

Karna: A Compliance- Checking Chatbot for Code Review and Analysis

M.S. Applied Data Science
Summer 2024 Capstone
Ankit Gubilgari, Aashai Avadhani, Nakul Vadlamudi, Tyler
Sapsford





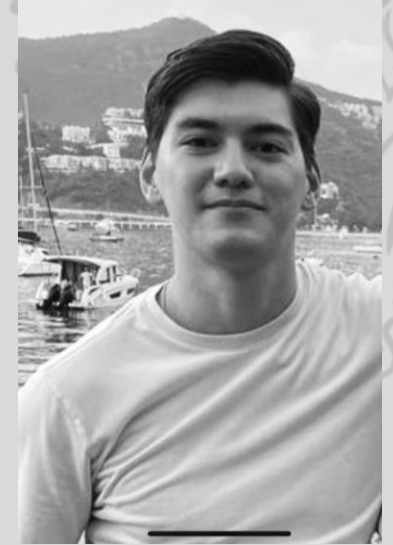
Ankit Gubiligari
Data Scientist
Costco IT



Aashai Avadhani
Data Scientist
Adobe



Nakul Vadlamudi
Analyst
E&Y



Tyler Sapsford
G10 Interest Rate Derivatives
Citadel - Global Fixed
Income

Meet the Capstone Team!

Executive Summary



Business Case

- Empower developers and data scientists to ensure compliance with privacy regulations.
- Reduce dependency on legal experts, accelerating development processes and cutting costs.



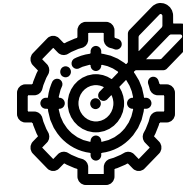
Source Data Overview

- Comprehensive dataset including GDPR, CPA, U.S. privacy legislation, and other legal texts, including code review templates.



Key Findings

- Immediate feedback on compliance status.
- Actionable insights for code modifications to meet regulatory standards.



Analysis Objective

- Develop a GenAI solution for real-time compliance checking.
- Focus on major privacy regulations: GDPR, CCPA, HIPAA, and the Privacy Act of 1974.



Methodology

- **Data Processing:** Legal texts and code snippets preparation.
- **Knowledge Extraction:** Key compliance rules identification.
- **Model Development:** Algorithms for compliance standards.
- **Validation:** Ensuring model accuracy and reliability.

How Should I Use A.I. Chatbots Like ChatGPT?

Large language models are already good at a wide variety of tasks.

HOME > TECH

Can ChatGPT be a doctor?

Asia Martin Apr 10, 2023, 3:48 AM

How Does ChatGPT Really Work?

Learning how a “large language model” operates.

Big Tech was moving cautiously on AI. Then came ChatGPT.

FORBES > INNOVATION

Google, Facebook and Microsoft helped build the scaffolding of AI. Smaller companies are taking it to the masses.

By Nitaasha Tikau, Gerrit De Vynck and Will Oremus

Updated February 20, 2023 at 10:43 a.m. EST. Published January 19, 2023 at 7:00 a.m. EST

TECHNOLOGY EXECUTIVE COUNCIL

How ChatGPT is changing the job hiring process, from the HR department to coders

PUBLISHED SAT APR 8 2023 11:00 AM EST. UPDATED APR 10 2023 11:00 AM EST. Kevin Travers, special to CNBC.com

Chatbots May ‘Hallucinate’ More Often Than Many Realize

When summarizing facts, ChatGPT technology makes things up about 3 percent of the time, according to research from a new start-up. A Google system's rate was 27 percent.

Here's How To Use AI—Like ChatGPT And Bard—For Everyday Tasks Like Creating A Budget, Finding A Ride Or Planning Meals

Arianna Johnson Forbes Staff
I cover the latest trends in science, tech and healthcare.

Generative A.I.'s Biggest Impact Will Be in Banking and Tech, Report Says

For some companies, the new technology is an opportunity to enhance productivity and profit. Will their workers benefit as well?

How Generative AI And ChatGPT Are Shaping The Contact Center

Muddu Sudhakar Forbes Council's Member
Forbes Technology Council
COUNCIL CHAIRMAN, Muddu Sudhakar

ChatGPT: Your AI-Powered Cryptocurrency Resource



Slamak Masnavi

10 Apr 2023 / In #AI, #cryptocurrency

What is GenAI?

The Industrial Economy Greets Generative AI With Open Arms

Generative AI is a business game-changer, but only if it's built on solid foundations

The technology is transforming a wide range of processes, including customer service, content creation and software coding.

HOME > TECH

A college student got a top grade for an essay written with the help of ChatGPT, report says

Beatrice Nolan Apr 10, 2023, 4:00 AM

Definition: GenAI refers to a class of artificial intelligence algorithms that can generate new content, such as text, images, or code, based on learned patterns from existing data.

Capabilities:

- **Content Creation:** Produces human-like text, realistic images, and complex code structures.
- **Enhanced Interactions:** Powers chatbots and virtual assistants for more natural and engaging user interactions.
- **Automated Processes:** Streamlines tasks by generating responses, summarizing information, and automating repetitive tasks.

How Should I Use A.I. Chatbots Like ChatGPT?

Large language models are already good at a wide variety of tasks.

HOME > TECH

Can ChatGPT be a doctor?

Asia Martin Apr 10, 2023, 3:48 AM

How Does ChatGPT Really Work?

Learning how a “large language model” operates.

Big Tech was moving cautiously on AI. Then came ChatGPT.

Google, Facebook and Microsoft helped build the scaffolding of AI. Smaller companies are taking it to the masses.

By Nitaasha Tikau, Gerrit De Vynck and Will Oremus

Updated February 10, 2023 at 10:43 a.m. EST. Published January 20, 2023 at 7:00 a.m. EST

TECHNOLOGY EXECUTIVE COUNCIL

How ChatGPT is changing the job hiring process, from the HR department to coders

PUBLISHED SAT APR 8 2023 11:00 AM EST. KEVIN TRAVERS, special to CNBC.com

FORBES > INNOVATION

BREAKING

Here's How To Use AI—Like ChatGPT And Bard—For Everyday Tasks Like Creating A Budget, Finding A House Or Planning Meals

Arianna Johnson Forbes Staff
I cover the latest trends in science, tech and healthcare.

Generative A.I.'s Biggest Impact Will Be in Banking and Tech, Report Says

For some companies, the new technology is an opportunity to enhance productivity and profit. Will their workers benefit as well?

FORBES > INNOVATION

How Generative AI And ChatGPT Are Shaping The Contact Center

Muddu Sudhakar Forbes Council's Member
Forbes Technology Council
COMMUNITY POST

ChatGPT: Your AI-Powered Cryptocurrency Resource



Slamak Masnavi

10 Apr 2023 / In #AI, #cryptocurrency

Our GenAI Solution:

Chatbots May 'Take Over' Jobs More Often Than Many Realize

When summarizing facts, ChatGPT technology makes things up about 3 percent of the time, according to research from a new start-up. A Google system's rate was 27 percent.

How to use ChatGPT to write code

Can ChatGPT replace programmers? What programming languages does ChatGPT know? We answer these and your other generative AI coding questions.

Generative AI is a business game-changer, but only if it's built on solid foundations

The technology is transforming a wide range of processes, including customer service, content creation and software coding.

HOME > TECH

A college student got a top grade for an essay written with the help of ChatGPT, report says

Beatrice Nolan Apr 10, 2023, 4:00 AM

Empowers Data Scientists and Software Engineers to own Privacy and compliance issues within their code

Karna: Our GenAI Solution

1

Market Opportunity

- Vanta, OneTrust: New category of privacy automation
- Liminal projects a TAM of \$2.2 billion in 2024, reaching \$5.4 billion by 2028 (CAGR of 19.3%).
- Mordor Intelligence estimates \$3.27 billion in 2024, reaching \$8.03 billion by 2029 (CAGR of 19.72%).
- Optimistic projection: \$13.92 billion by 2031.
- Market experiencing significant growth

2

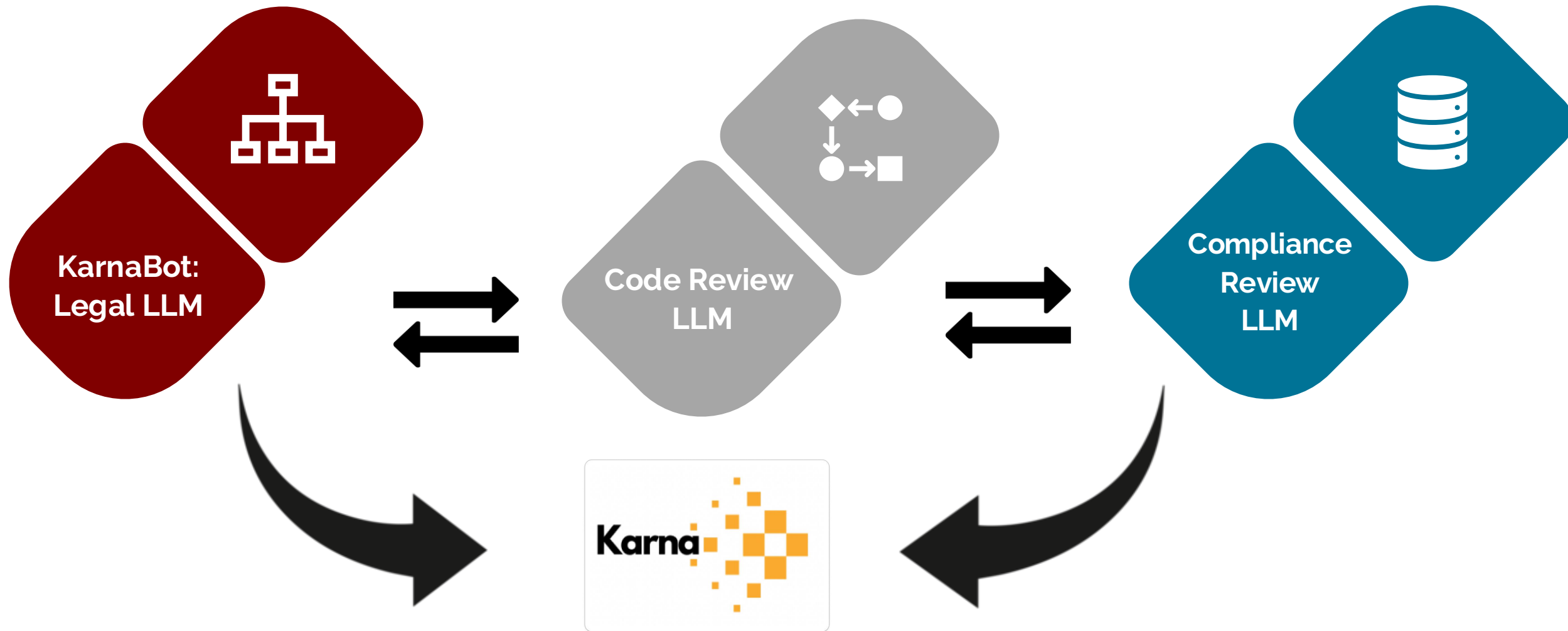
Motivation

- **Challenges: Complex Privacy Legislation**
- **Traditional need for legal expert consultations**
- **Inefficient Back and Forth between legal and data professionals**

Karna:

- **GenAI-Driven Compliance checks**
- **Integration into workflows and code bases**
- **Focus on Innovation, reduce legal complexities**

Three-step Approach to Producing a Functional Solution




KarnaBot: Our Legal and Privacy Expert


Sample Query:

Back

KarnaBot

 Can I use a person's last name in my code in regards to HIPAA?

Example Output:

 Overview of HIPAA

The Health Insurance Portability and Accountability Act (HIPAA) is a federal law that protects the privacy of individually identifiable health information. It applies to healthcare providers, health plans, and healthcare clearinghouses.

Compliance Risk

Using a person's last name in your code could potentially violate HIPAA if it is used in a way that could identify the individual. For example, if you are using the last name to create a unique identifier for the individual, this could be considered a violation of HIPAA.

Key Compliance Tasks

- Ensure that any personally identifiable health information (PHI) is encrypted at rest and in transit.
- Implement access controls to limit who can access PHI.
- Train employees on HIPAA compliance.
- Conduct regular security audits to identify and mitigate any potential vulnerabilities.

Potential Vulnerabilities

- Storing PHI in plaintext
- Allowing unauthorized access to PHI
- Failing to encrypt PHI
- Not training employees on HIPAA compliance

Legal Text and Associated Laws

HIPAA Privacy Rule

45 CFR § 164.514(b)(2)(i) requires covered entities to implement reasonable and appropriate safeguards to protect the privacy of PHI.

Examples of Non-Compliance

- Storing a patient's last name in a database without encryption
- Allowing a receptionist to access a patient's medical records without authorization
- Failing to train employees on HIPAA compliance

Example Code for Compliance

```
# Encrypt PHI before storing it in a database
encrypted_phi = encrypt(phi)
```

Example Code for Non-Compliance

```
# Store PHI in plaintext in a database
phi = "John Doe"
```

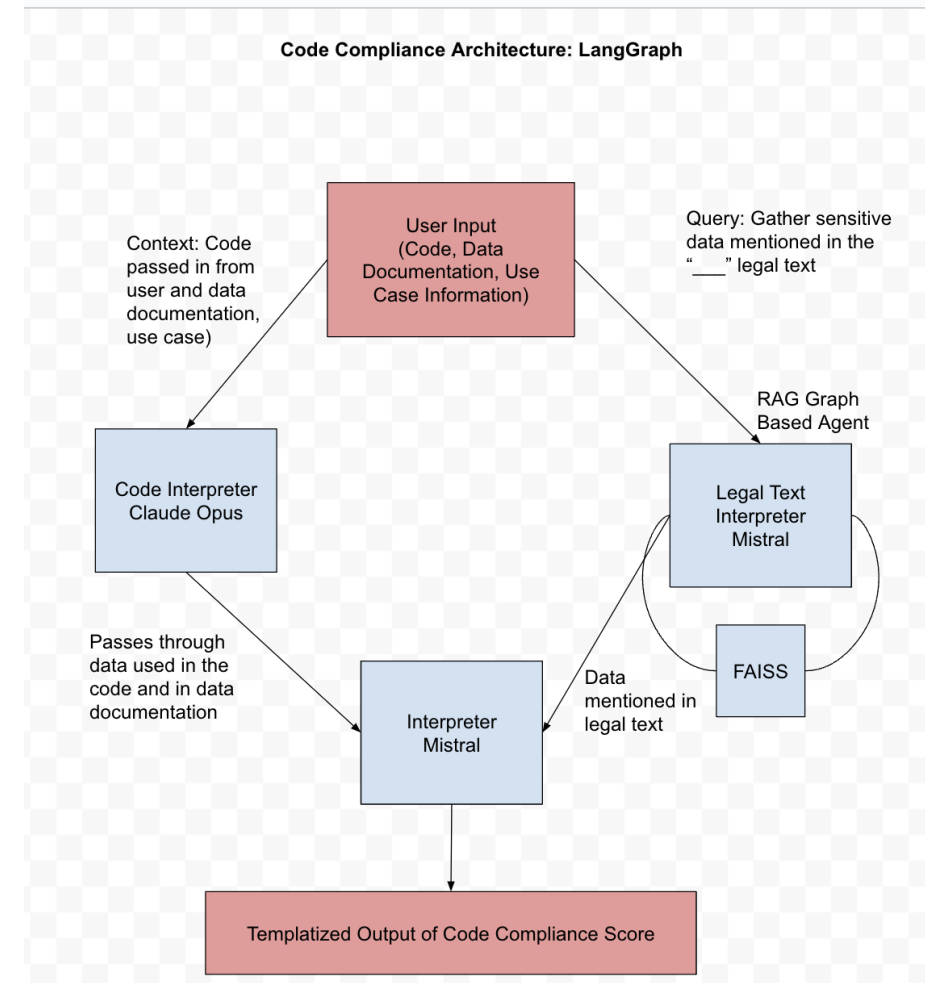


Model reads
information
from vector store to
generate outputs to
the given query
based on the prompt.

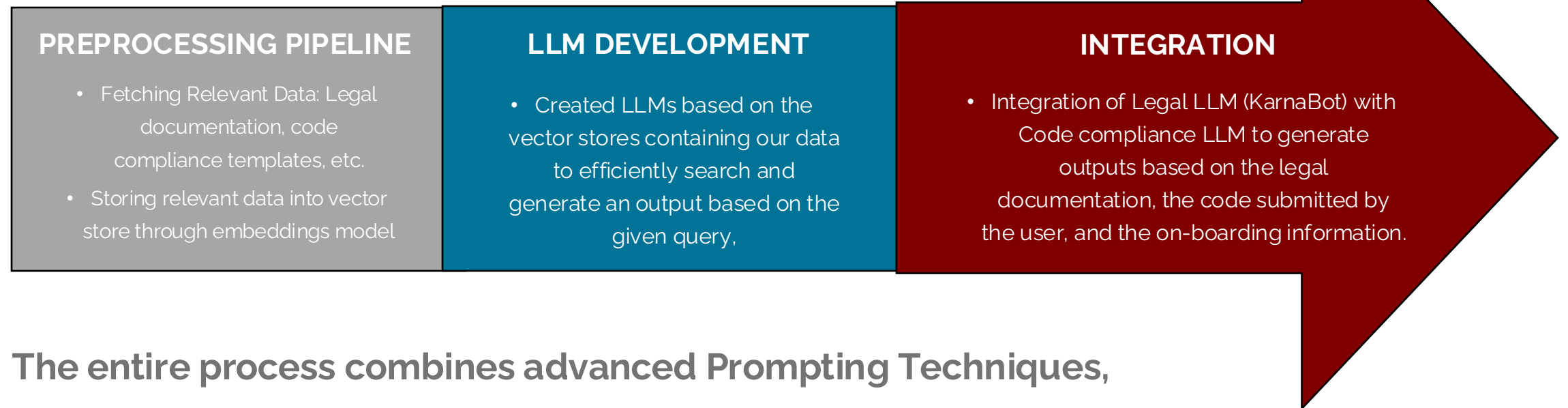
Code Compliance Architecture

Using a LangGraph approach we are able to conduct compliance assessments on code for any developer or privacy counsel

Use Langgraph and combination of Mistral, Anthropic Claude Model, and Gemini to orchestrate code compliance for a software project



RAG Pipeline Using LangChain, Gemini Pro, FAISS, Mistral, and Anthropic



The entire process combines advanced Prompting Techniques,

Multiple LLM models like Mistral, Anthropic, and Gemini

Storing the latest regulatory compliance laws recently passed (GDPR, EU AI Act, HIPAA)

Combined through a code compliance architecture through LangGraph

The Karna logo is displayed on a white rectangular background. It features the word "Karna" in a bold, black, sans-serif font. To the right of the text is a graphic composed of several orange squares of varying sizes, arranged in a pattern that suggests a stylized sunburst or a cluster of data points.

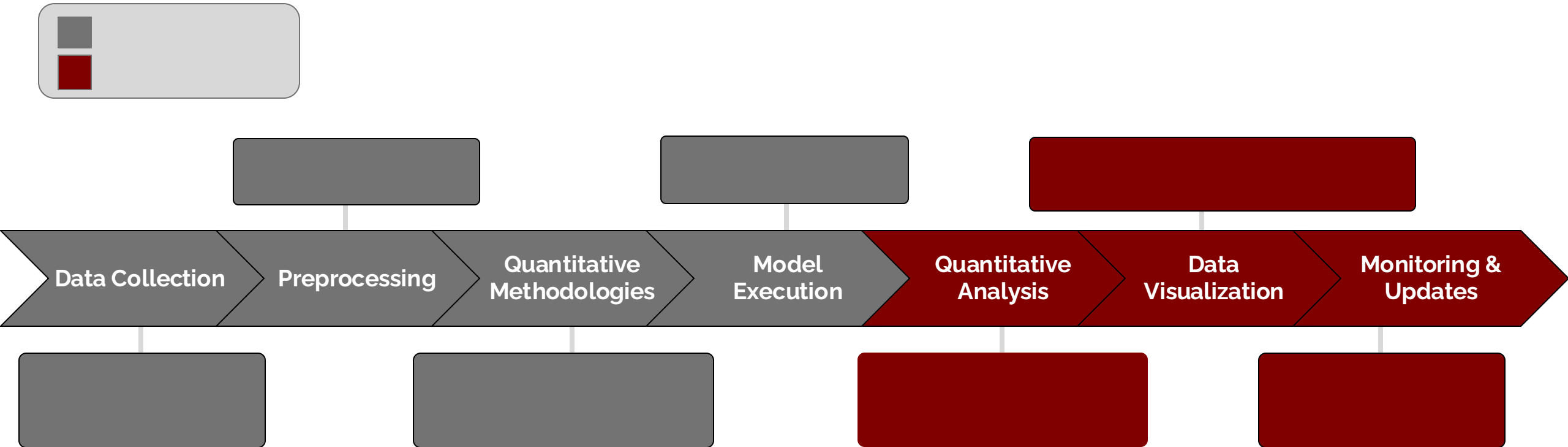
Karna

Thank You – Questions?

Appendix



Karna Timeline and Future Steps



References

