Curso de LangChain

¿Cómo crear aplicaciones con LLM?



LangChain

LLM y modelos de chat







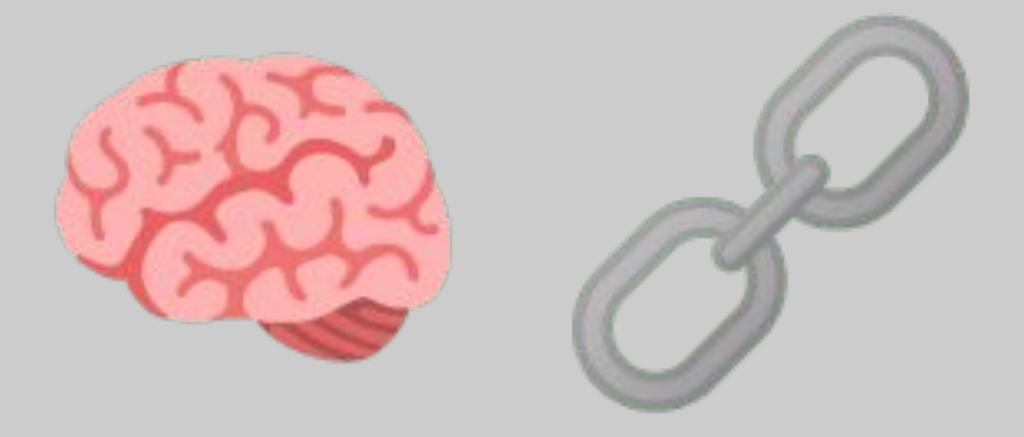




Indexar documentos



Memorias y cadenas



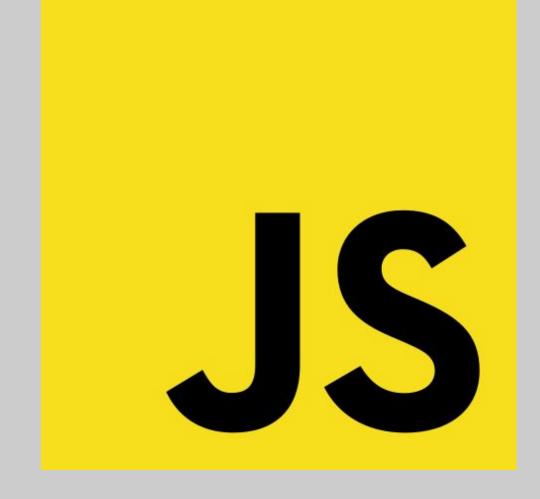
Unir modelos



Estructura de Langchain

Disponible en



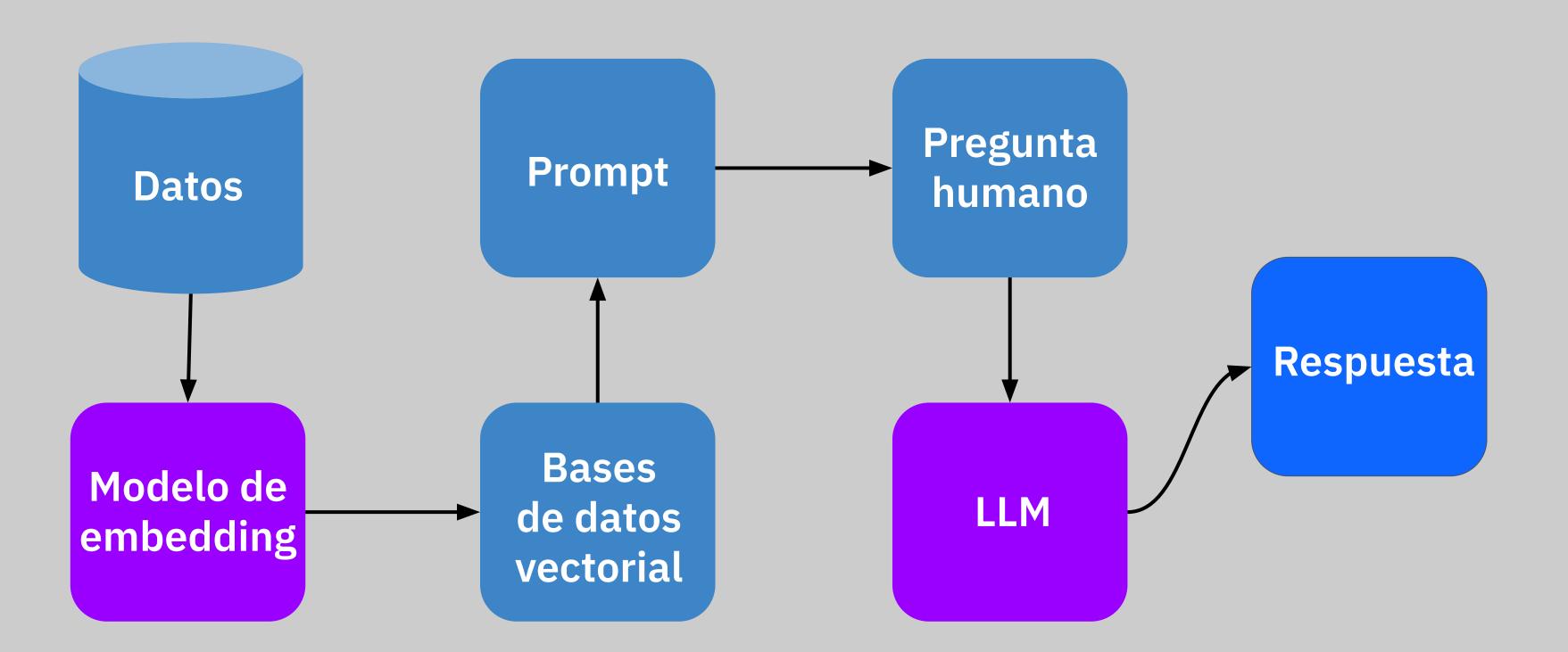


Estructura

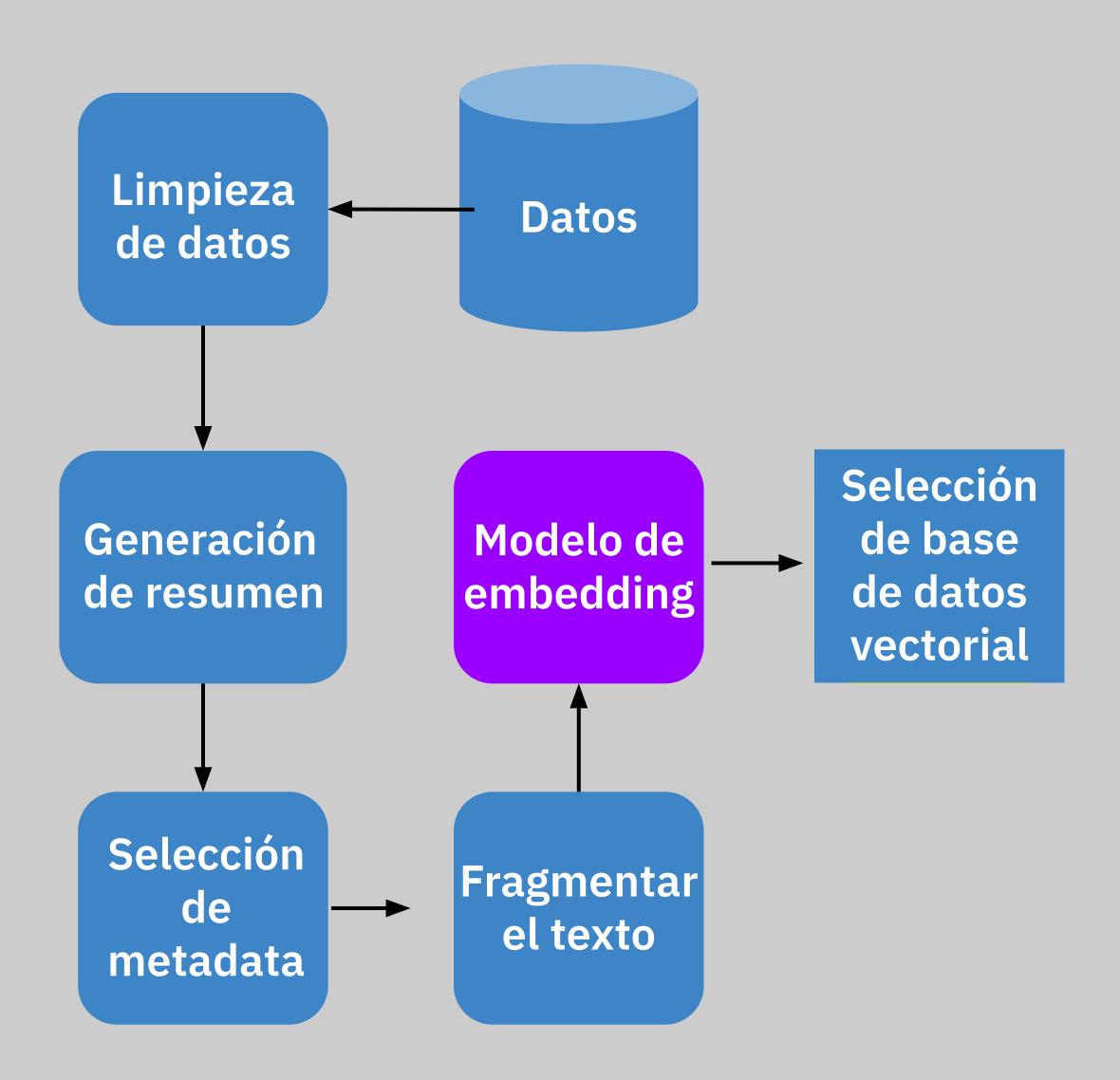
Claves:

- 1. Conexión con modelos
- 2. Conexión con datos
- 3. Encadenamiento de procesos

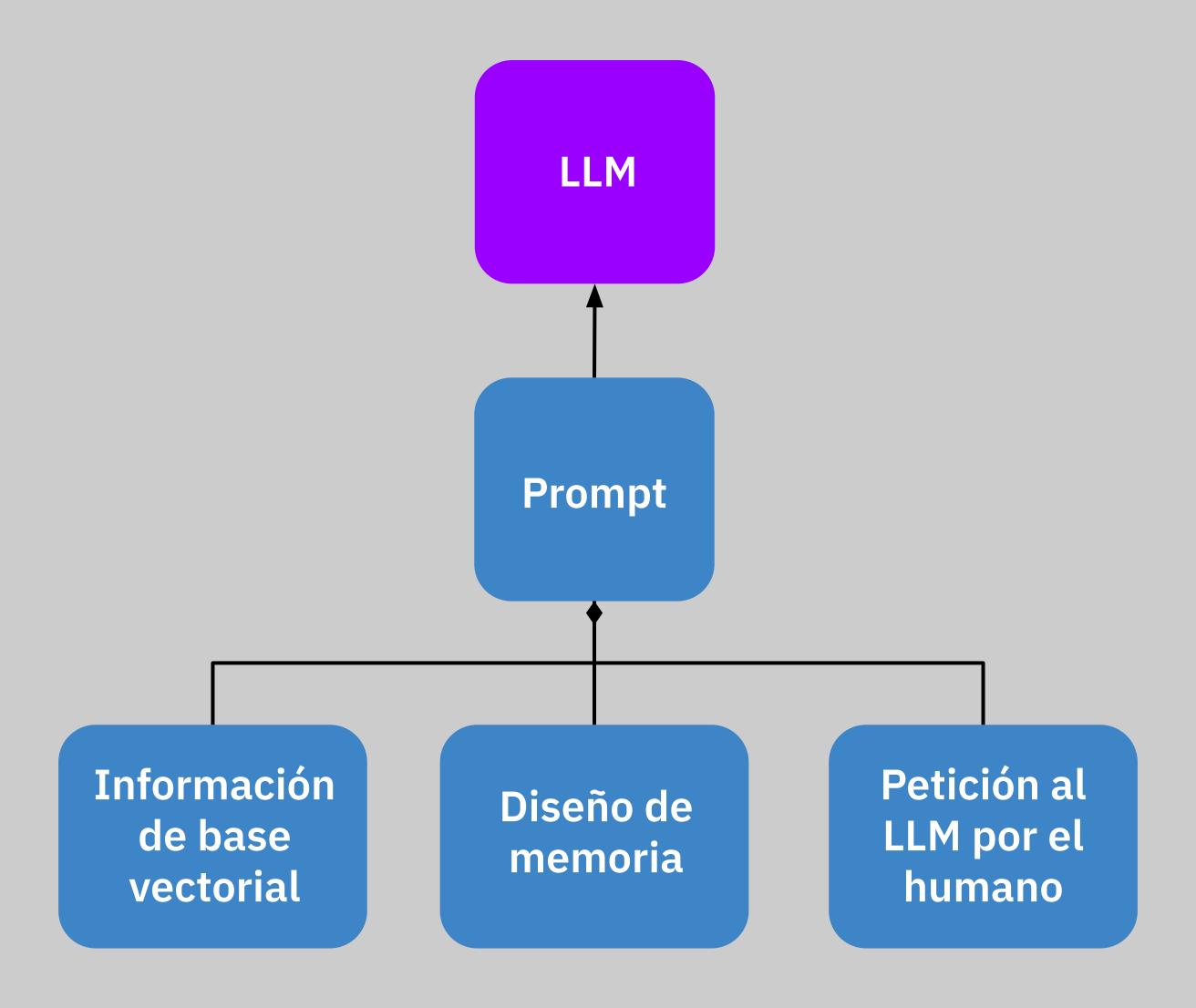
Fundamentos



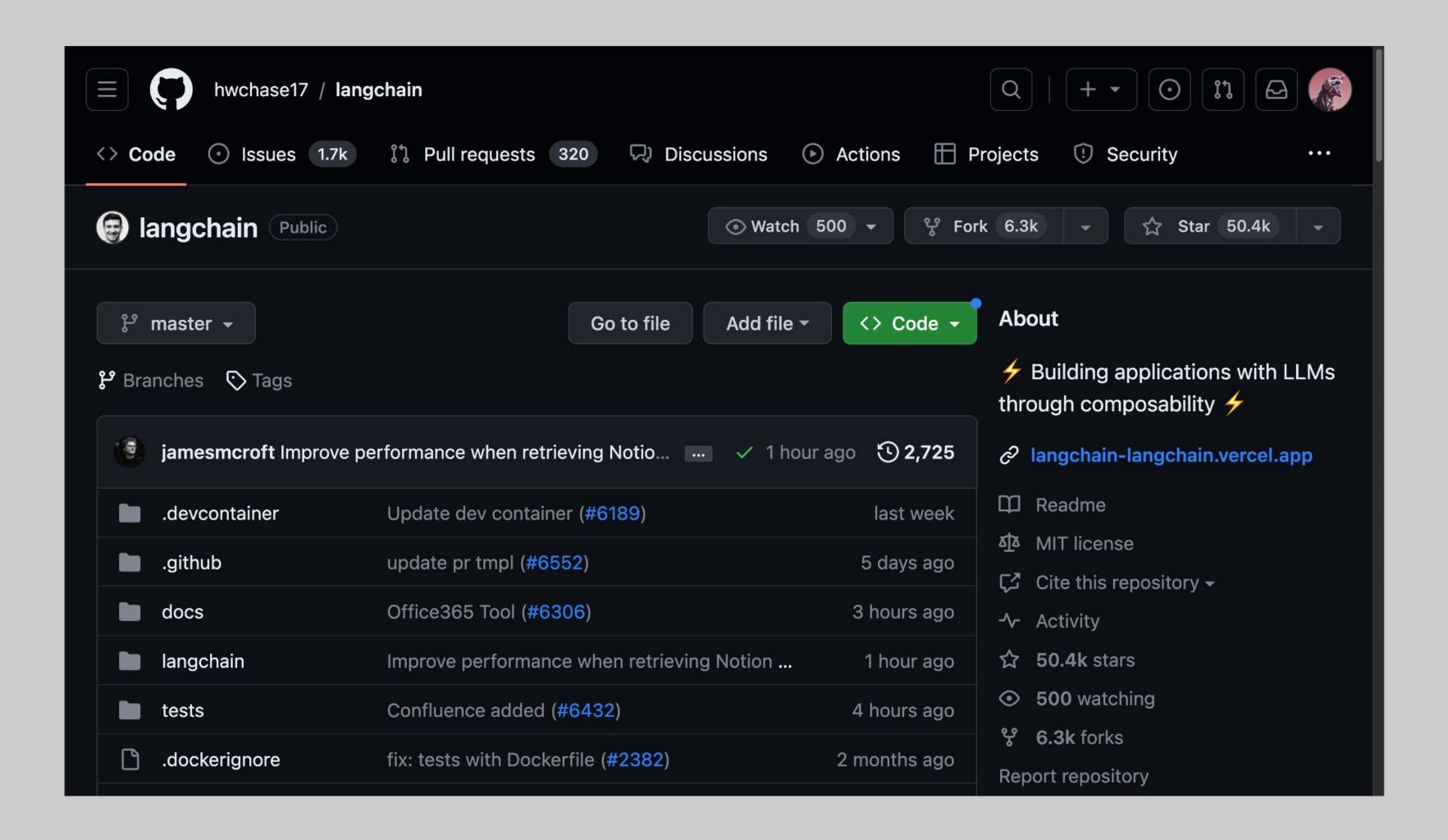
Ejemplo del procesamiento de datos



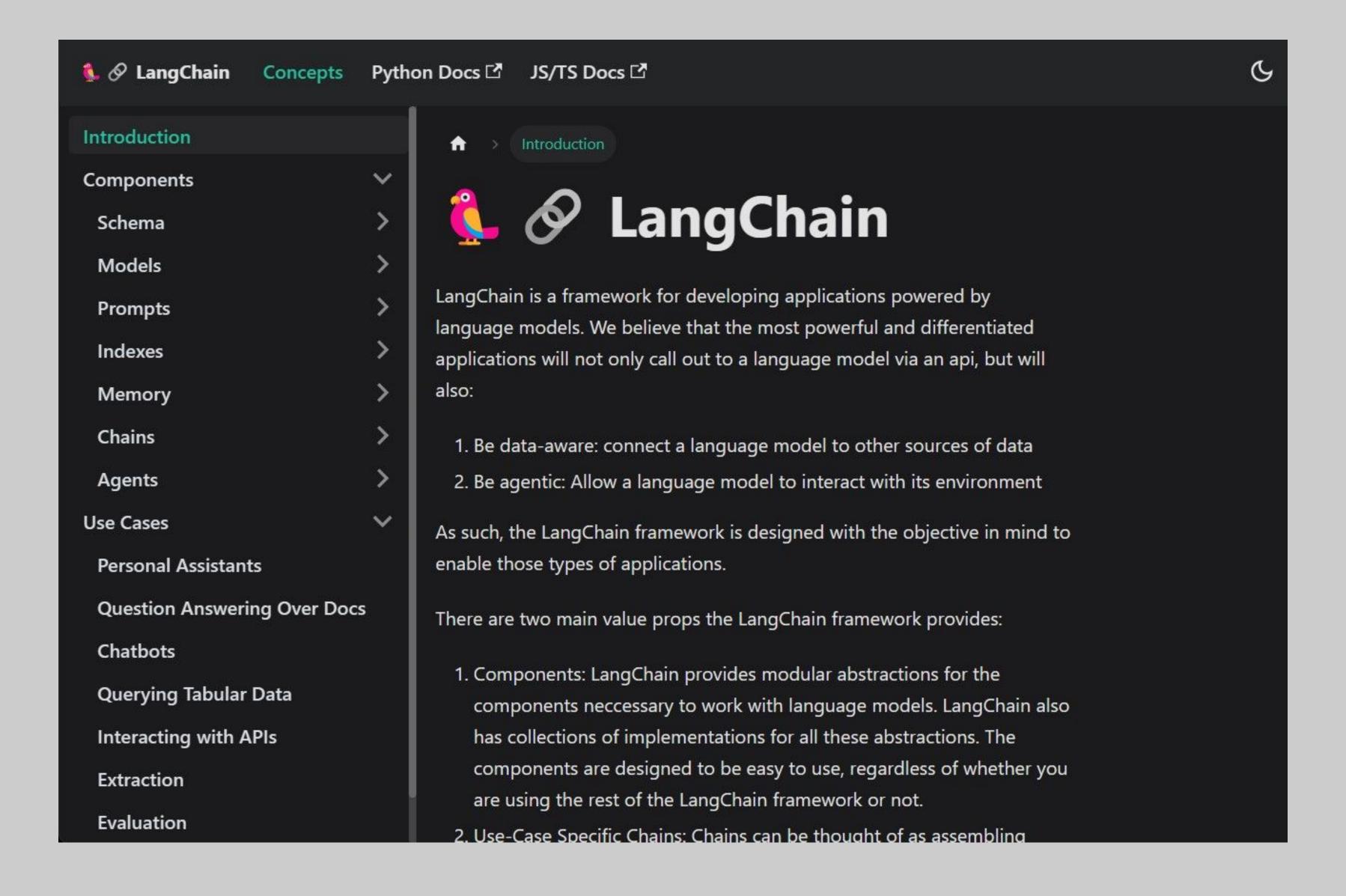
Ej. de uso de LLM



1. Repositorio en GitHub



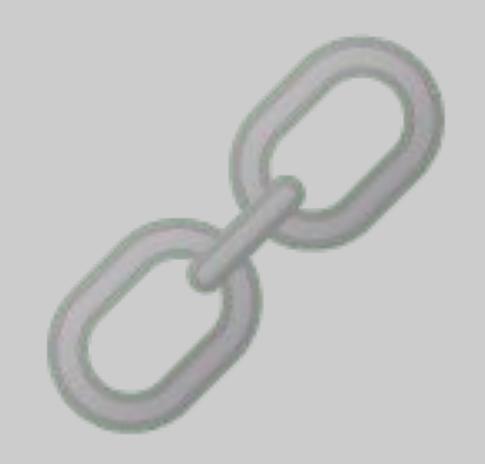
2. Documentación



Cadenas

LangChain

¿Qué es una cadena y cómo se relaciona con los modelos?



¿Cómo funcionan?

Tipos de cadenas

- Utility
- Foundational
 - o LLM
 - Transformation
 - Sequential

Indexes oindices

¿Cómo manejar documentos en LangChain?

Conexión a datos de documentos









Clases de indices

- Document loaders
- Text splitters
- Vectorstore

¿Por qué usar indices?

- LLM pueden interactuar con otros datos.
- Indexar datos no estructurados.
- Consultas precisas.

Uso de embeddings y bases de datos vectoriales

Embedding class

- Varios modelos de embedding.
- Interfaz estándar para todos los modelos.

Embedding class methods

- embed_documents
- embed_query

Embedding integrations









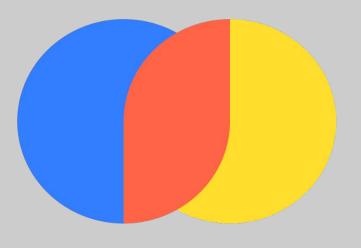


Vector stores

- Almacenar embedded data.
- Consultas a la data de los docs.

Vector store integrations







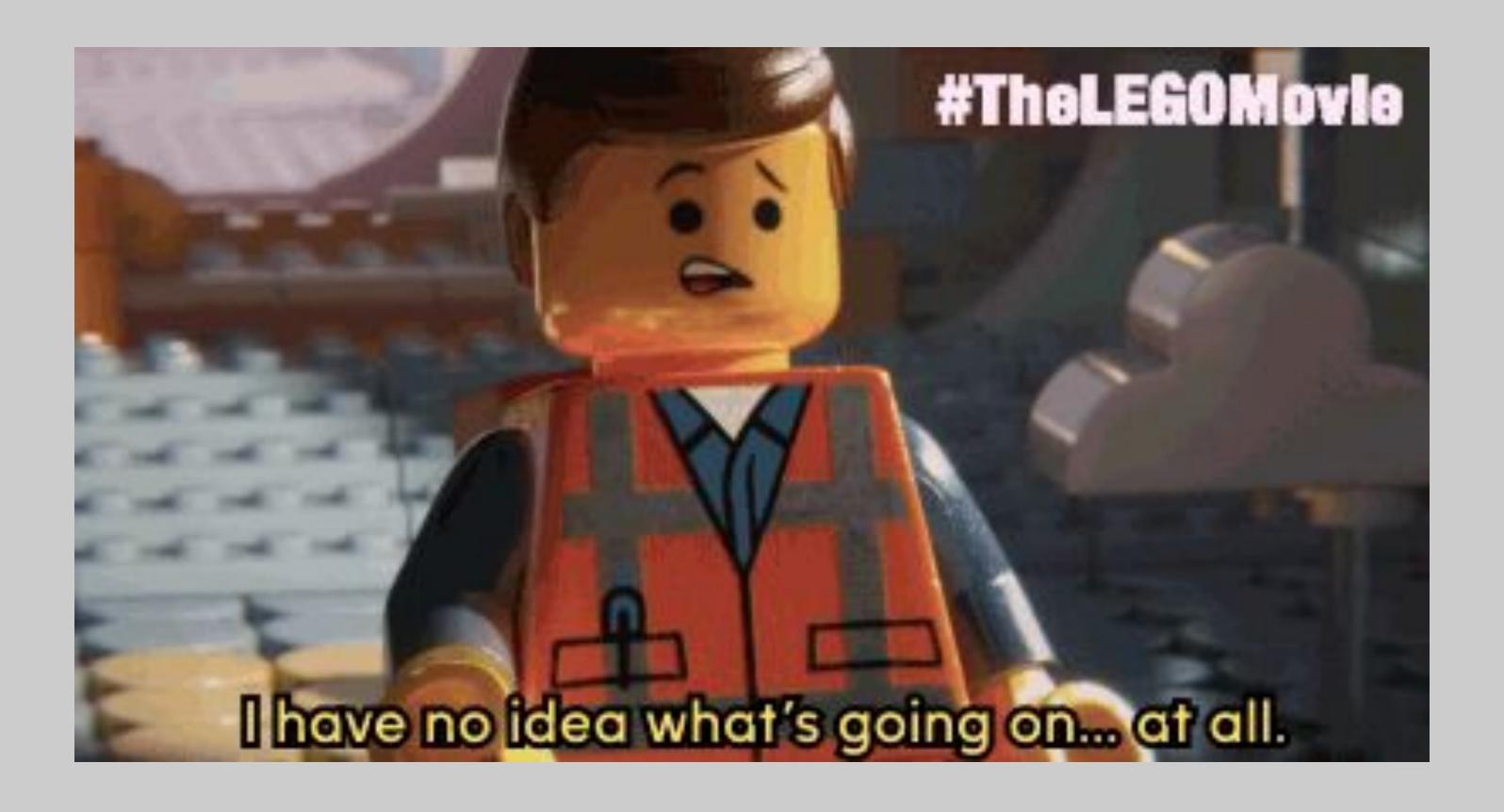






Memoria en LangChain

¿Para qué se utiliza?



Un chatbot por defecto no recuerda lo que le dijiste.

Tipos de memoria

- Corto plazo
- Largo plazo

from langchain.memory import ChatMessageHistory

history = ChatMessageHistory()

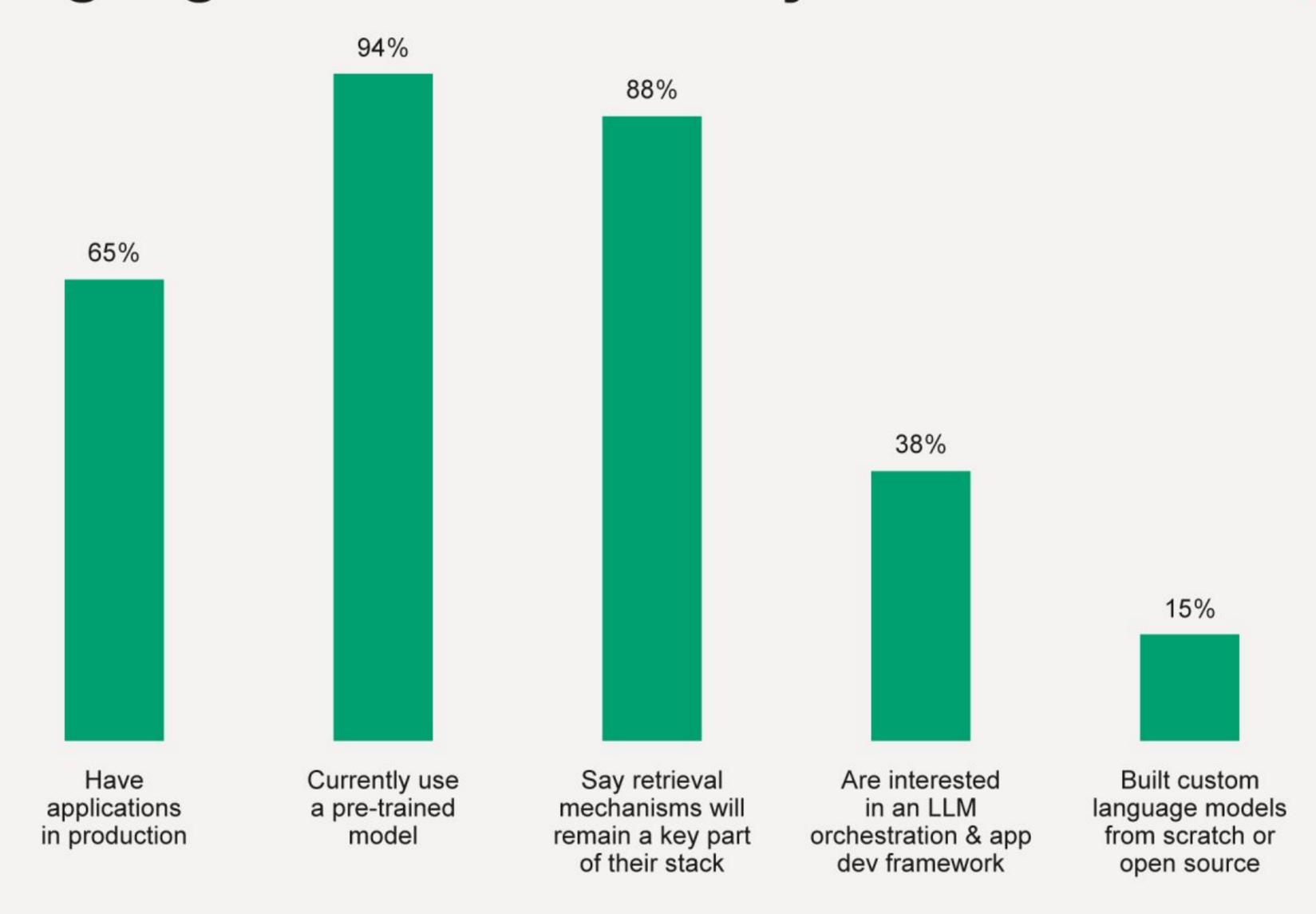
history.add_user_message("hi!")

history.add_ai_message("whats up?")

LangChain y LLM en evolución constante

Language model API survey

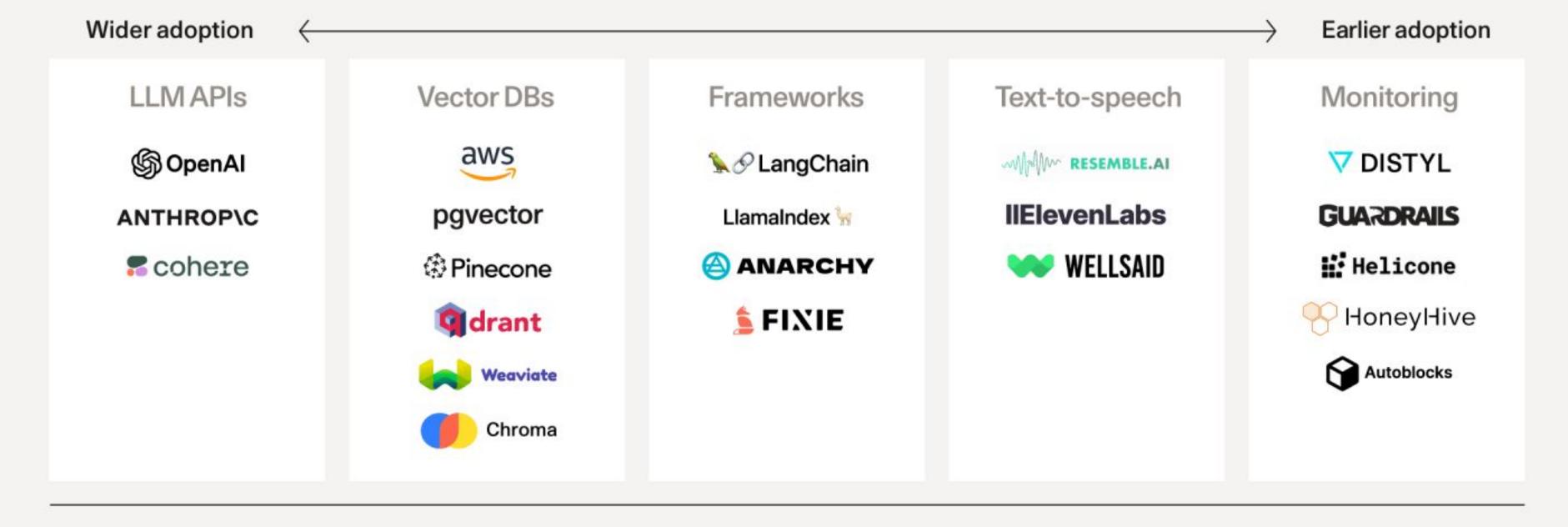




Fuente: Sequoia Capital - The New Language Model Stack

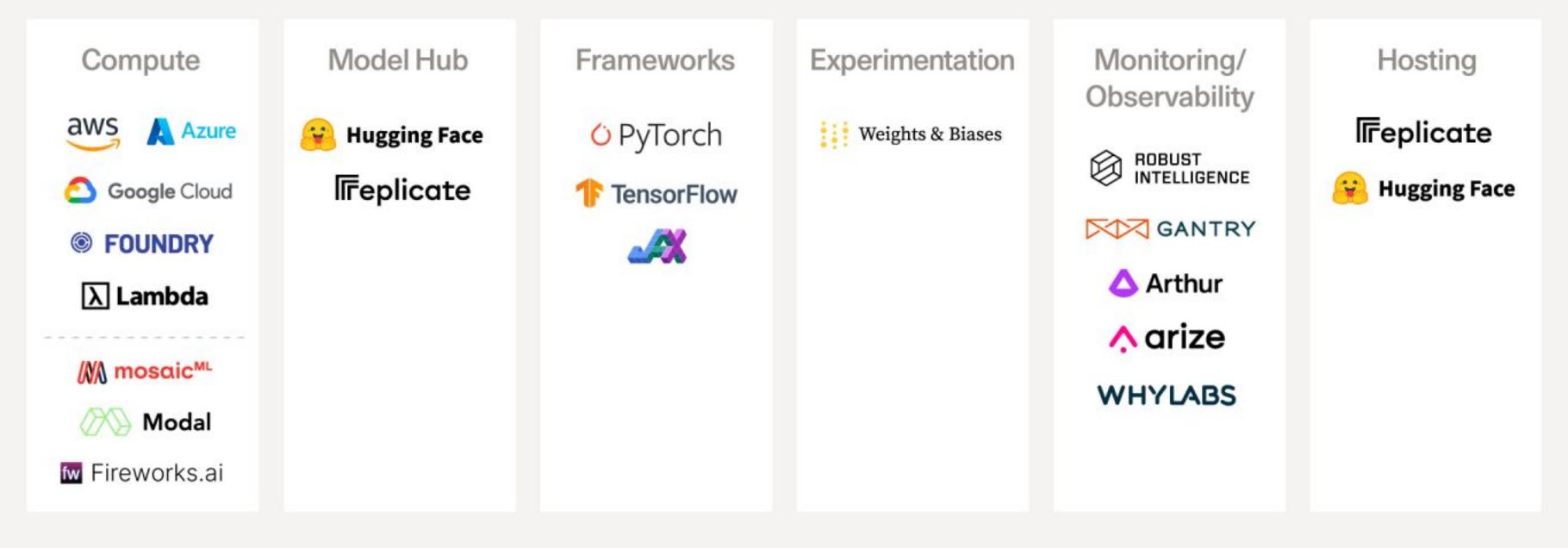
LLM API Stack





Fuente: Sequoia Capital - The New Language Model Stack

Custom Model Training / Tuning Stack



Note: this is not exhaustive. This is based on customer-reported usage and perception.

Fuente: Sequoia Capital - The New Language Model Stack