

CHAT WITH YOUR DATA



TEAM MEMBERS

TEAM 3

- **Abdullah Abdelaziz**
- **Mohamed Allam**
- **Saief El-SHAHED**
- **AbdelRahman**



TABLE OF CONTENTS

01 Project Overview

02 Introduction to LLMs

03 What is RAG

04 RAG vs. Fine-Tuning

05 Our Approach

06 Results

07 Interactive Demo

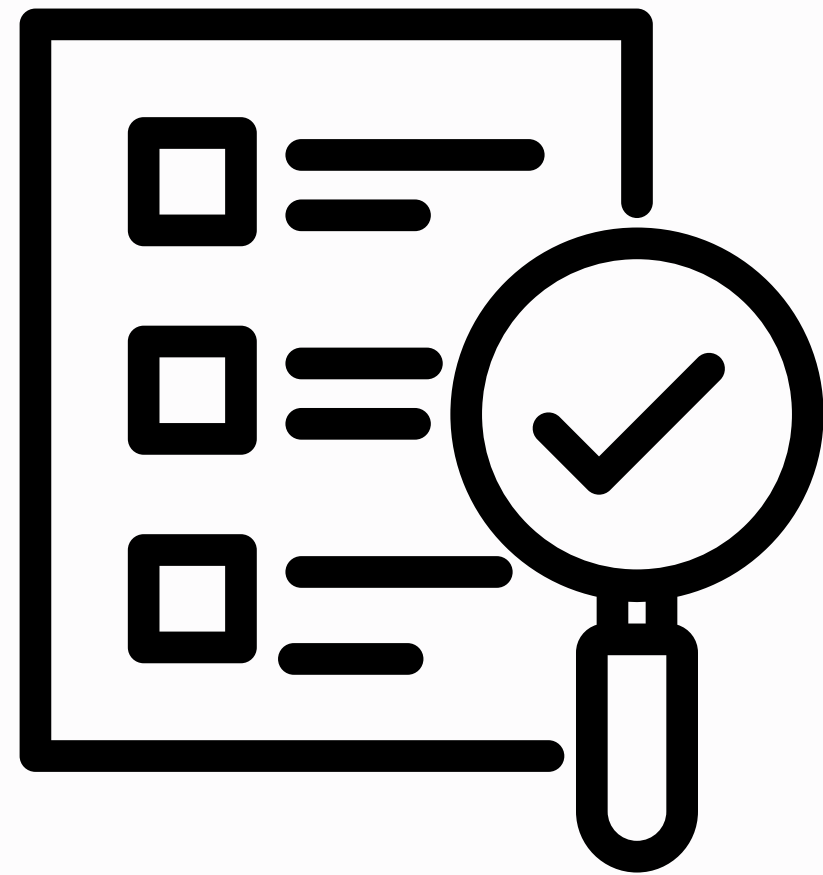
08 Challenges and Learnings

-- Future Enhancements

PROJECT OVERVIEW



CHAT WITH DATA



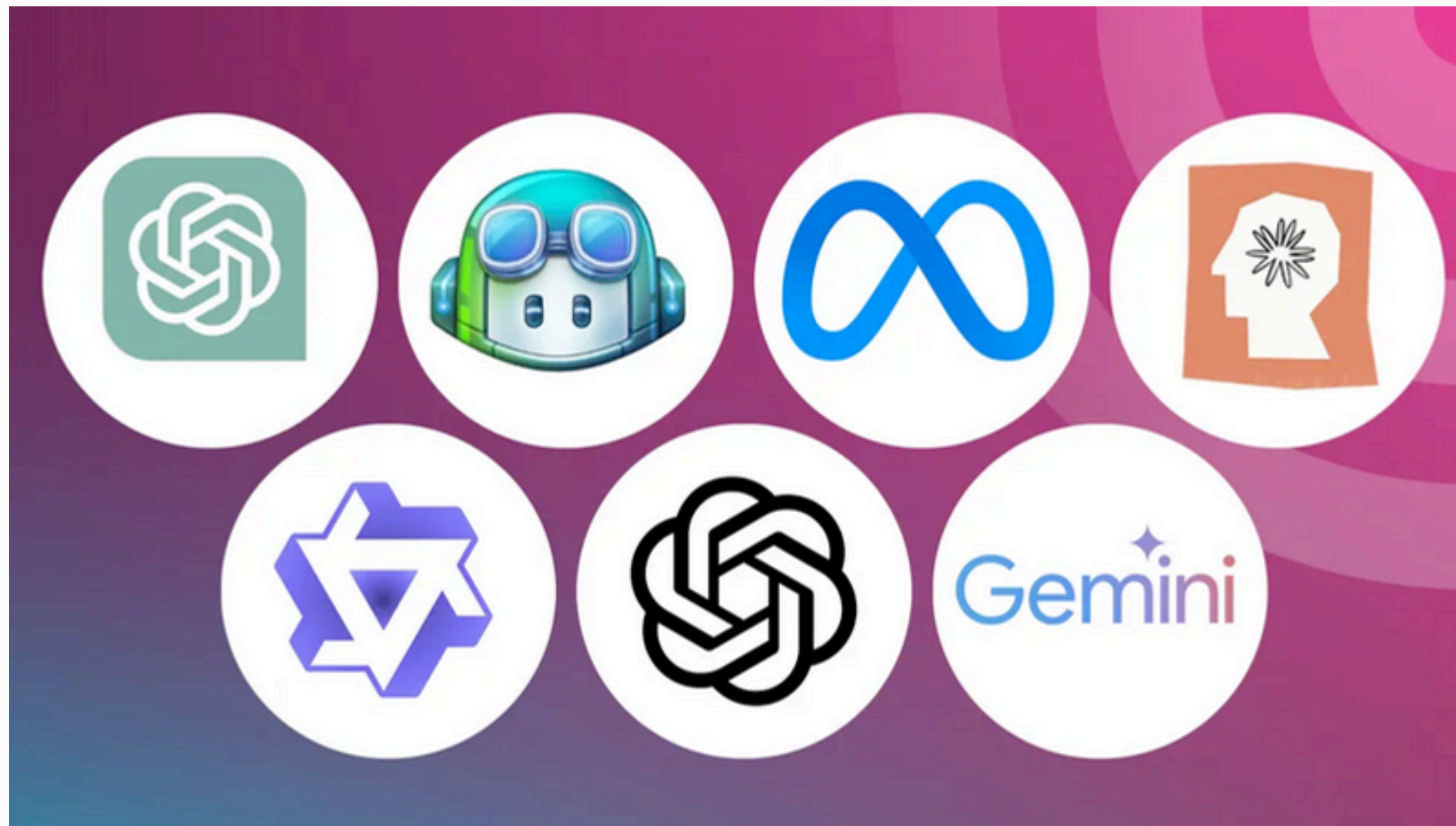
INTERACTING WITH US CRIMINAL CODE



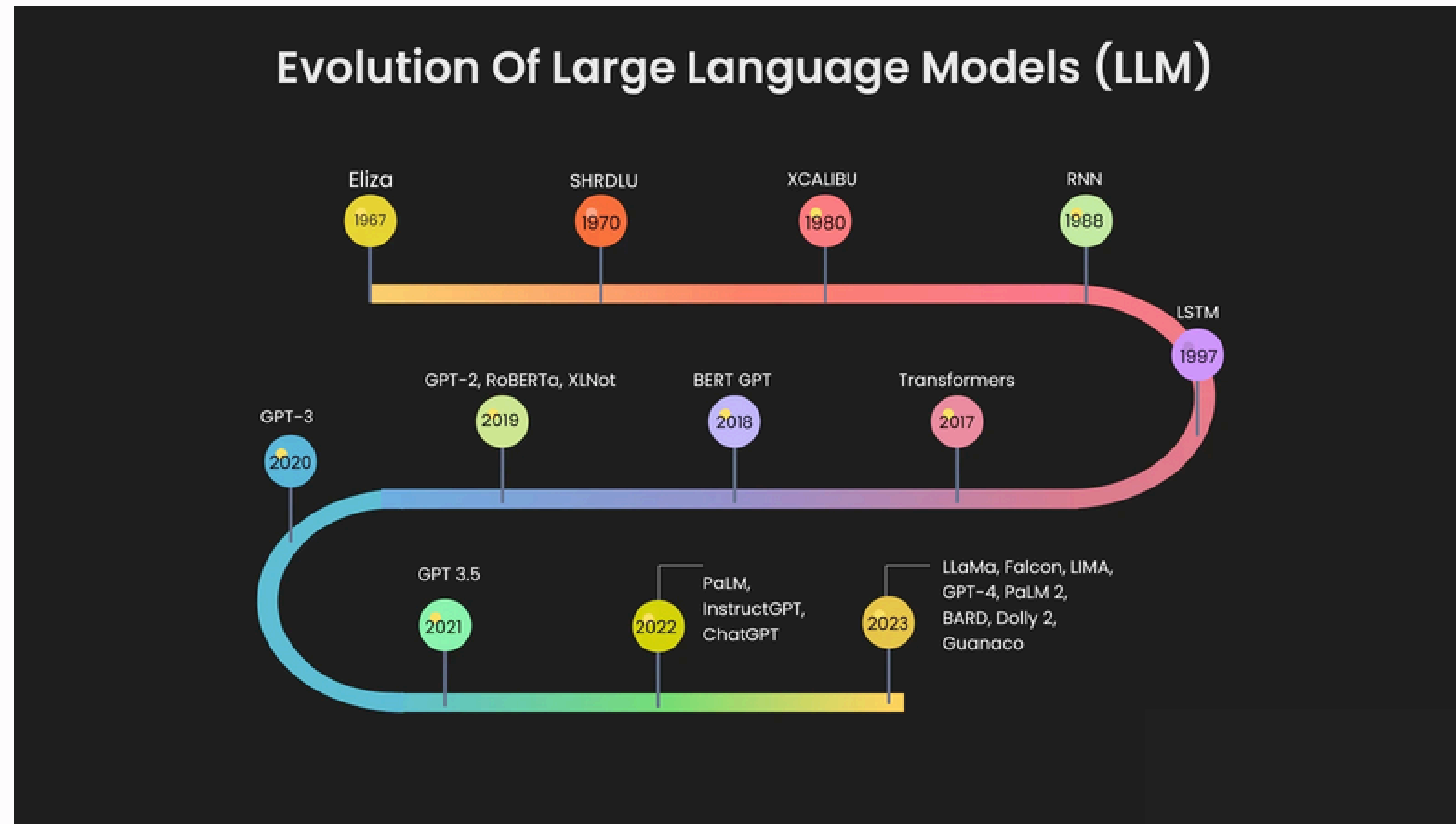
LARGE LANGUAGE MODELS



LARGE LANGUAGE MODELS (LLMs)



REVOLUTION OF LLMs



RETRIEVAL-AUGMENTED GENERATION



WHAT IS RAG

RAG (Retrieval-Augmented Generation) is an AI framework that combines the strengths of traditional information retrieval systems (such as search and databases) with the capabilities of generative large language models (LLMs). By combining your data and world knowledge with LLM language skills, grounded generation is more accurate, up-to-date, and relevant to your specific needs

RAG **VS** FINE-TUNING



WHAT IS FINE-TUNING ?

- Fine-tuning in deep learning is a form of transfer learning. It involves taking a pre-trained model, which has been trained on a large dataset for a general task such as image recognition or natural language understanding, and making minor adjustments to its internal parameters.
- The goal is to optimize the model's performance on a new, related task without starting the training process from scratch.

WHY WE CHOOSE RAG

Summary Table: RAG vs. Fine-tuning

Feature	RAG	Fine-tuning
Adaptability to Dynamic Information	✓ Adapts well with access to latest information	✗ May require updates to stay relevant
Customization and Linguistic Style	✗ Limited customization based on retrieved data	✓ High degree of personalization possible
Data Efficiency and Requirements	✓ Leverages external datasets, less labeled data needed	✗ Requires substantial, task-specific training data
Efficiency and Scalability	✓ Cost-effective and scalable with external data	✗ Higher initial resource investment required
Domain-Specific Performance	✓ Broad topical coverage, versatile	✓ Deep, precise domain expertise



APPROACH

APPROACH

- 01 Data Collection: Building the Legal Knowledge Base**
- 02 Integrating Llama 3 for Natural Language Understanding**
- 03 Implementing the Retrieval-Augmented Generation Pipeline**
- 04 Building the Frontend**
- 05 Testing and Refining the System**

DATA COLLECTION: BUILDING THE LEGAL KNOWLEDGE BASE

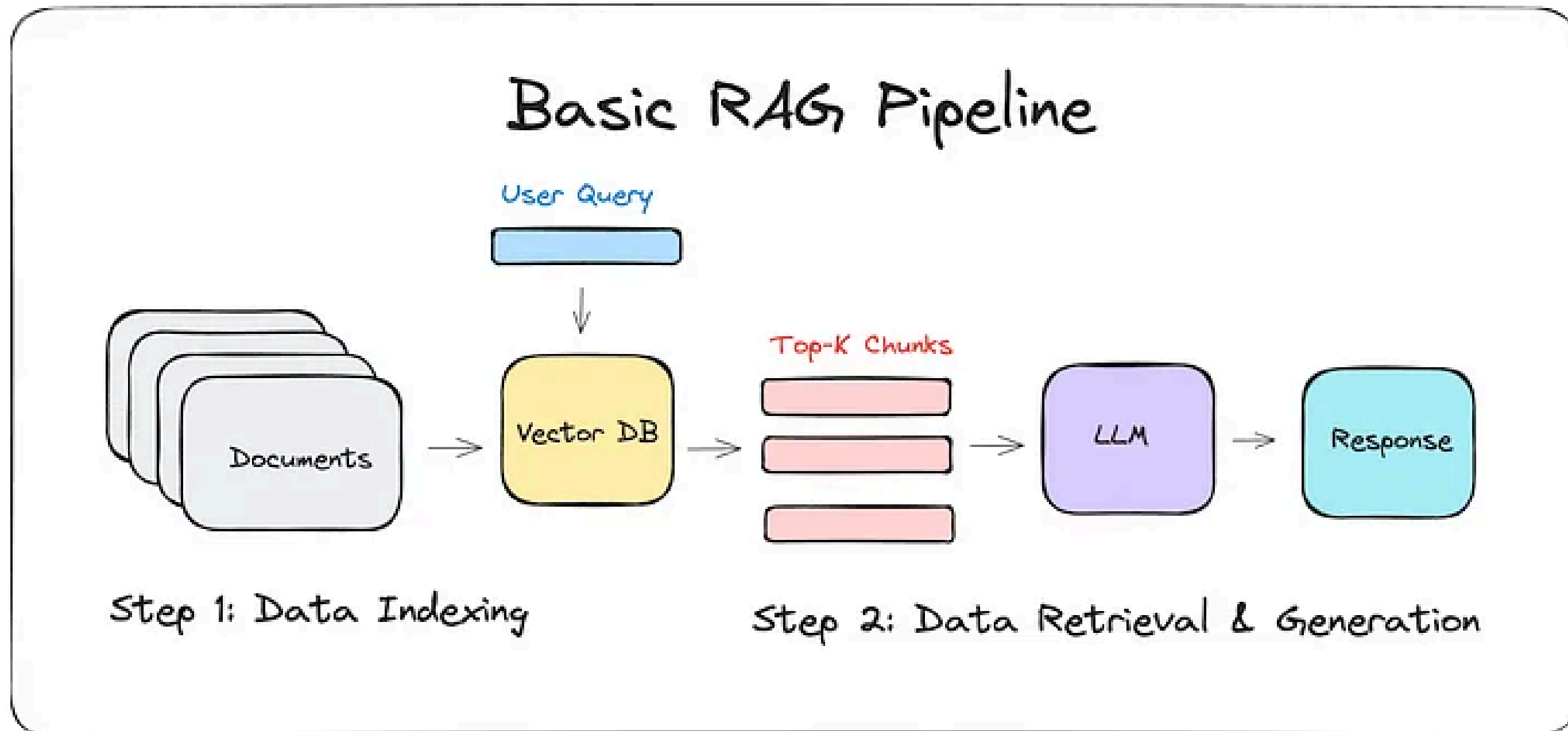
- **Data Understanding**
- **Data Collecting**
- **Data preprocessing**

INTEGRATING LLAMA 3 FOR NATURAL LANGUAGE UNDERSTANDING



Hugging Face

IMPLEMENTING THE RETRIEVAL-AUGMENTED GENERATION PIPELINE



BUILDING THE FRONTEND



VS





Chat With Data

DEPI Genrative AI 'TEAM 3'

Uploed a valid PDF file

📎 Upload a PDF



Drop File Here

- or -

Click to Upload

File State

Question

Answer



PDF Question Answering System

Upload a PDF file



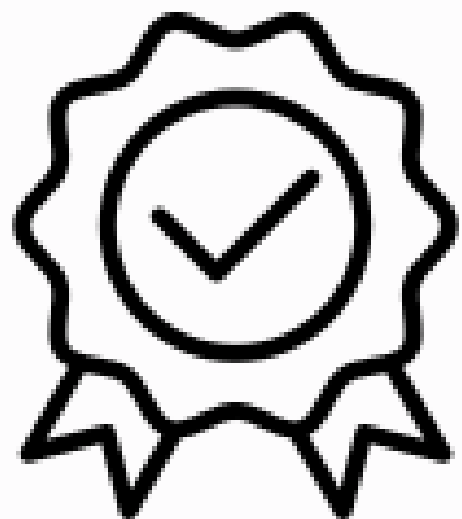
Drag and drop file here
Limit 200MB per file • PDF

Browse files

What is the Law of Murder?

Submit

TESTING AND REFINING THE SYSTEM



RESULTS





INTERACTIVE DEMO: CHATTING WITH YOUR OWN DATA

CHALLENGES & LEARNINGS

FUTURE ENHANCEMENTS

GENERAL PURPOSE APP

THANK YOU!

DO YOU HAVE ANY **QUESTIONS?**

