# Assignment – 3: Retrieval-Augmented Generation (RAG) using LangChain Task 2: RAG Pipeline Diagram and Explanation

Below is a detailed explanation and diagram of the Retrieval-Augmented Generation (RAG) system pipeline:

1. 1. User Query:

A natural language question asked by the user, e.g., 'What are the objectives of this course?'

1. 2. Retriever:

Searches a vector store to find the most relevant document chunks for the user's query.

1. 3. Vector Store:

Stores pre-processed document embeddings. It returns the most similar chunks based on query embeddings.

1. 4. Relevant Chunks:

These are the most relevant parts of documents retrieved from the vector store.

1. 5. LLM (Large Language Model):

Takes the user query and retrieved chunks to generate a context-aware answer.

1. 6. Final Answer:

The model provides the final answer using both internal knowledge and retrieved content.

RAG System Architecture Diagram:

