Code Interpreter Powered RAG

The program is designed to enhance RAG's capabilities by connecting to the TogetherAI API for the Large Language Model (LLM) and using LlamaIndex Cohere embeddings and Cohere rerank to locate and process documents. It can retrieve relevant documentation locally and generate a program sequence for execution. When this sequence is run, it produces output text that the system uses for its responses.

A key feature of the program is its integration of a full code interpreter. This interpreter significantly improves RAG's ability to process information accurately, especially for numeric computations. It also allows RAG to execute user-requested functions, making it more versatile and useful.

By leveraging the TogetherAI API and Cohere technology, the program enhances RAG's functionality in providing detailed and accurate information across various topics. The addition of the code interpreter expands RAG's capabilities, making it a more powerful tool for users seeking both information and computational assistance.