**Helpmate-AI-project**

**Problem Statement:**

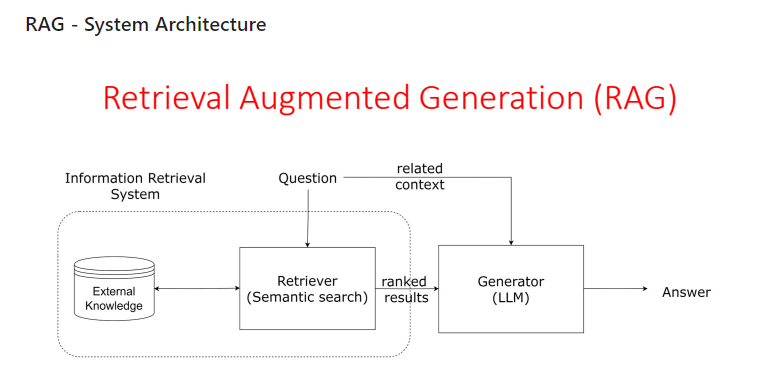
The goal of the project will be to build a robust generative search system capable of effectively and accurately answering questions from various insurance policy documents.

**Solution Strategy:**

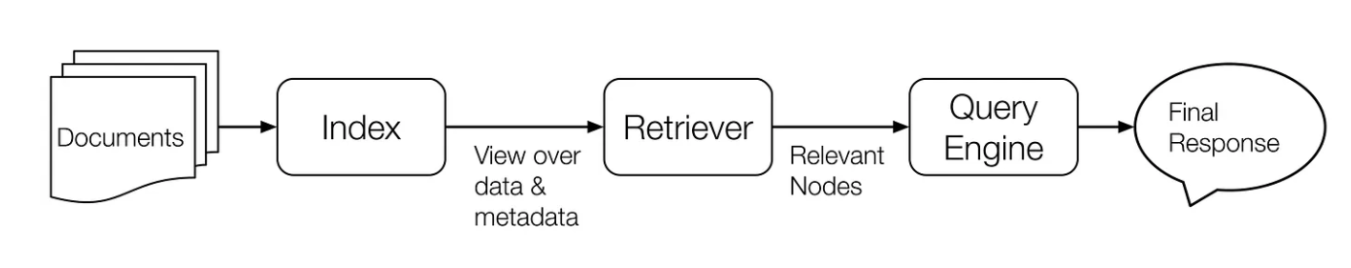
* Users would get responses from the insurance documents.
* Use LlamaIndex to build the generative search application.
* If they want to refer to the original page from which the bot is responding, the bot should provide a citation as well. Goal - Solving the above two requirements well in the POC would ensure that the accuracy of the overall model is good and therefore further improvisations and customizations make sense.

**Data Used:**

Dataset is a set of Insurance Documents provided by Upgrad

[](https://private-user-images.githubusercontent.com/43923258/366770300-cd01793b-6fd9-4c79-9fb7-db4117319bc2.png?jwt=eyJhbGciOiJIUzI1NiIsInR5cCI6IkpXVCJ9..vLatUKby_xgpmy-GS-CY-nAjbf87hGaOxJa-cVwS3g0)

**Tools used:**

LlamaIndex (only for now) has been used due to its powerful query engine, fast data processing using data loaders and directory readers as well as easier and faster implementation using fewer lines of code. [](https://private-user-images.githubusercontent.com/43923258/366764242-37e626b2-4b6f-4b3e-bfd7-b283ceea52d6.png?jwt=eyJhbGciOiJIUzI1NiIsInR5cCI6IkpXVCJ9..-ZQG5d9S0519I9dEAQC_heOBorwBbpyQF5MEKg4AIos)

Documents: These are the "books" in your library.

Index: It's the "library" of your data - Stores your data.

Retriever: It's the "librarian" that finds relevant data - Finds data.

Response Synthesizer: It's the "storyteller" that creates a response - Makes responses.

QueryEngine: It's the "director" that makes everything work together - Coordinates everything.