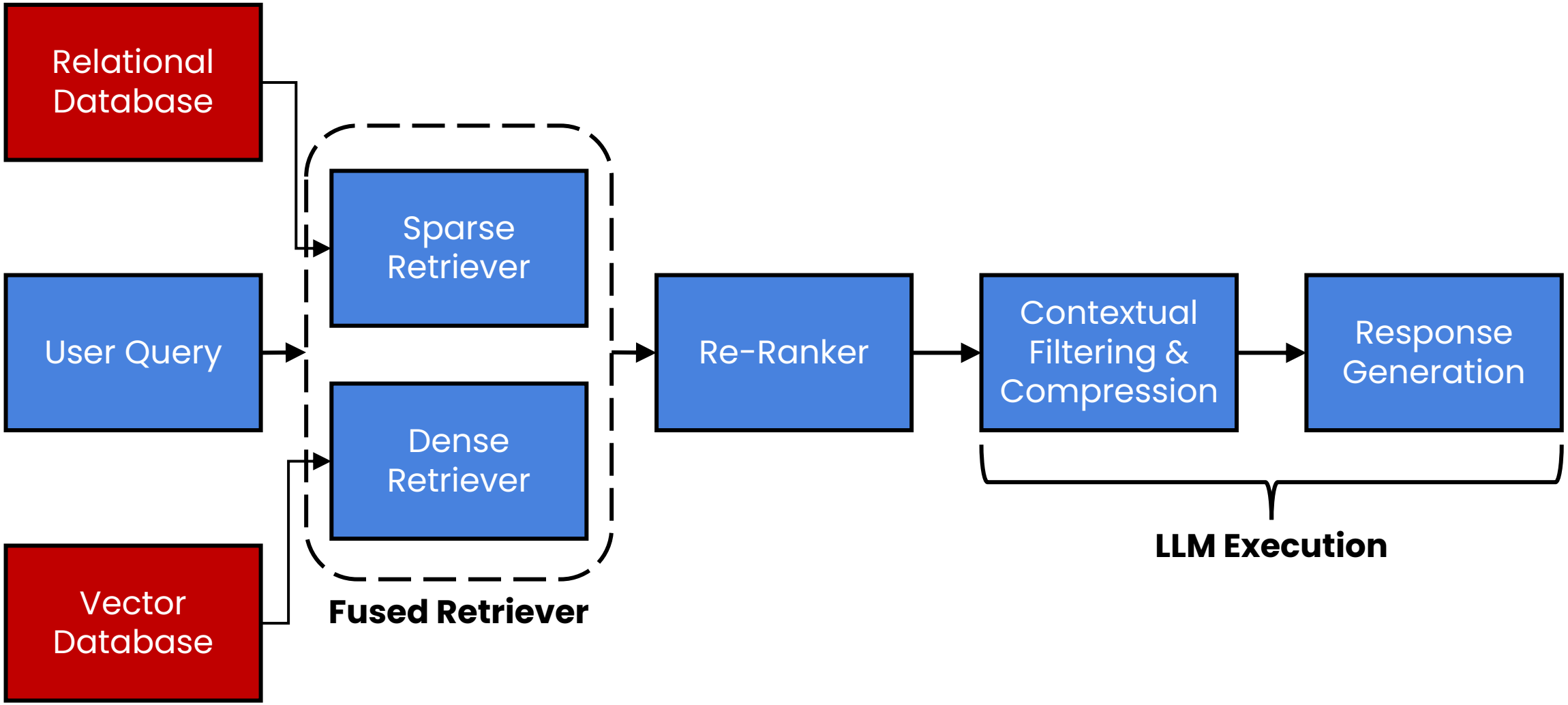
Data Sources:

- Official Report on parliamentary debates can be obtained from https://www.parliament.gov.sg/

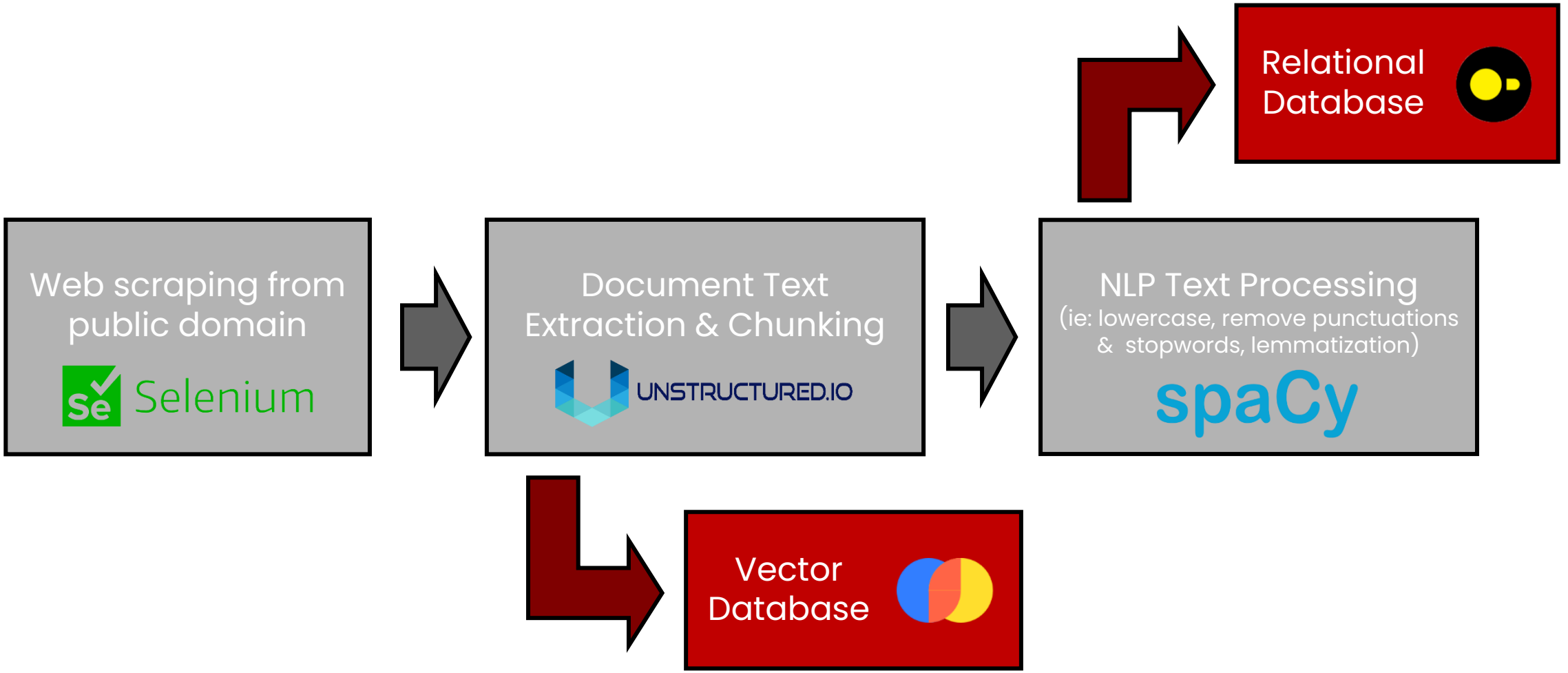
System Overview



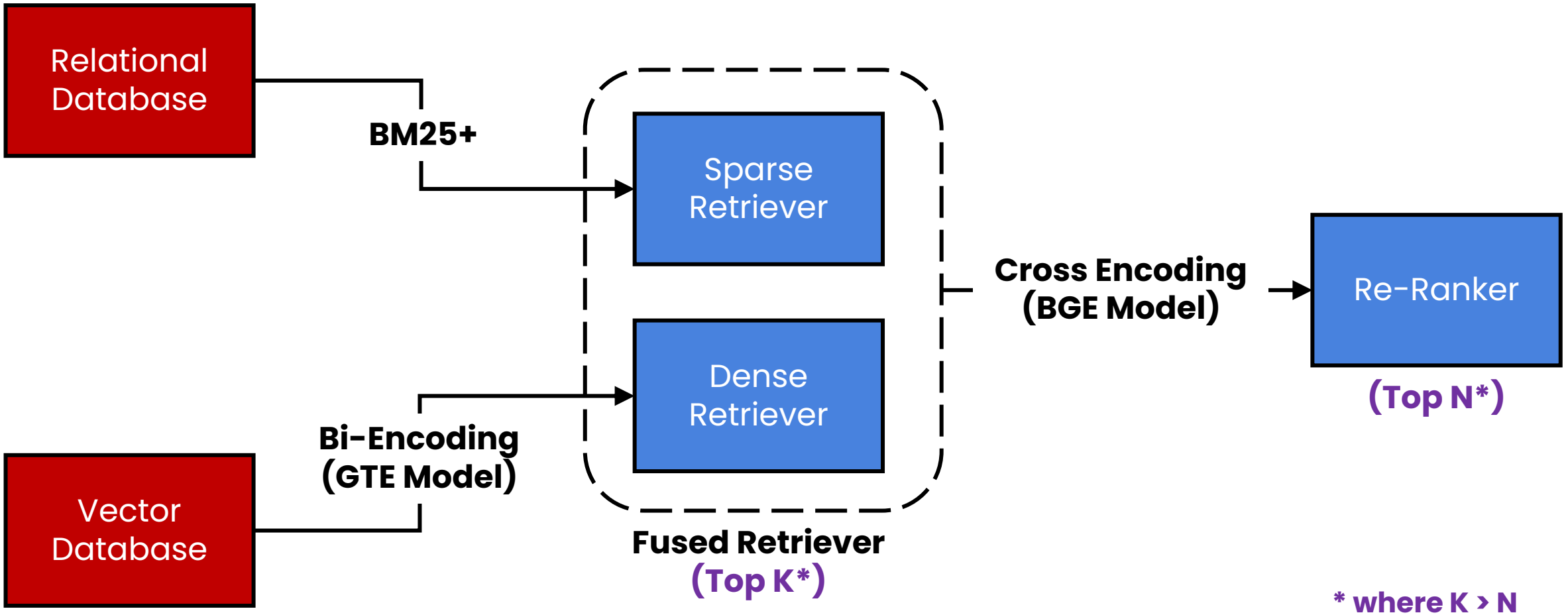
System Design – RAG Architecture



Implementation – Retrieval Databases

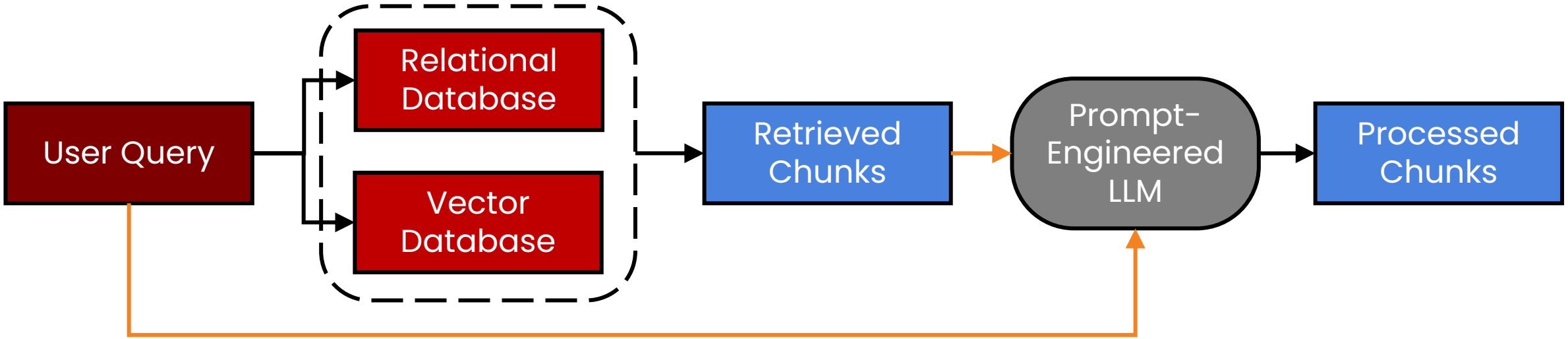


Implementation – Fused Retriever + Re-Ranker



Implementation – Contextual Filtering & Compression

- Retrieved chunks may contain information spanning multiple topics, including sub-chunks that are **not relevant** to users' query
- Retrieved chunks are processed by extracting sub-chunks relevant to the query through **prompt engineering**, before feeding into LLM for response generation



Future Work

- Improve text generator output by incorporating a **sub-query engine** that splits complex queries into simpler queries for processing
- Incorporate a **voting mechanism** that allows users to provide that feedback on the quality of responses provided by the FAQ Bot
- Create a more polished **user interface** with additional features (ie: Accept voice inputs, Provide top N queries made)

Conclusion

- In summary, ParliaGPT is an FAQ Bot that **transforms** the way that the general public, students and academics can gain access to parliamentary debate documents

The End

Thank You! :)