



Introducción en Gen-IA

www.datapath.ai



Miguel Cotrina

- **Perfil Académico**
 - Bachiller de Ingeniería de software por la Universidad Tecnológica del Perú.
 - Estudiante de Maestría en ciencia de datos por la Universidad Ricardo Palma.
- **Perfil Profesional**
 - Arquitecto de datos en Clínica Internacional
 - Consultor de Big Data, cloud e IA en empresas privadas y públicas



/mcotrina



mac@digisoftperu.com



Agenda

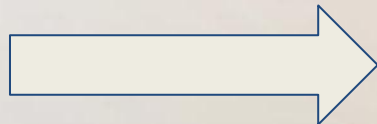
1. Gen-AI
2. RAG
3. Demostracion

Gen-AI

Gen - AI

Los modelos LLM responden de acuerdo a los datos con los que fueron pre entrenados

Crea una función para sumar en python



GPT(LLM)



Claro, aquí tienes una función simple en Python que suma dos números y devuelve el resultado:

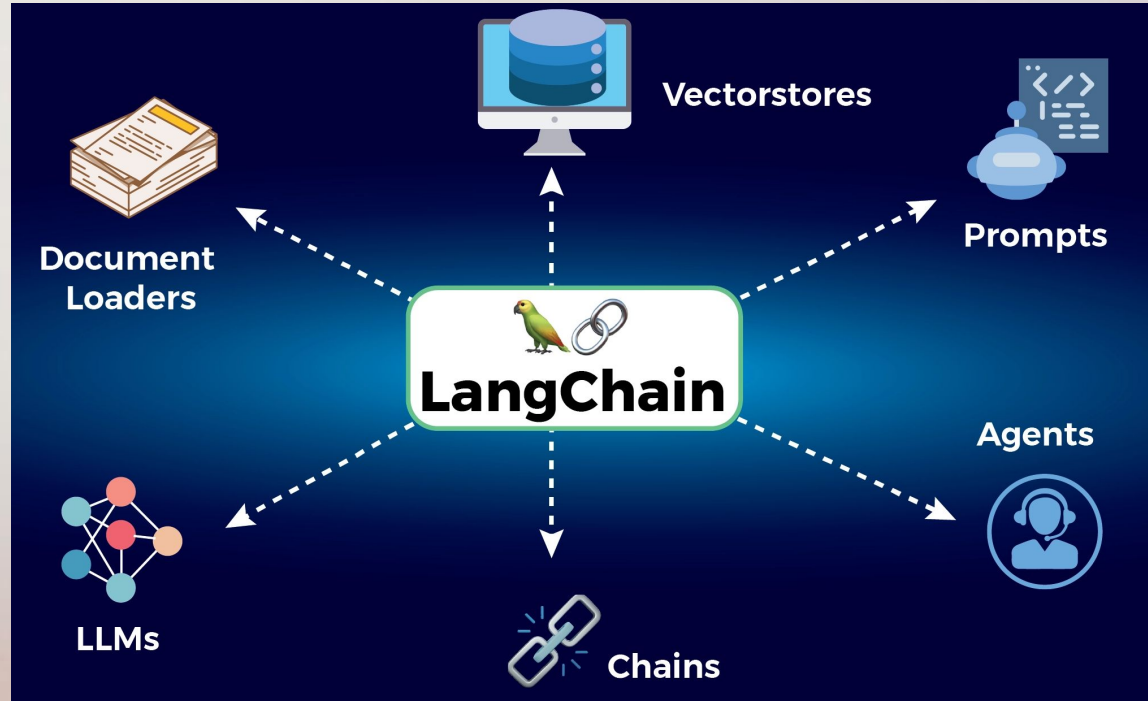
```
def suma(a, b):  
    return a + b
```

LangChain

LangChain es un Framework para desarrollar aplicaciones impulsadas por modelos de lenguaje grandes (LLM).

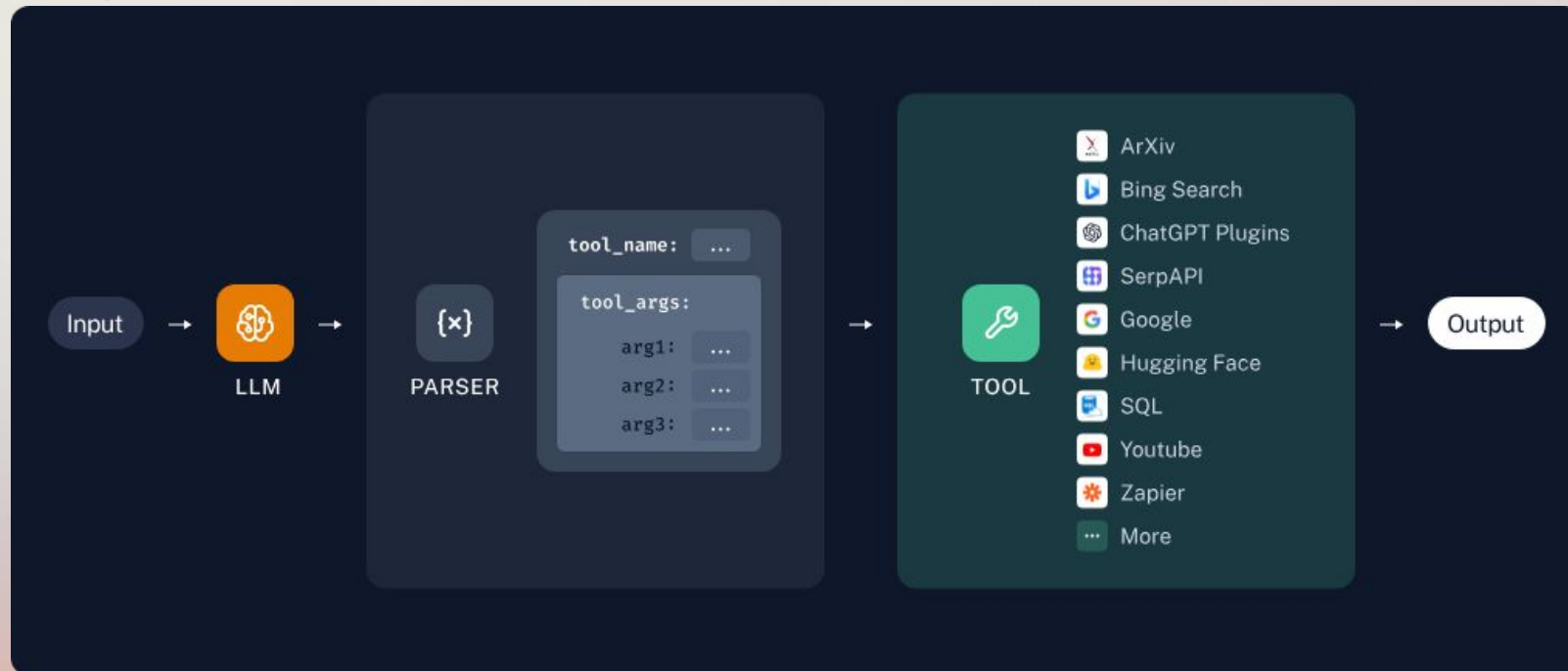
Los bloques sobre los que se base son:

- Componentes básicos
- Integraciones third-party
- Langgraph(workflows)
- LangSmith(product)



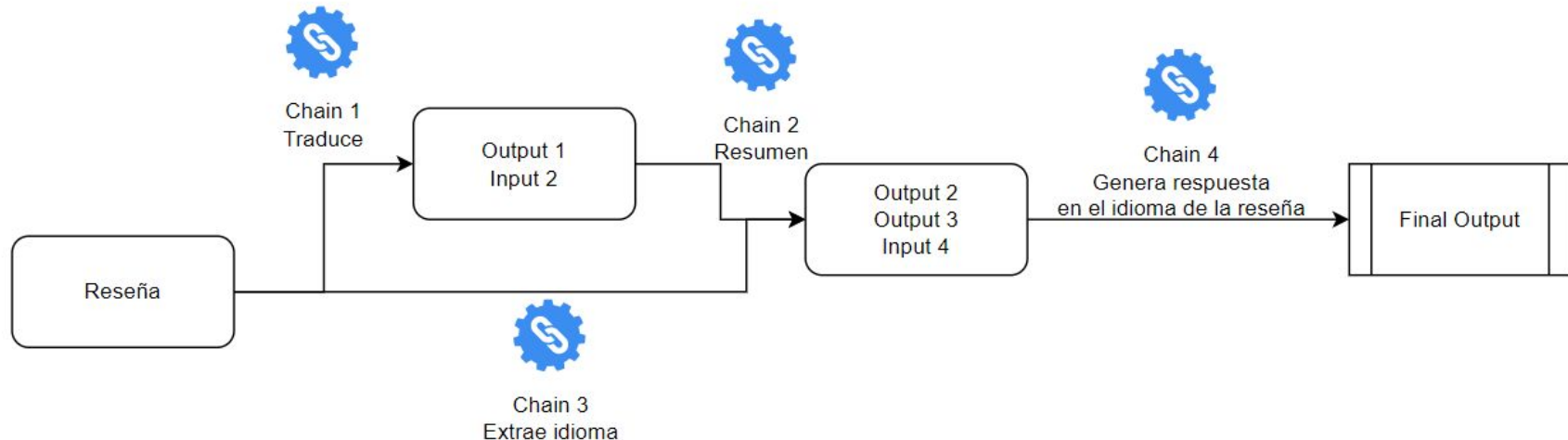
Secuencia de cadenas

Arquitectura de una secuencia de cadena



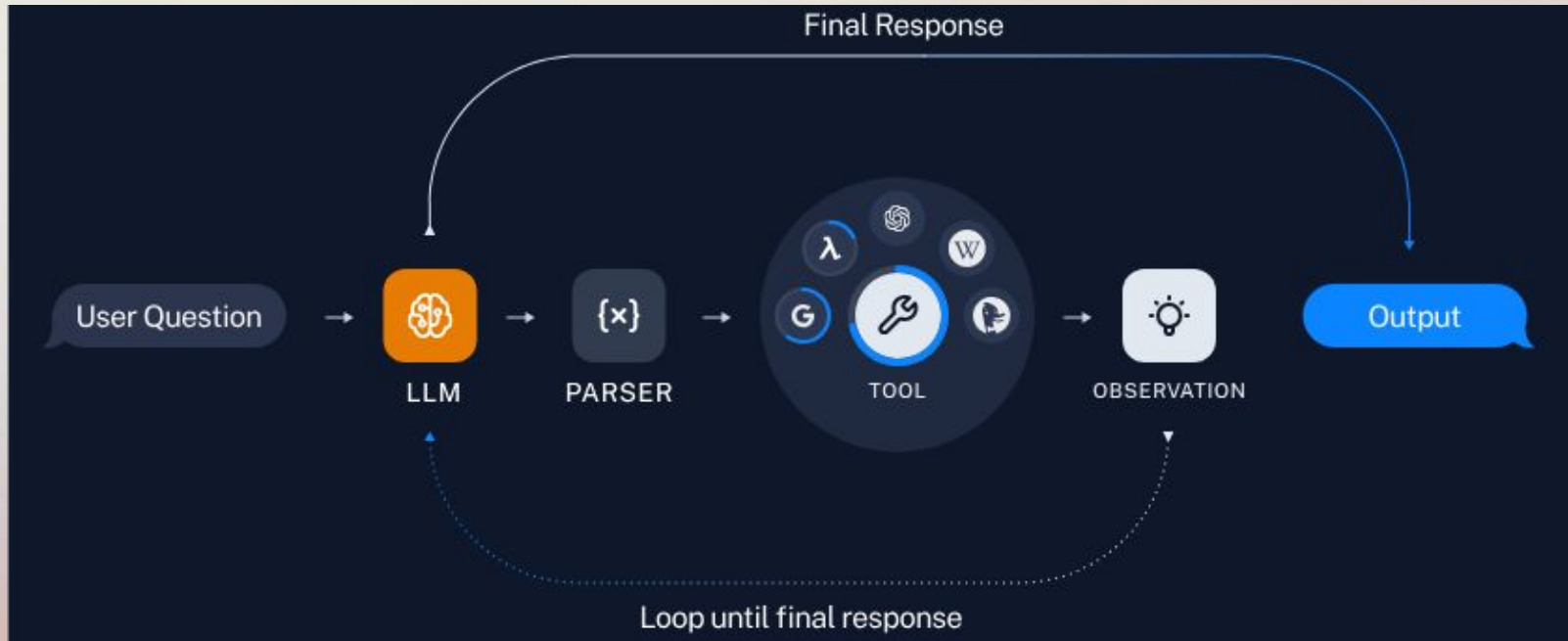
Secuencia de cadenas

- Una empresa multinacional recibe múltiples reseñas en varios idiomas de sus productos, los cuales son respondidos por un equipo de personas.
- La multinacional requiere una solución que permita respuestas automáticas en el **idioma original** y guardar un **resumen traducido**.



Agentes

Arquitectura de agentes



RAG

Retrieval Augmented Generation

- (RAG) Es una forma como usamos datos externos de información para mejorar la precisión de los Modelo de lenguaje grande(LLM)
- **Resaltemos que RAG:** no es una forma de entrenamiento o Fine-tune, sino una forma de enriquecer LLM con datos precisos

Componente principales

Vectorización
semántica(Embeddings)

Almacenamiento
Vectorial(Vector Store)

Retrieval Augmented Generation

Data connection

Source



Load

XXXXXXXXXX
XXXXXXXXXX
XXXXXXXXXX
XXXXXXXXXX
XXXXXXXXXX
XXXXXXXXXX

Transform

XXXXXXXXXX
XXXXXXXXXX

XXXXXXXXXX
XXXXXXXXXX

XXXXXXXXXX
XXXXXXXXXX

Embed

0.5, 0.2....0.1, 0.9
-0.1, 0.4....1.4, 5.9

0.2, 0.7....2.1, -1.2
4.1, 3.4....-1.5, 2.5

5.5, -0.3....0.8, 2.3
2.1, 0.1....-1.7, 0.9

Store



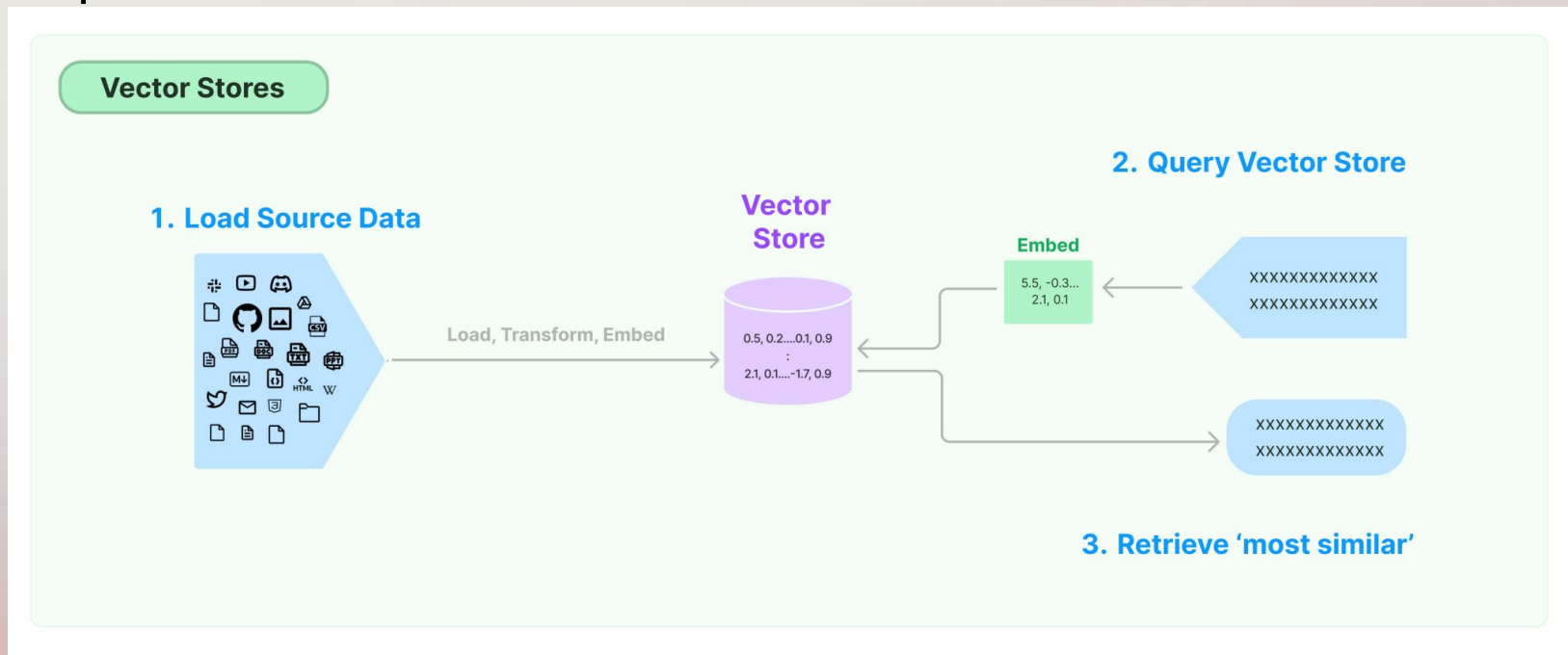
0.5, 0.2....0.1, 0.9
:
2.1, 0.1....-1.7, 0.9

Retrieve

XXXXXXXXXX
XXXXXXXXXX

Retrieval Augmented Generation

Arquitectura de RAG



Demostración

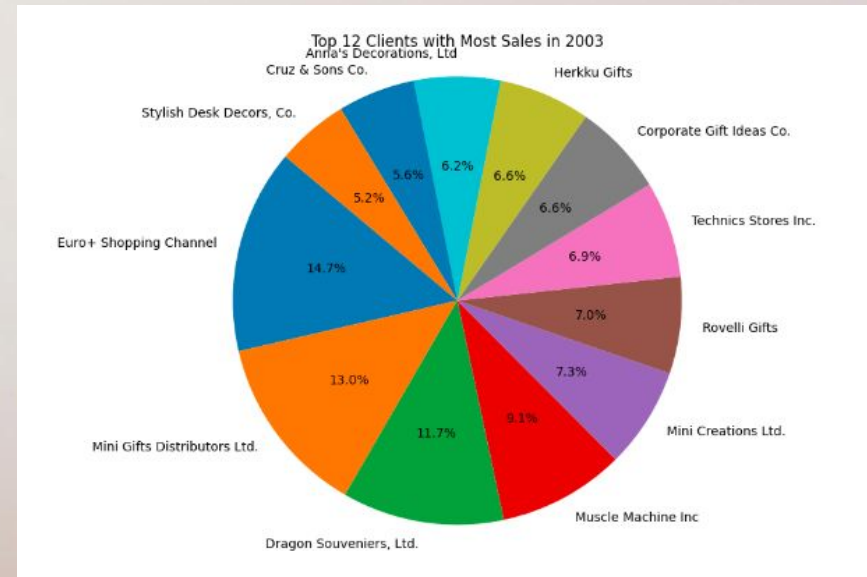
www.datapath.ai

Agente SQL

Entrada:

Gráfica un reporte en forma de torta del top 12 de cliente con más ventas en el 2003 en mi base de datos

Salida:



Gracias

Aprende, aplica y crece

www.datapath.ai