

# Recap: Custom LLM Testing and Evaluation

## 1 Understand Custom LLM Architecture

Understand how Custom LLM's are being built on RAG Architecture

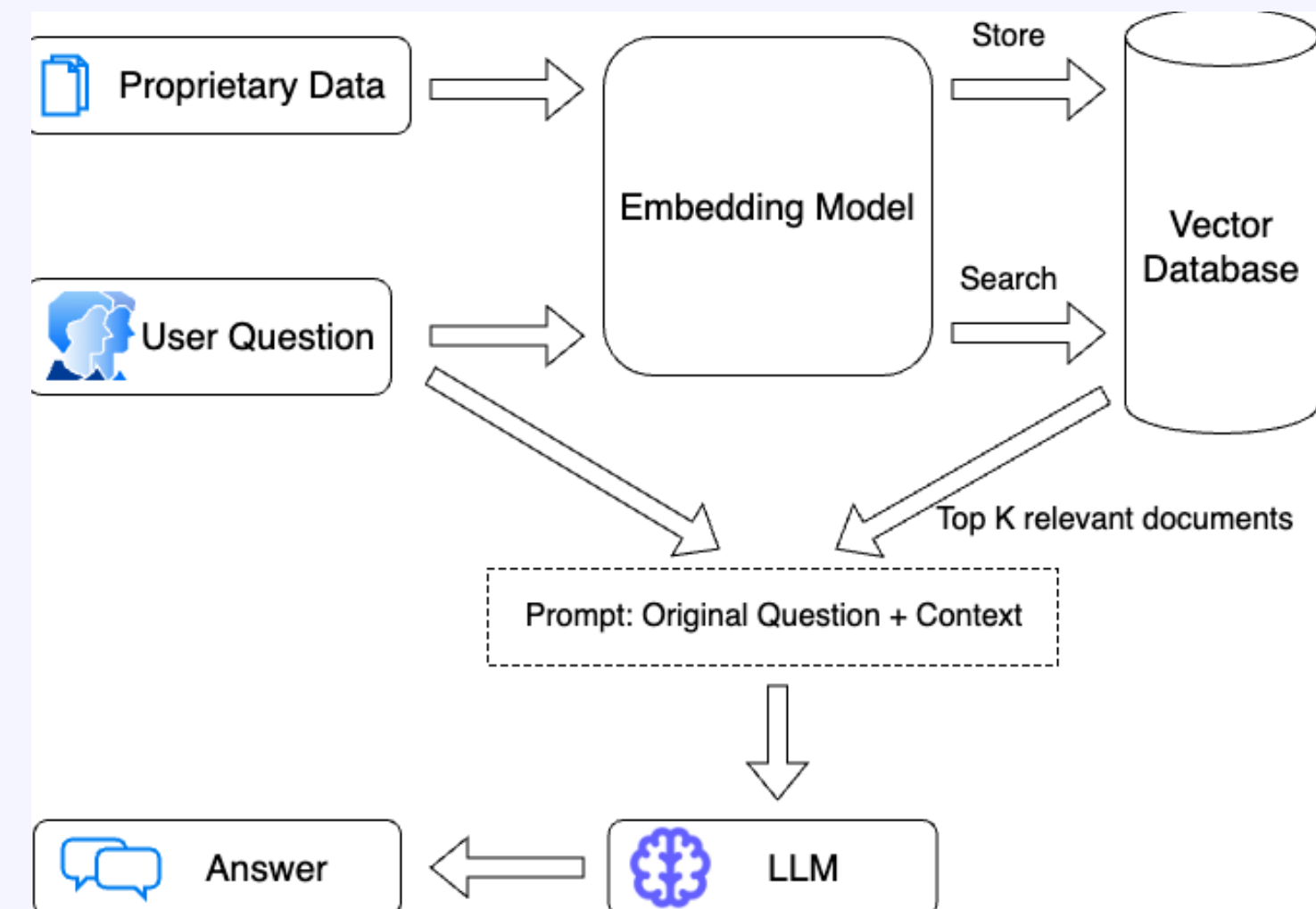
- Context Precision
- Context Recall
- Faithfulness
- Factual Correctness
- Response Relevancy
- Topic Adherence
- Rubrics Score

## 2 Testing Scope

Testing Scope at Retrieval, Augmentation and Generation Modules

## 3 LLM Evaluation Metrics

Scripts to assert LLM Evaluation Metrics:



## 1 Multi Conversational Scenarios

Testing LLM Multi Conversational Scenarios:

## 2 Multiple Metrics Evaluation

Evaluating Multiple metrics using EvalDataSet

## 3 Synthetic Test Data

Generating Synthetic Test Data to evaluate LLM's

## 4 Test Optimization

Optimizing Tests into Pytest Framework Standards.