

LangChain | ChatGroq | FAISS | Hugging Face

Q&A CHATBOT | RAG

LLAMA 3

NAME : KARTAVYA MASTER

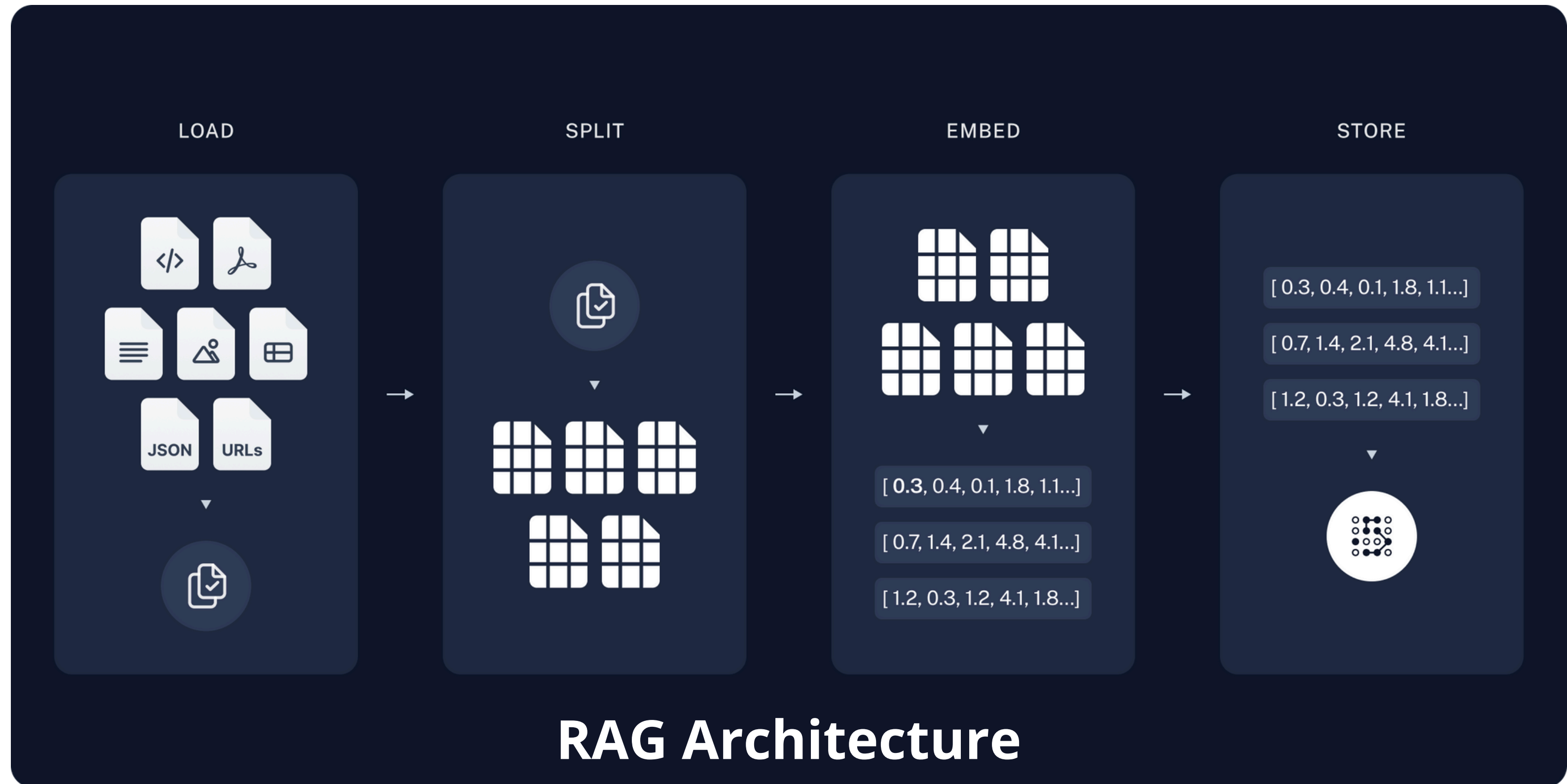


Technologies Used

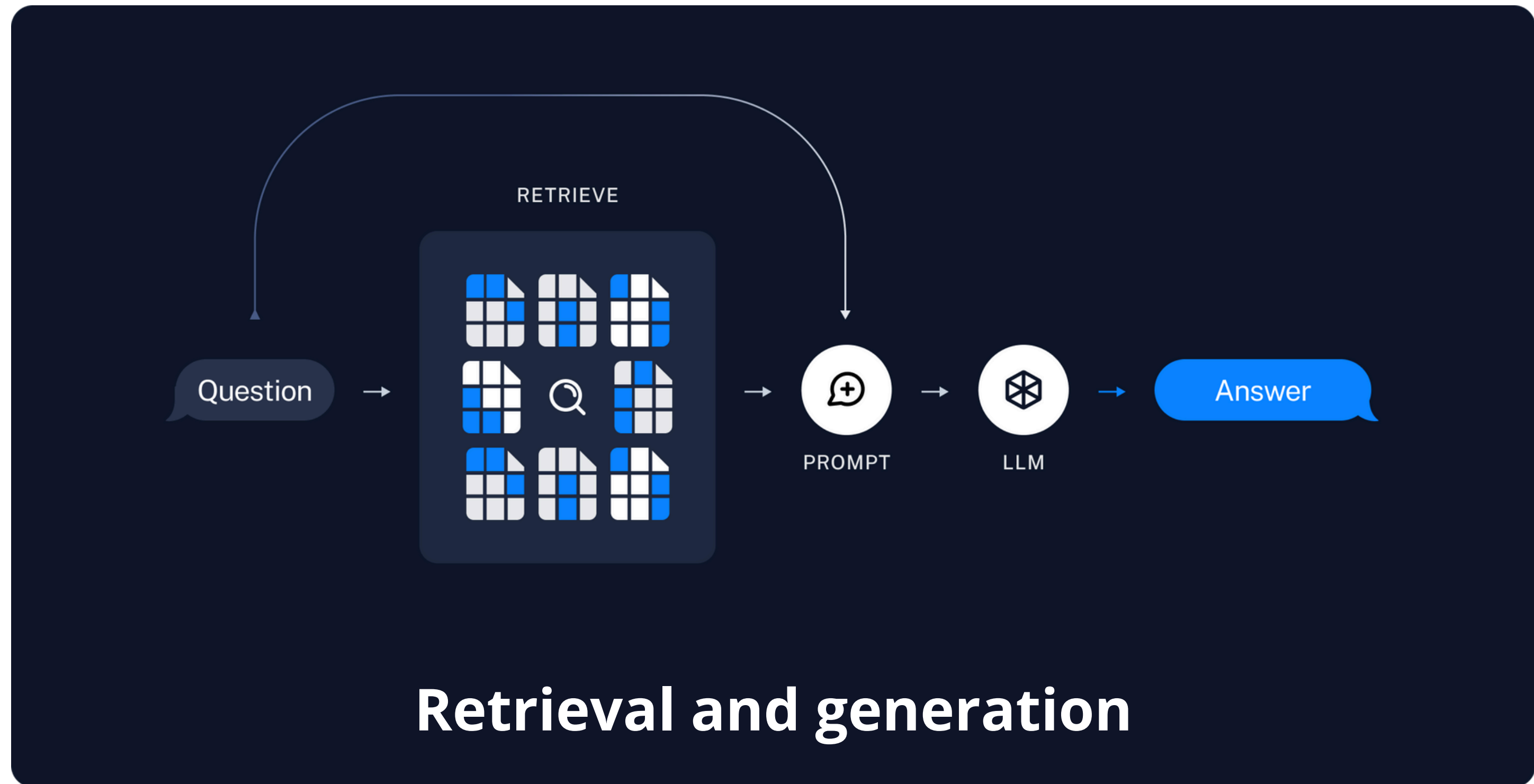


- **LangChain**: The backbone of the chatbot, orchestrating complex conversational flows.
- **ChatGroq**: Our primary chat model, ensuring responsive and engaging interactions.
- **LLAMA3** by Meta: Recently published by Meta, this powerful LLM enhanced our chatbot's language capabilities.
- **FAISS** by Facebook AI: Implemented as the Vector Database, it provided efficient and scalable similarity search.
- **Huggingface Embeddings**: Enabled precise and meaningful text embeddings, crucial for accurate Q&A matching.

Steps Performed



Steps Performed



Code Link :

EXAMPLES

ChatGroq with LLAMA3 Demo ✨ 🔗

Choose a PDF file



Drag and drop file here

Limit 200MB per file • PDF

Browse files



Idiomatic Expression In IELTS WRITING.pdf 324.4KB



http://localhost:8000/uploads/Idiomatic_Expression_In_IELTS_WRITING.pdf

Enter the Question from your Mind:

what is document all about ?

Document Embeddings

Based on the provided context, it appears that the document is about clarifying the meanings of various words and phrases related to various fields and concepts. The examples provided illustrate the different meanings of words and phrases, such as "to provide clarification", "to create a specialized area", "to set the stage", "to paint a comprehensive picture", and others.

The context suggests that the document is intended to provide a clear understanding of the different meanings of these words and phrases, helping readers to better comprehend and apply them in their own context.

Developed by KARTAVYA MASTER 🌐

PORTFOLIO : [CLICK ME](#)

Q&A ChatBot with RAG (LLAMA3)



Choose a PDF file



Drag and drop file here

Limit 200MB per file • PDF

Browse files



attention.pdf 0.5MB



<http://localhost:8000/uploads/attention.pdf>

Document Embeddings

Enter the Question from your Mind:

why attention ? write in 100 words

The authors propose the Transformer, a new sequence transduction model that relies solely on attention mechanisms, dispensing with recurrence and convolutions. The authors motivate the use of attention by explaining that sequence transduction problems, such as language modeling and machine translation, can be viewed as a complex sequence-to-sequence prediction task. They argue that attention mechanisms are better suited for such tasks because they allow the model to focus on the relevant parts of the input sequence, rather than relying on recurrence and convolutional operations.

Developed by KARTAVYA MASTER 🌐

PORTFOLIO : [CLICK ME](#)



Links

Streamlit Web

Hugging Face

GitHub



THANK YOU

"Generative AI is teaching us that the way you speak is actually code itself"
– Lisa Huang,