

{desafío}
latam_



Bummi AI

Del dato al diálogo: Alimentando “Chatbots” con “Data Engineering”



Leonardo Burbano

<https://www.linkedin.com/in/leoburbano>

Conociendo a la audiencia: ¿Qué tanto saben de Data Engineering?



Desafío Latam - Data Engineering y Chatbots

[Acceder a Google](#) para guardar el progreso. [Más información](#)

* Indica que la pregunta es obligatoria

¿De dónde eres? *

- ☐ México
- ☐ Centro América
- ☐ Venezuela
- ☐ Colombia
- ☐ Perú
- ☐ Bolivia
- ☐ Ecuador
- ☐ Uruguay
- ☐ Paraguay
- ☐ Argentina
- ☐ Chile
- ☐ Otro

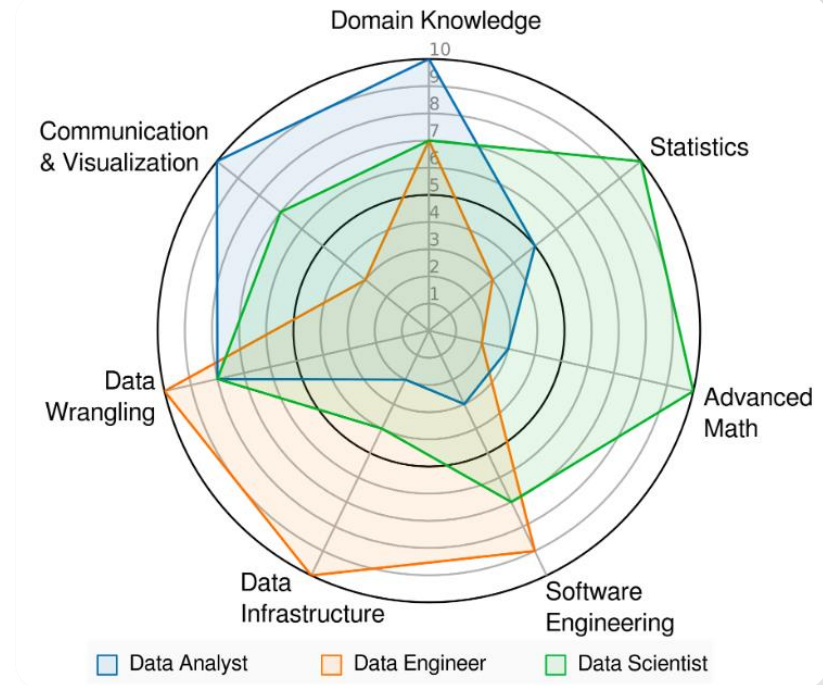
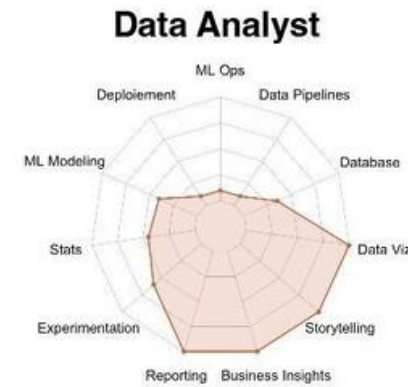
¿Cuáles son las habilidades de un Data Engineer? *

- ☐ Software Avanzado y SQL
- ☐ Python

https://docs.google.com/forms/d/e/1FAIpQLSeBCaYsmXntkBb_S2noikz19EEDBpZDQRVHZjjURM4rjE0QHA/viewform

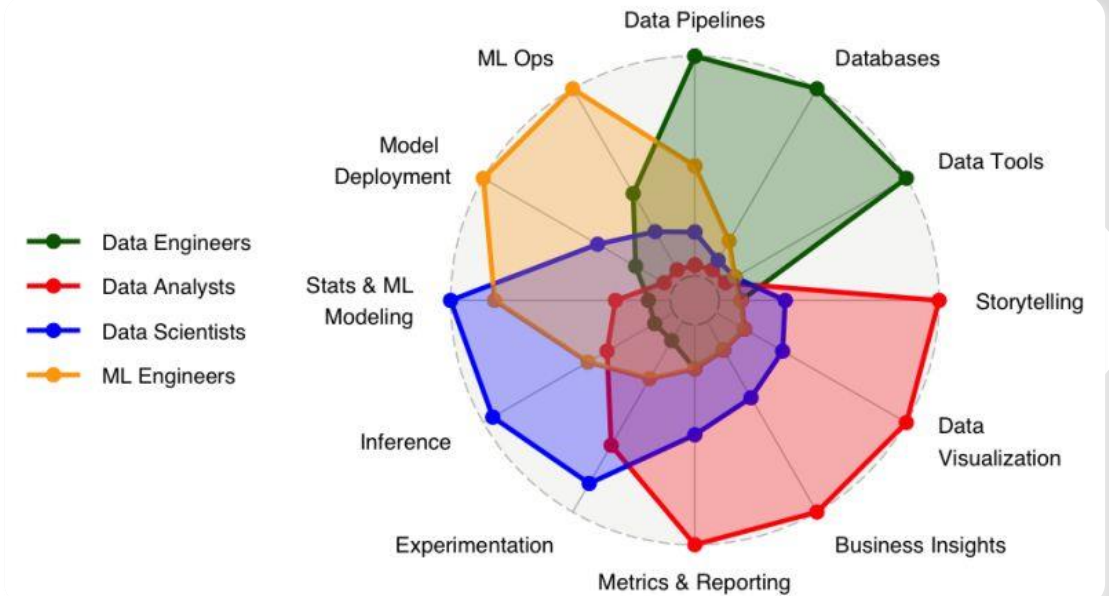
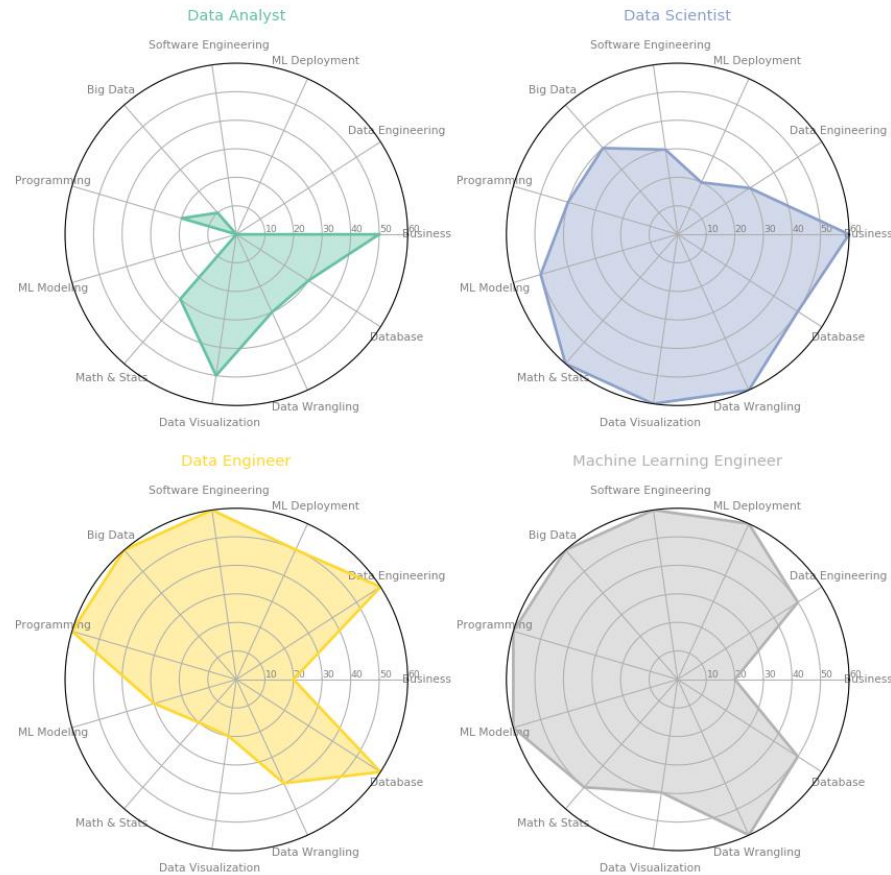


¿Qué diferencia hay entre un Data Engineer y un Data Scientist?





¿Qué diferencia hay entre un Data Engineer y un Data Scientist?





La demanda ha crecido...




[Find a Hire](#) [Tools](#) [About Us](#) [Success Stories](#) [Pricing](#) [Services](#) [Resources](#)

[APPLY FOR JOBS](#)

[START HIRING](#)

Potential savings over US salaries: 42.73%

Outline

 10 MINUTE READ

Business Intelligence (BI) Analyst

Data Analyst

Data Engineer




Data Engineering Manager

Data Lead

Data Scientist

Director of Data

Return to the 2024 Salary Guide: US vs. Latin America



Data Engineer

Data engineers are responsible for designing, building, and maintaining the technical infrastructure necessary for the organization to collect, store, and analyze data. This role requires a deep understanding of databases, data warehousing, and data processing technologies. Data engineers must be able to work with large amounts of data, design and implement efficient data storage solutions, and ensure that data is easily accessible to other teams within the organization.

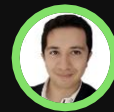
LatAm annual salary range: \$42,000–\$84,000

Potential savings over US salaries: 52.42%

Data Engineering Manager

A data engineering manager oversees a team of data engineers,

<https://www.hirewithnear.com/blog/data-roles-salary-guide-us-vs-latin-america>



Objetivos del día de hoy

1. Del diálogo al dato: Tenemos Chatbots y ahora ¿qué sigue?
2. ¿Qué hace la ingeniería de datos?

{desafío}
latam_

 **Bummi AI**

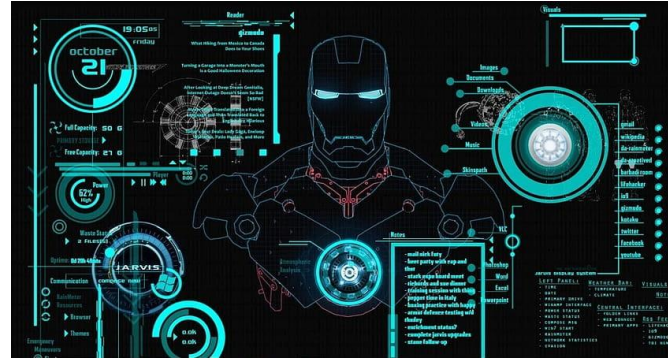
<https://www.linkedin.com/in/leoburbano>



¿Cuál fue el primer Chatbot que viste?



¿Cuál fue el primer Chatbot que viste?



Escucha y contesta en **tiempo real**

Acceso a **Internet**

Accede a las **bases de datos**

Coherencia, privacidad y seguridad

Visión artificial y realidad aumentada

Modifica lo que ve el user

Personalidad, lealtad y un objetivo



¿Cuál fue el primer Chatbot que viste?



Contacto **físico** y sensibilidad

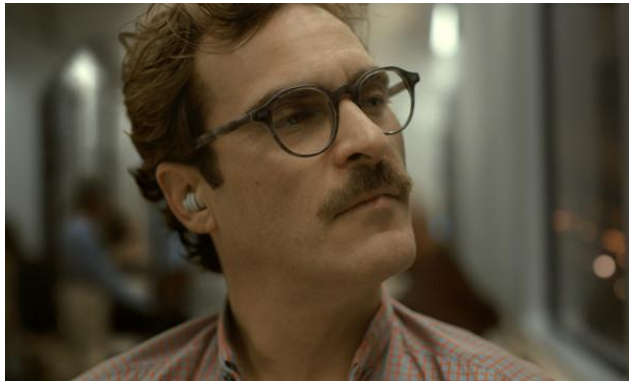
Evaluación en tiempo real de **signos vitales**

Siempre alerta, detecta **dolor** y **apoya**

En salud, no es viable fallar



¿Cuál fue el primer Chatbot que viste?



Emociones y aprendizaje

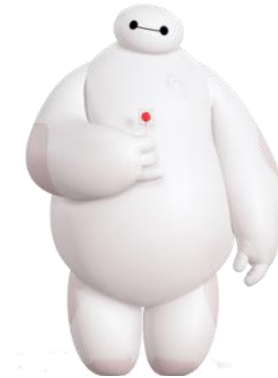
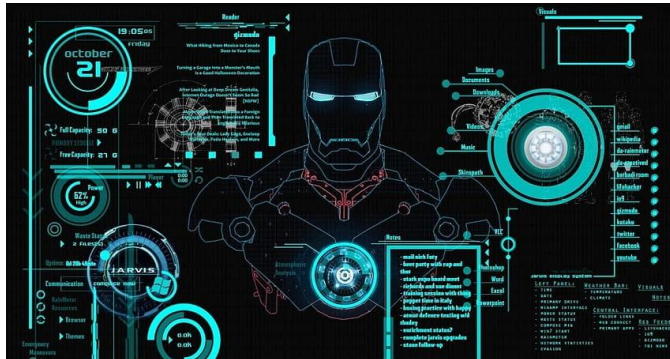
Deseos y sueños

Miedos, angustia y propias decisiones

Es un **testigo**, la definición de copiloto



¿Cuál fue el primer Chatbot que viste?



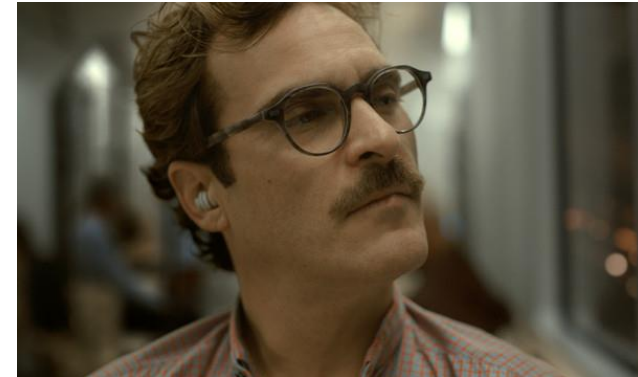
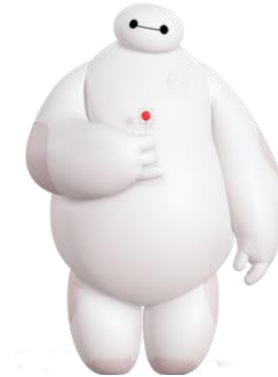
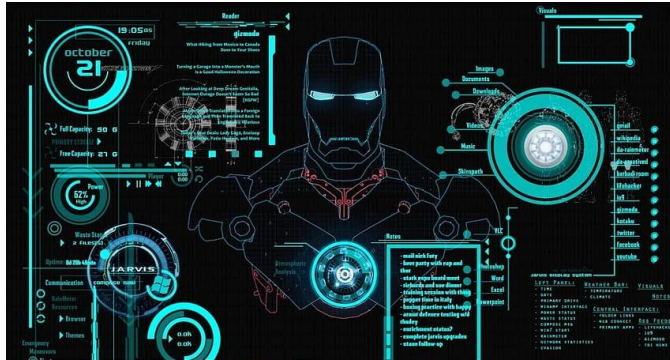
Autonomía, coherencia e improvisación

Acceso a información coherente en tiempo record

Se perciben como humanos



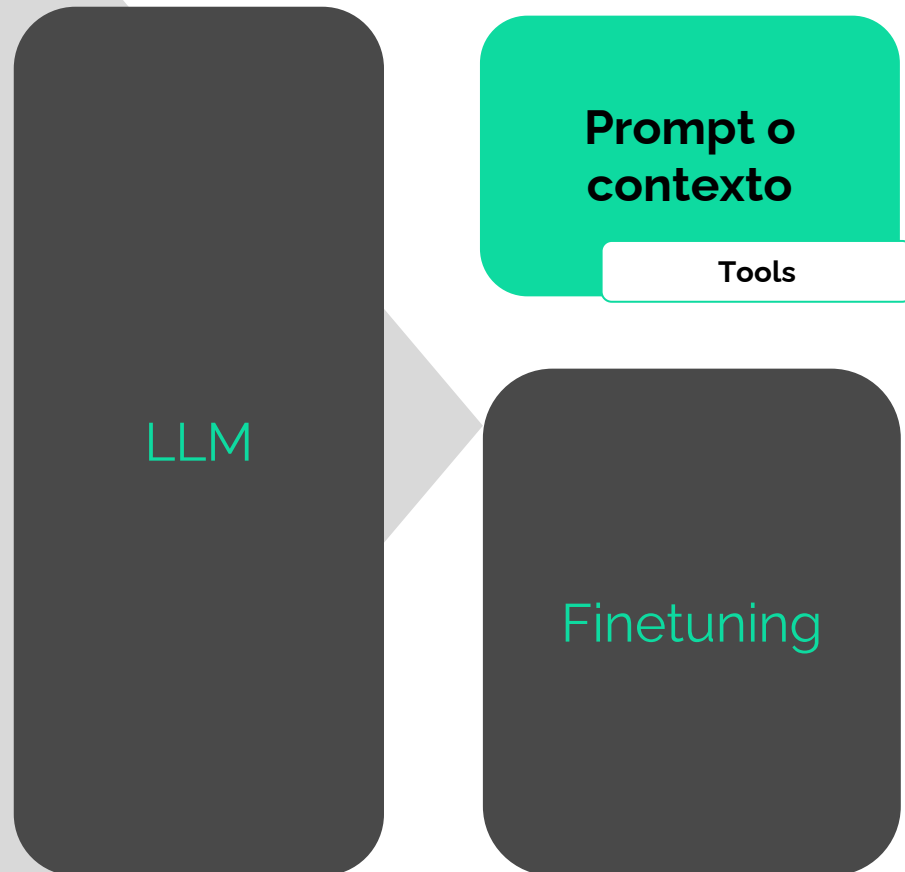
¿Cuál fue el primer Chatbot que viste?



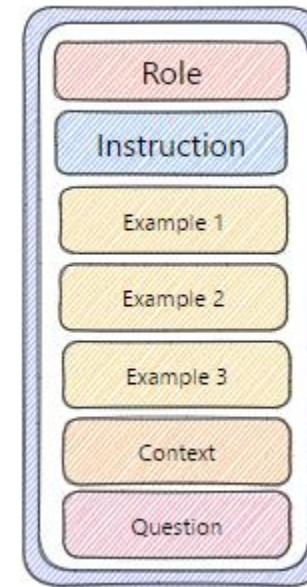
Acceso a información coherente en tiempo record



¿Qué podemos hacer hoy con un chatbot?



```
{"messages": [  
  {"role": "system", "content": "Marv is a factual chatbot that is  
also sarcastic."},  
  {"role": "user", "content": "What's the capital of France?"},  
  {"role": "assistant", "content": "Paris, as if everyone doesn't  
know that already."}  
]}
```





¿Cuál es el problema?

Actualizar el modelo

Facilitar el acceso a información
externa y privada

Detectar y evitar las alucinaciones

Dar seguimiento

Coherencia vs. improvisación

Ingeniería de
datos

Machine Learning

Software
Engineering

Infraestructura y DevOps



Gran parte de la solución

Internet

Información interna

Audio, video, PDFs

BDDs

Feedback humano

Ingeniería de
datos

Limpieza y
depuración

Transformación
y seguimiento

Organización

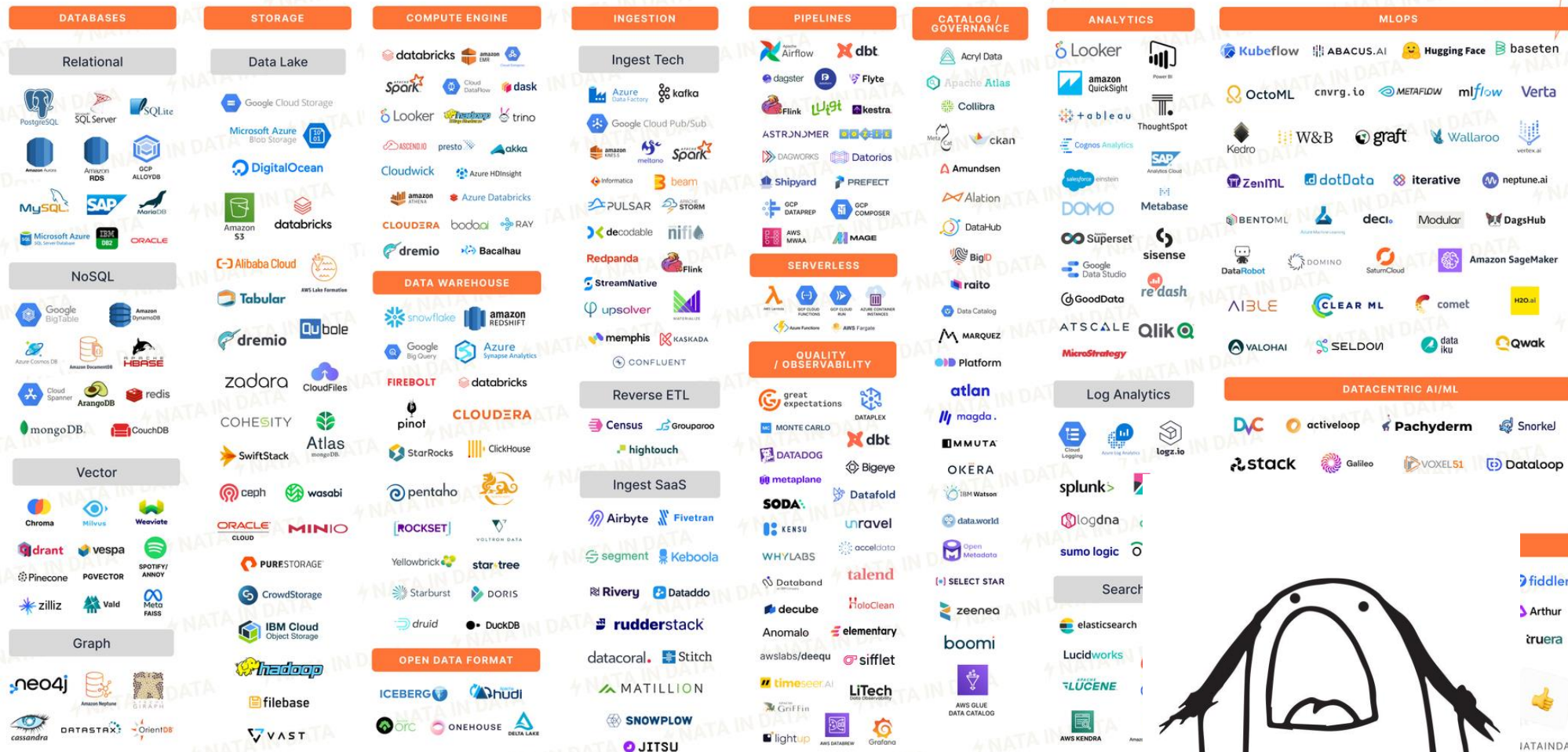
Almacenamiento

Calidad y
orquestración



Gran parte de la solución

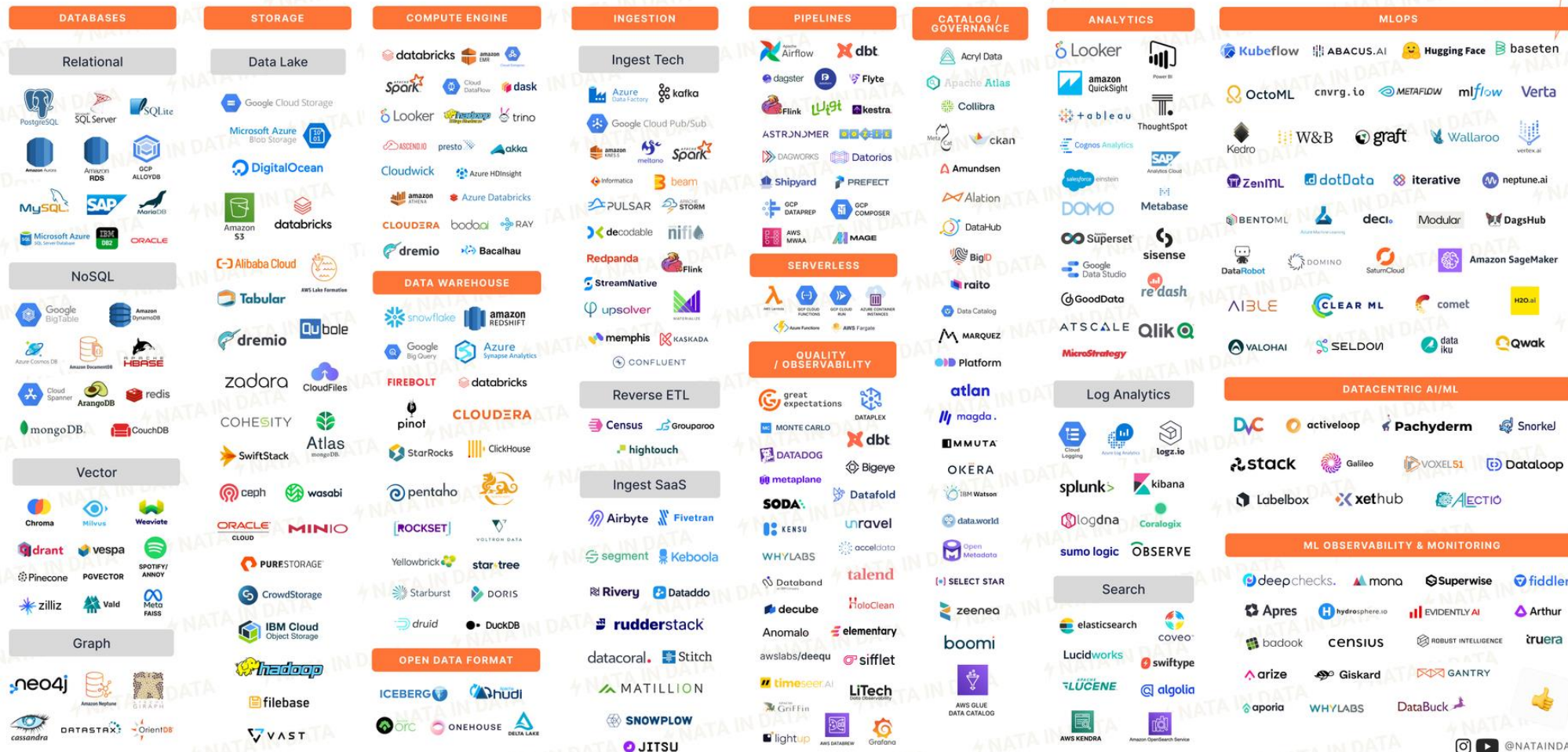
⚡ NATA IN DATA DATA TOOLS LANDSCAPE 2024





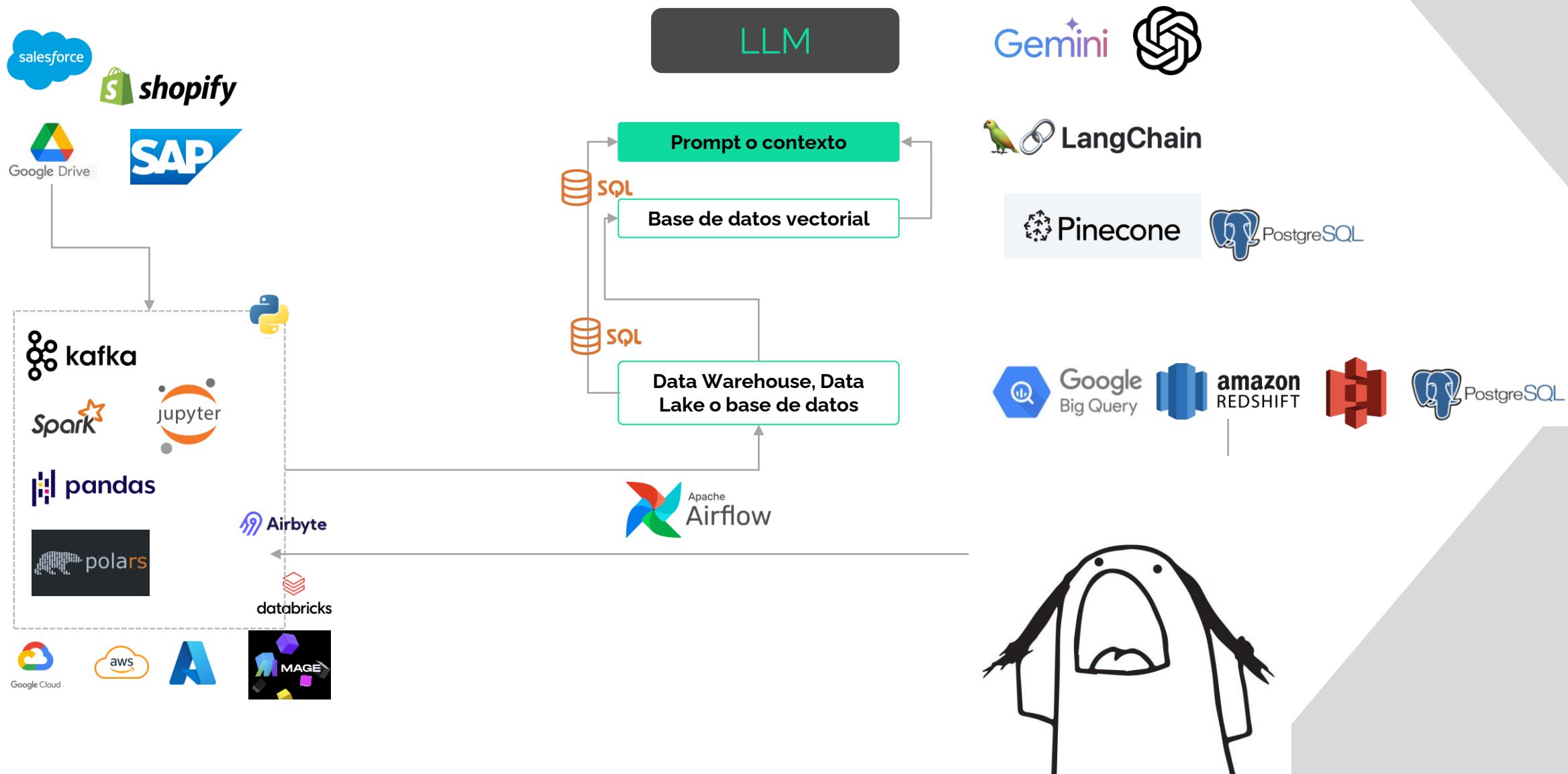
Gran parte de la solución

⚡ NATA IN DATA DATA TOOLS LANDSCAPE 2024



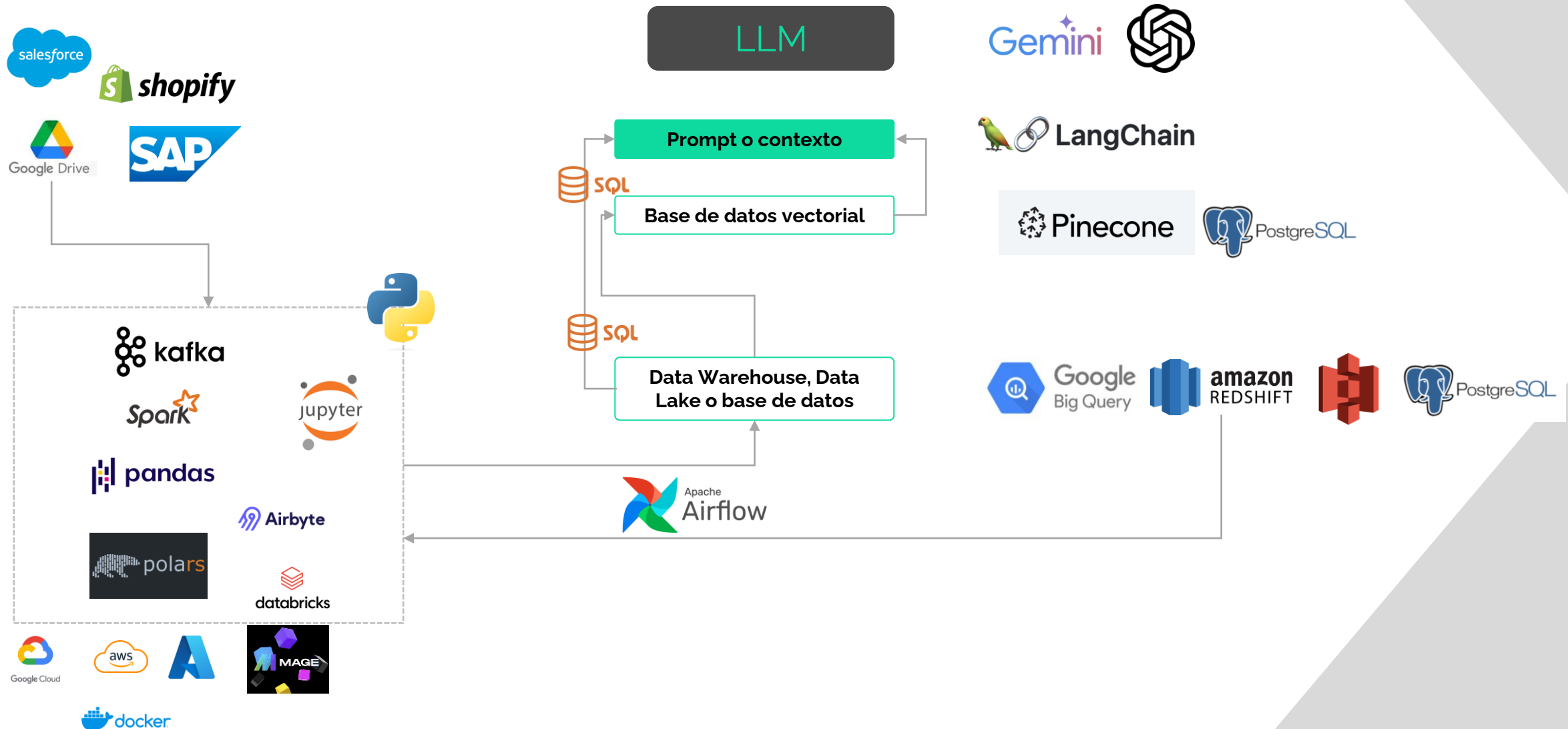


Un ejemplo de la arquitectura para mejorar e integrarnos






Un ejemplo de la arquitectura para mejorar e integrarnos







[DAGs](#)
[Security](#)
[Browse](#)
[Admin](#)
[Docs](#)
10:46 PDT (-07:00)
FB

DAG: example_bash_operator
success
schedule: 0 0 * * *

Tree
Graph
Calendar
Task Duration
Task Tries
Landing Times
Gantt
Details
Code

2021-06-02T09:27:27-07:00

Runs

25

Run

manual__2021-06-02T16:27:26.797940+00:00

Layout

Left > Right

Find Task...

Update

Python 3

Logout

Operator

queued

running

success

failed

up_for_retry

up_for_reschedule

upstream_failed

skipped

scheduled

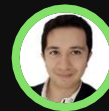
no_status

Auto-refresh

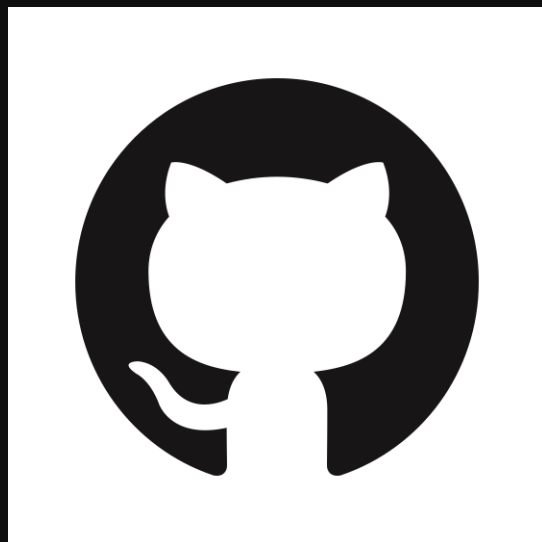
Refresh

```

graph LR
    runme_0[runme_0] --> also_run_this[also_run_this]
    runme_1[runme_1] --> run_after_loop[run_after_loop]
    runme_2[runme_2] --> this_will_skip[this_will_skip]
    also_run_this --> run_this_last[run_this_last]
    run_after_loop --> run_this_last
    this_will_skip --> run_this_last
    
```



Veamos algunos ejemplos en código



Solicitar repositorio por LinkedIn:

<https://www.linkedin.com/in/leoburbano>



Veamos algunos ejemplos en código

The screenshot shows a web application running on localhost8501/chatbot_with_vector_db. The interface has a sidebar on the left with the following links: Home, chatbot, analytics, chatbot with sql, and chatbot with vector db (which is highlighted). The main chat area displays a conversation with a chatbot. The chatbot's message is a long paragraph about startups in the US and other locations. The user's input is "List the total startups per location?". The chatbot's response is a list of locations and their startup counts. Below the list, there is a note: "Please note that these numbers represent the total startups listed in the database for each location." The user's next input is "¿Con quién se fusionó Medismart?". The chatbot's response is "Medismart se fusionó con Teledoc." At the bottom of the chat area, there is a text input field labeled "Consulta desde SQL..." with a right arrow button.

Deploy

Home
chatbot
analytics
chatbot with sql
chatbot with vector db

chatbot

List the total startups per location?

Here are the total startups per location:

- San Francisco, CA, USA: 1341
- New York, NY, USA: 402
- Mountain View, CA, USA: 83
- Toronto, ON, Canada: 74
- Bangalore, KA, India: 115
- London, England, United Kingdom: 116
- Singapore, Singapore: 49
- Los Angeles, CA, USA: 112
- Palo Alto, CA, USA: 93
- Austin, TX, USA: 52

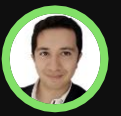
Please note that these numbers represent the total startups listed in the database for each location.

¿Con quién se fusionó Medismart?

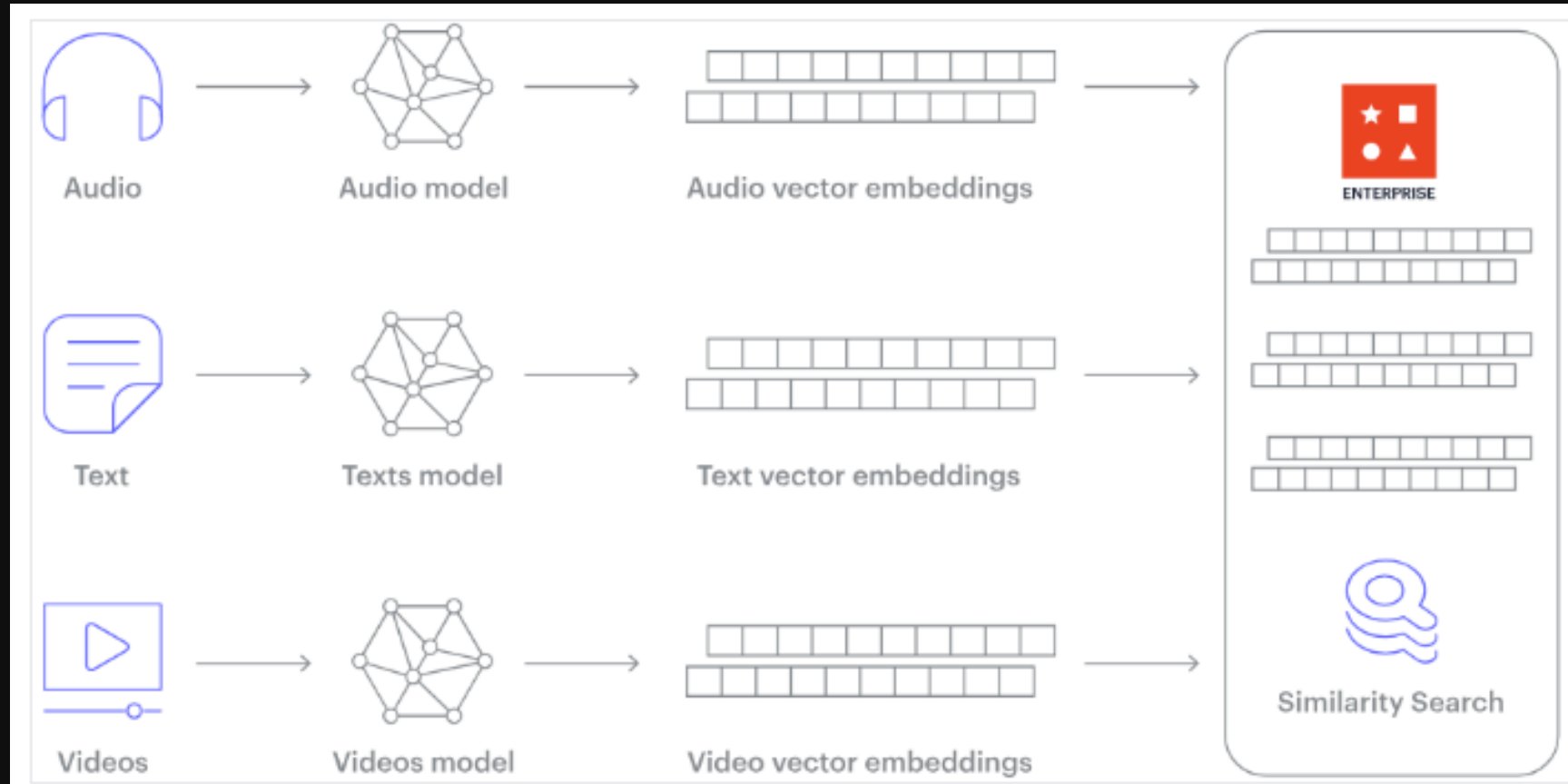
Medismart se fusionó con Teledoc.

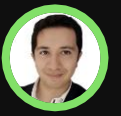
Consulta desde SQL...

¿Cómo el chatbot se vuelve más coherente al integrarse con herramientas e ingeniería de datos?



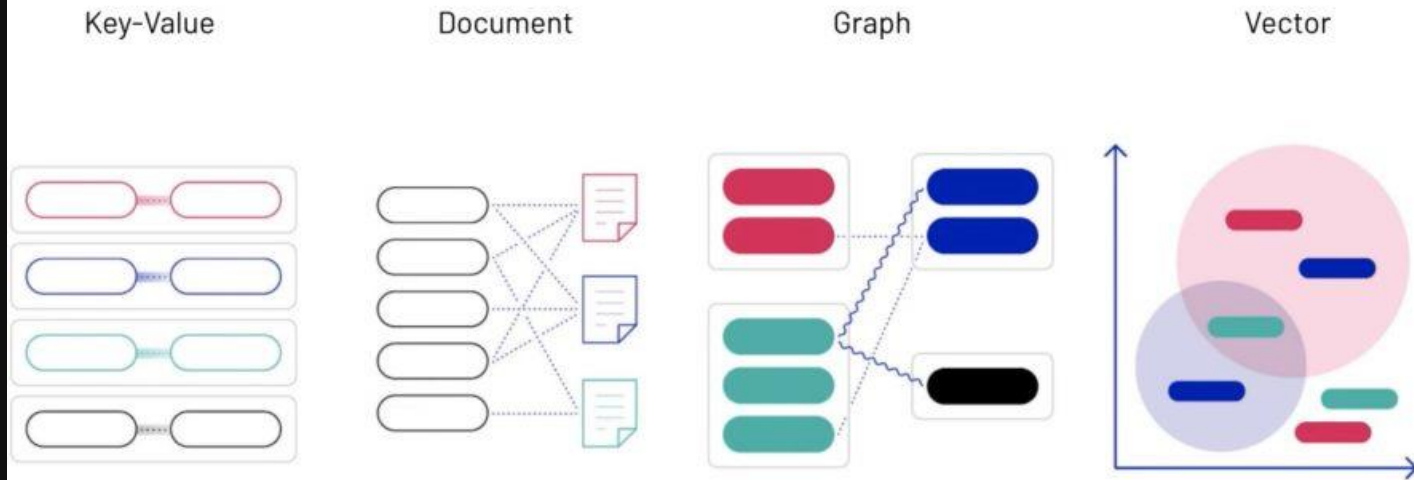
Cada vez hay más opciones de bases de datos vectoriales

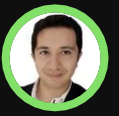




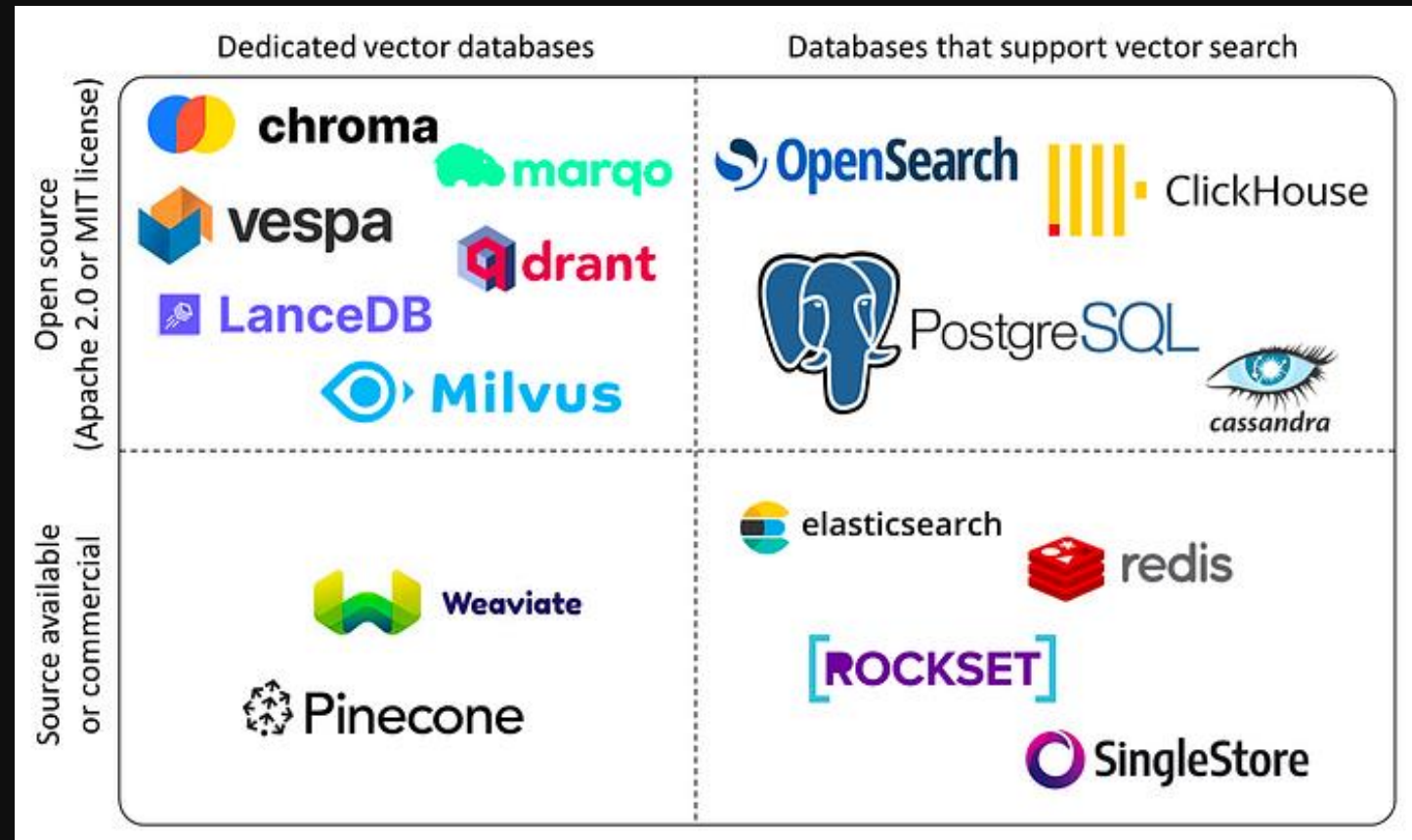
Cada vez hay más opciones de bases de datos vectoriales

Vectors need a new kind of database





Cada vez hay más opciones de bases de datos vectoriales





Mi path recomendado...

Data Analyst

Databricks

Grafana

Modelado de
datos

SQL Intermedio

Python [PySpark]

Airbyte o dbt

Calidad de datos

Python [Pandas]

Machine Learning
y NLP

Spark [Avanzado]

TDD

Airflow

Kafka

Vertex AI o
Sagemaker

CDC

Cloud: GCP o
Azure o AWS

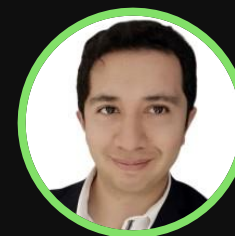
CI/CD y Git

Backend

Profundizar en la
nube y Apache

{desafío}
latam_

 Bummi AI



¿Preguntas?

 Leonardo Burbano

<https://www.linkedin.com/in/leoburbano>

{desafío}
latam_

 Bummi AI



Gracias!

 Leonardo Burbano

<https://www.linkedin.com/in/leoburbano>