Introducing "InSight"

"Insight" is an innovative app that revolutionizes the research process by leveraging cutting-edge technologies like LLM, RAG, Weaviate, and Langchain. It simplifies knowledge discovery and empowers researchers.

Features of InSight

LLM: Language Understanding

Utilize advanced language understanding capabilities to extract valuable insights and connections from your research.

RAG: QnA and Topic Modeling

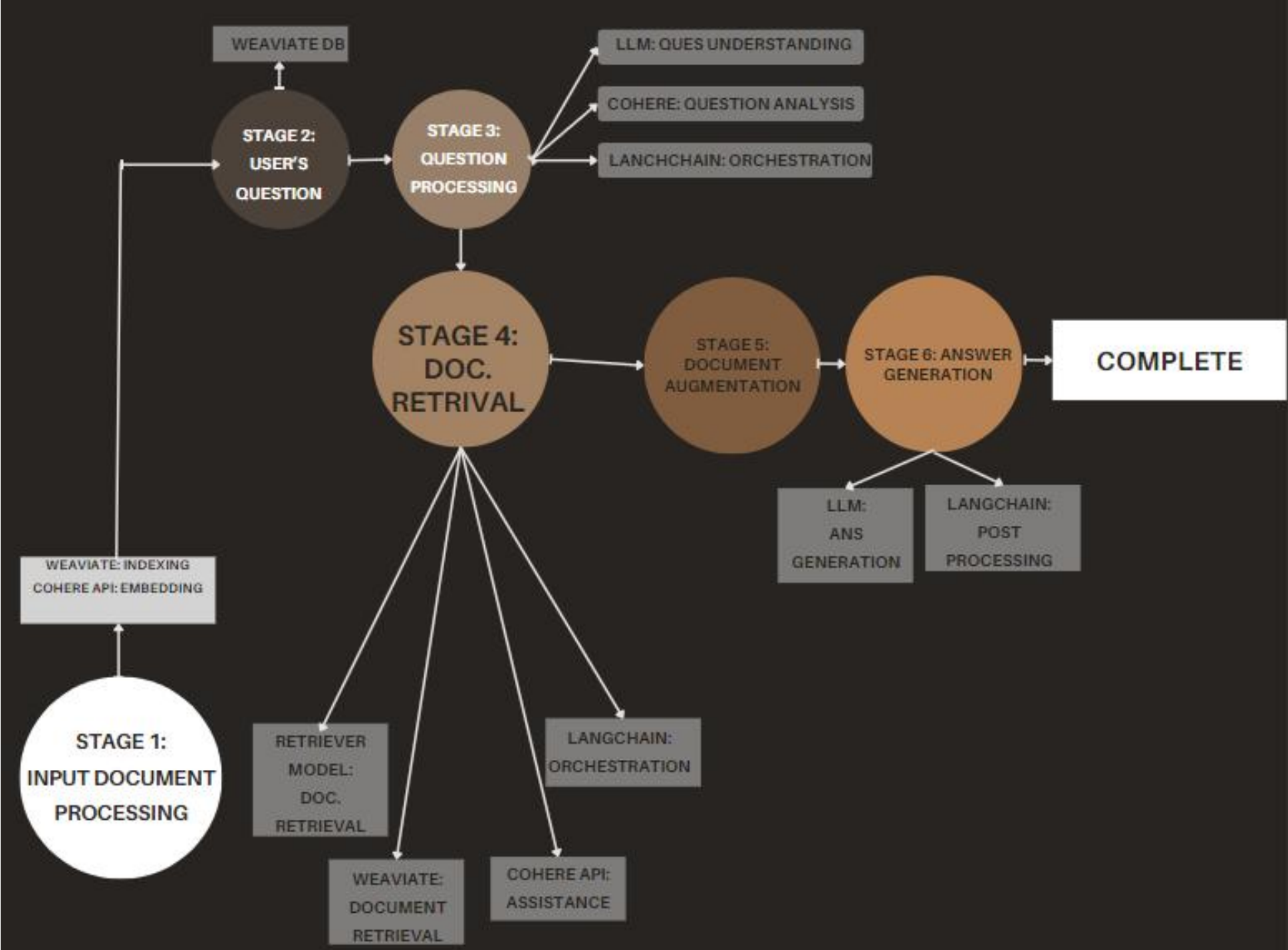Get quick answers to your research questions and explore relevant topics through AI-powered QnA and topic modeling.

Summarization Capabilities

Easily summarize extensive research material into concise and digestible snippets for efficient analysis.

Weaviate Knowledge Graph

Construct a robust knowledge graph using Weaviate, enabling seamless navigation and visualization of research connections.

Functional Diagram:

How it works :

Research Insight's Document-Based Question-Answering system is powered by Cohere's RAG (Retrieval-Augmented Generation) model. This model allows users to input questions and receive coherent, contextually accurate answers by analyzing connected documents. "Research Insight" integrates Cohere's powerful technology, which includes advanced text embedding capabilities.
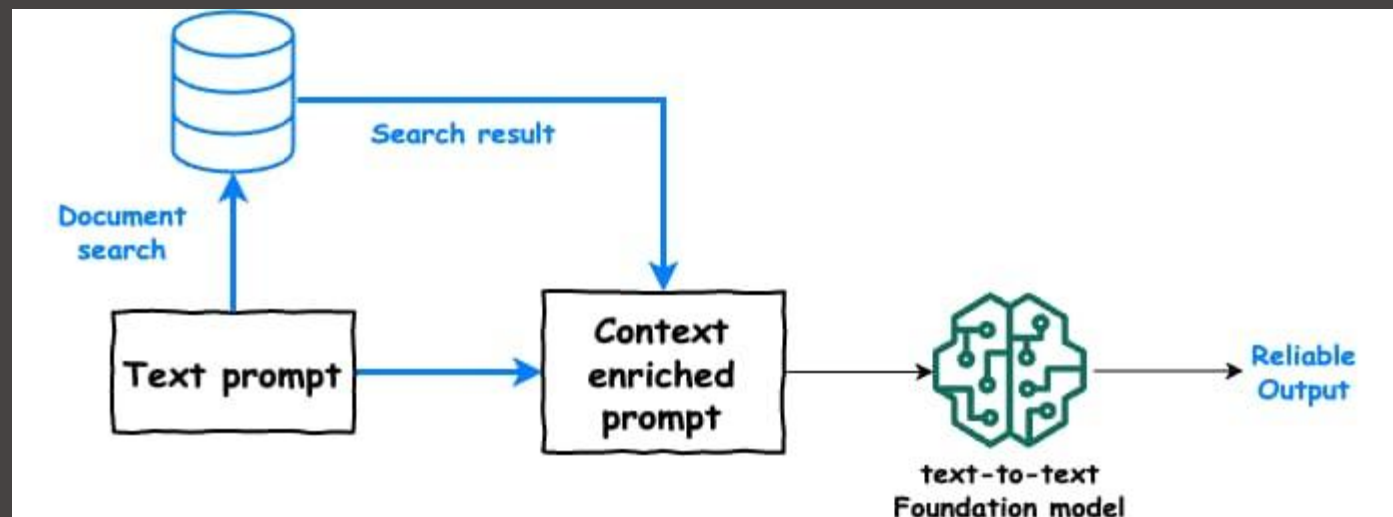
- This integration allows us to retrieve meaningful text embeddings that form the basis of our topic modeling.

- Here, we are using out LLM stack containing langchain, weaviate, cohere api etc to retrieve suitable answers based on the user's prompt

- we are using RAG model to retrieve precise answers for our questions through the input documents

Utilizing Langchain and Weaviate Database

- We rely on Langchain to maintain the quality and consistency of research data throughout its lifecycle.

- This framework is vital for creating a robust and reliable environment for our users.

- Weaviate serves as the knowledge repository for "InSight."

- It stores and organizes research data in a way that facilitates efficient search and discovery.

- Weaviate enables the seamless exploration of research topics and themes, providing a knowledge graph that connects and structures information.

Retrieval-Augmented Generation (RAG)

Retrieval-Augmented Generation (RAG) is a powerful technique that combines information retrieval and language generation to generate relevant and coherent text.

Role of RAG in Our Use Case

Large Language Models make semantic search—a viable option. This capability is powered by text embeddings, which takes a document and represents its contextual information by a list of numbers, also called vectors.

This Document Question Answering application takes a user's documents and generates an answer by searching the most semantically similar sections of the document using the Cohere chat and embed endpoint.

The Embed endpoint is responsible for converting the document chunks into contextual embeddings which are further saved into a Weaviate cluster.

Benefits of Research Insight

1 Simplifies Knowledge Discovery

Streamline the process of knowledge discovery by efficiently organizing, analyzing, and connecting research information.

2 Revolutionizes Research Process

Empower researchers with cutting-edge technologies to enhance productivity, accuracy, and innovation in their work.

3 Invaluable Tool for Researchers

Transform the way researchers approach their work, providing a valuable tool that saves time and enhances the quality of their research.

Project Link: https://insightbot.streamlit.app/