**Problem Statement** - Build a robust generative search system capable of effectively and accurately answering questions from a list of HDFC Insurance policy documents.

**Solution Strategy** - The solution should address the following requirements:

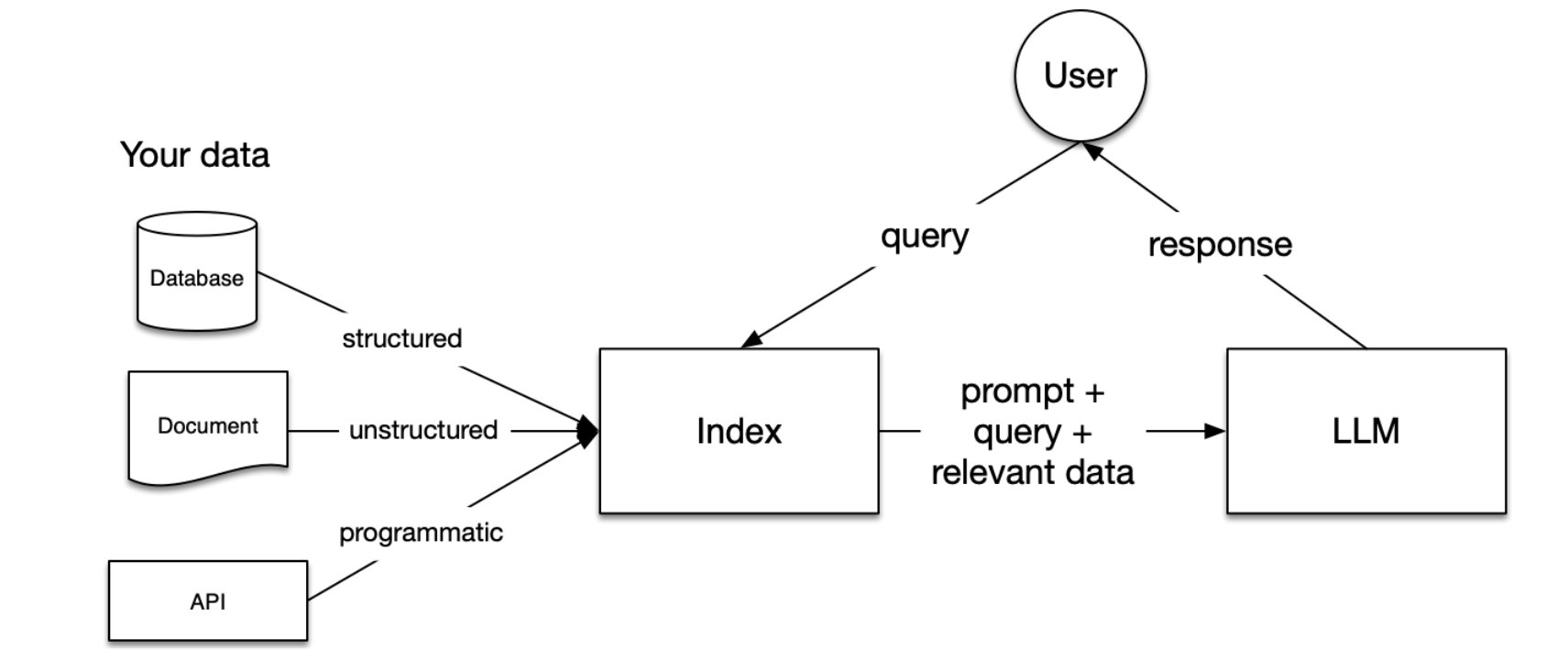
- Users should receive accurate responses from the insurance policy knowledge base.

- When users perform queries, the system must be able to respond accurately and effectively.

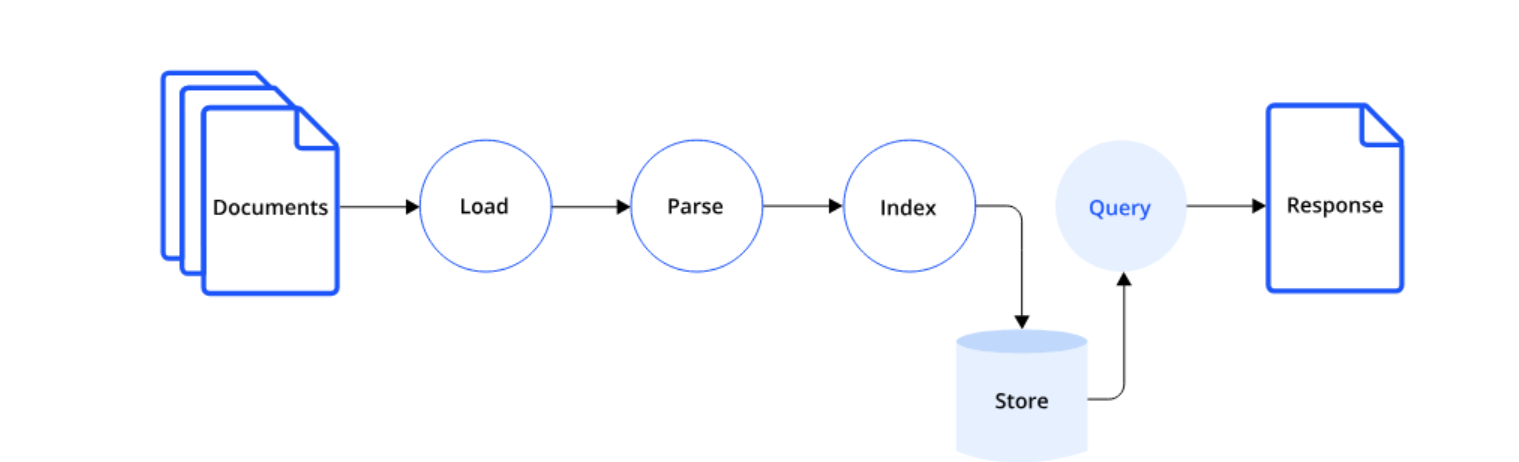
**Goal** - Successfully meeting the above requirements will ensure that the overall model performs well and delivers high accuracy.

**Data Used** - HDFC Insurance policy documents stored in a single folder.

**Tools Used** - LlamaIndex and ChatGPT are utilized due to their powerful query engines, efficient data processing using data loaders and directory readers, and their ease of implementation with minimal lines of code.



**Benefits of Llama Index:**



Llama Index is a cutting-edge data framework designed to streamline the development of **Retrieval-Augmented Generation (RAG)** applications using large language models (LLMs). It provides a robust platform that enables developers to integrate various data sources seamlessly with

**Versatile Data Connectors**: LlamaIndex offers robust connectors that allow seamless ingestion of data from a wide range of sources and formats.

**Data Synthesis**: It can efficiently synthesize information from multiple documents or heterogeneous data sources, enabling comprehensive insights.

**Extensive Integrations**: LlamaIndex supports numerous integrations, including vector stores, ChatGPT plugins, tracing tools, LangChain, and more, enhancing its versatility and functionality.

**Conclusion:**

The development of the HDFC insurance policy Q/A bot (Semantic spotter) was a successful, integrating multiple advanced AI techniques, including Retrieval-Augmented Generation (RAG), prompt customization and Llama index and caching mechanisms. The process offered several key insights into building intelligent systems capable of extracting and presenting information from large, unstructured document sets.

**Steps to execute:**

* Update API Key using path = '/content/drive/MyDrive/GENAI/SemanticSearch/OpenAI\_API\_Key.txt
* Unzip and Read Docs from path /content/drive/MyDrive/GENAI/SemanticSearch/PolicyDocuments



* Once execution started, Insert question “What is the cornea transplantation cost in Surgicare plan?” when initialize\_conv() executes. After ans appears press exit to continue with the run
* Insert Good/Bad feedback when prompted by testing\_pipeline()