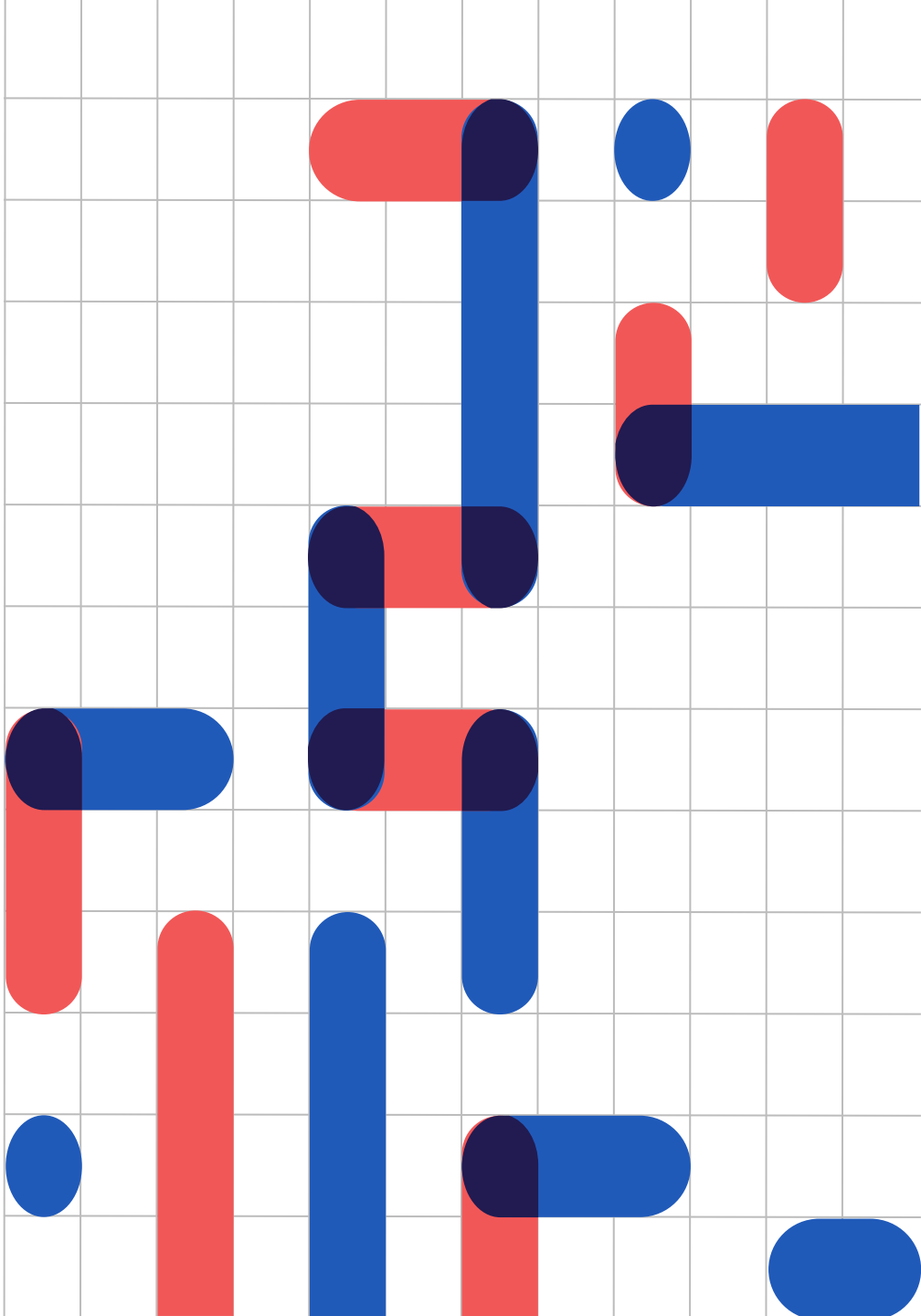
Agentes IA: el futuro de la interacción con sistemas y datos

PyData Granada

Mariia Chizhikova



Me presento

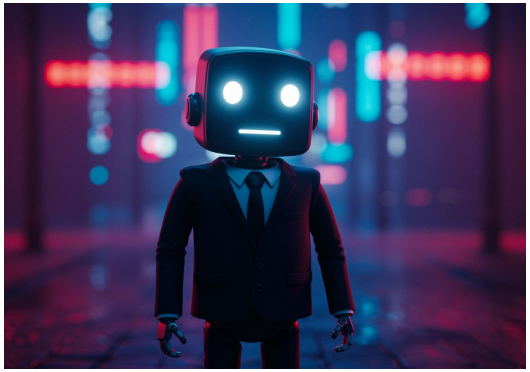
↘

- Graduada en Filología Hispánica
- Máster en Tecnologías del Lenguaje
- Ingeniera de modelos de lenguaje
- Divulgadora inteligencia artificial
- Podcaster: Naturalmente IA

@mariia.en.ia

AutonomIA

Agente... ¿especial? ¿inmobiliario? ¿de seguros? ¿de viajes?



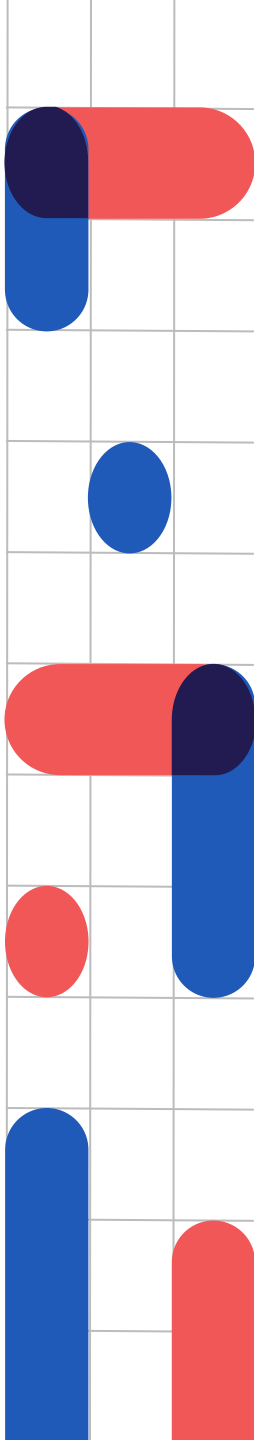
¿Qué es un agente IA?

Los agentes IA son sistemas autónomos que pueden:

- interactuar con su entorno
- recopilar y analizar datos
- ejecutar acciones para conseguir objetivos predefinidos

Componentes esenciales

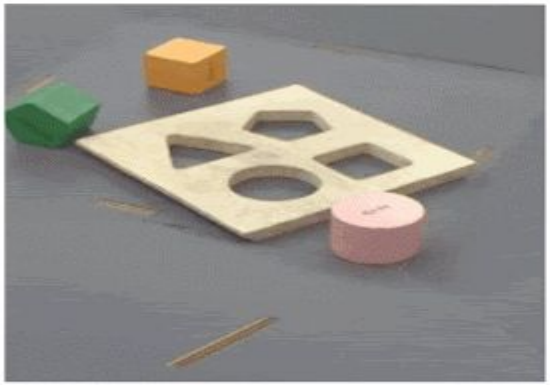
- **Entorno**
  - Físico (robótica)
  - Digital (software)
- **Sistemas de percepción**
  - Físicos (cámaras, LiDAR, micrófonos)
  - Digitales (endpoints de API, llamadas a BBDD)
- **Motor de toma de decisiones**
  - Sistemas basadas en reglas (termostato)
  - Modelos de ML/DL
  - Aprendizaje por refuerzo
- **Mecanismos de acción**
  - Físicos (motores, brazos robóticos)
  - Digitales (llamadas de API)

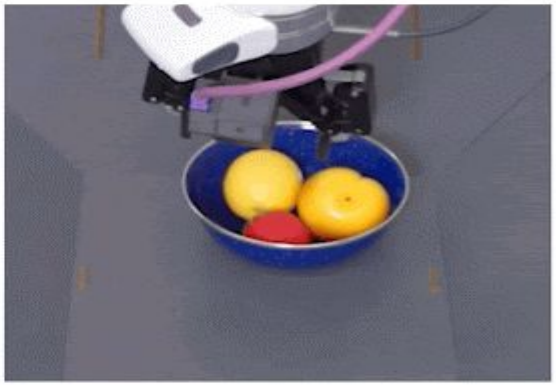


Ejemplos

- Chatbots
- AlphaGo y otros agentes “jugadores”
- Agentes robóticos

Robocat de Google DeepMind (2023)

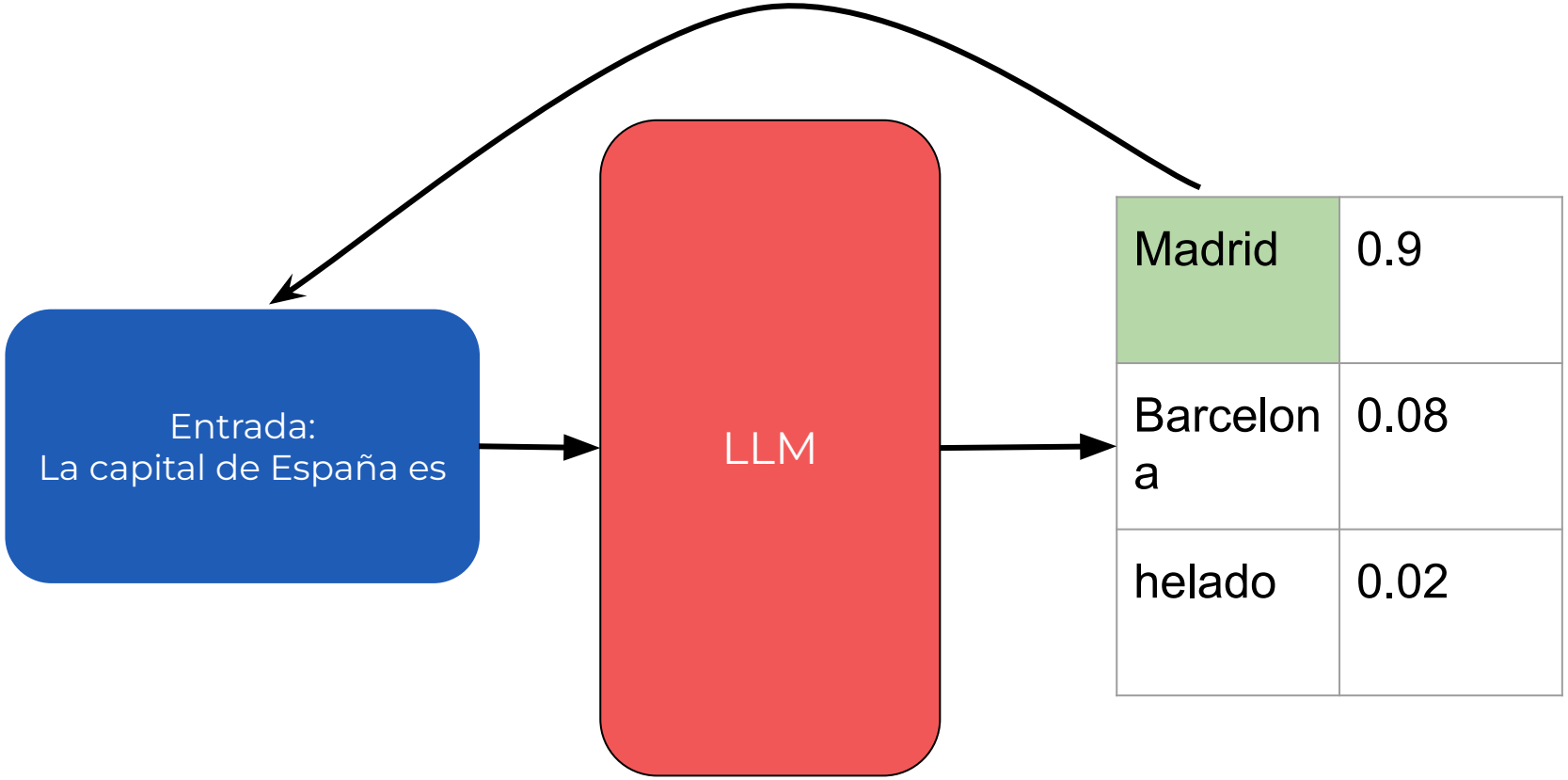




El maravilloso mundo de los LLMs

Más allá de la generación del texto

¿Cómo funcionan?



El “razonamiento” de los LLMs

- Chain of thought - *Cadena de pensamiento*

Entrada al modelo

Q: Juan tiene 5 pelotas de tenis Va a comprar 2 paquetes de pelotas más. Cada paquete lleva 3 pelotas. ¿Cuántas pelotas de tenis tiene ahora?
A: La respuesta es 11
Q: En la cafetería había 23 manzanas. SI han usado 20 para hacer la comida y compraron 6 más, cuántas manzanas tienen ahora?

Salida

A: La respuesta es 17

Entrada al modelo

Q: Juan tiene 5 pelotas de tenis Va a comprar 2 paquetes de pelotas más. Cada paquete lleva 3 pelotas. ¿Cuántas pelotas de tenis tiene ahora?
A: Roger empezó teniendo 5 pelotas. 2 paquetes de 3 suman 6. 5+6=11 La respuesta es 11
Q: En la cafetería había 23 manzanas. SI han usado 20 para hacer la comida y compraron 6 más, cuántas manzanas tienen ahora?

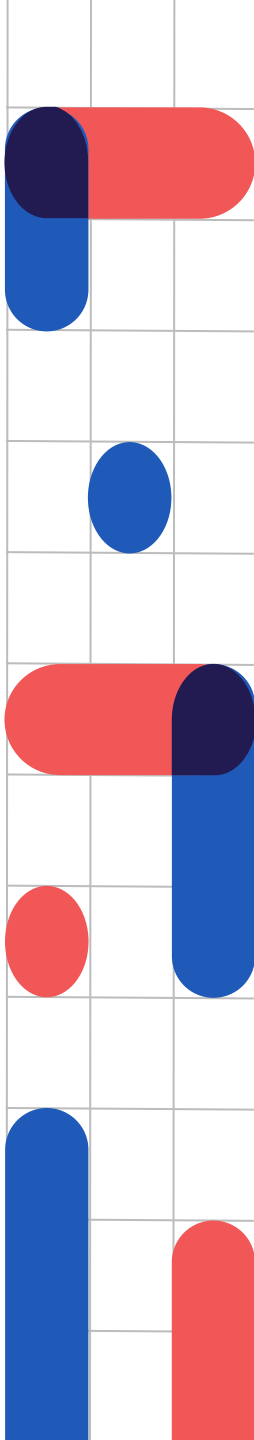
Salida

A: La cafetería tenía 23 manzanas. Usaron 20: 23-20=3. Compraron 6 más: 3+6=9. La respuesta es 9

Es una técnica útil para

Todo.

... o casi todo.

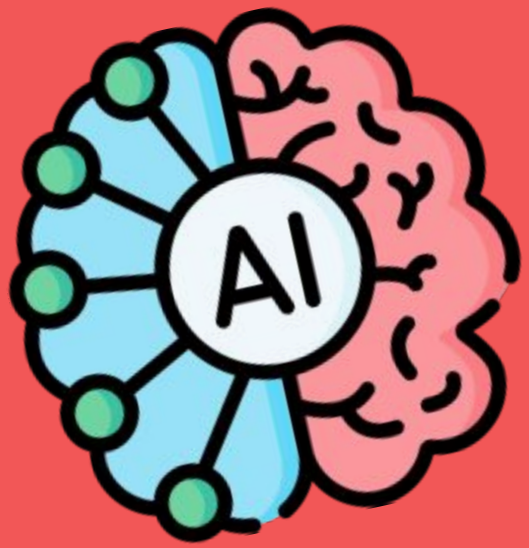


Agentes LLM

~~No~~ eres lo que dices, eres lo que haces

Partes esenciales

El entorno

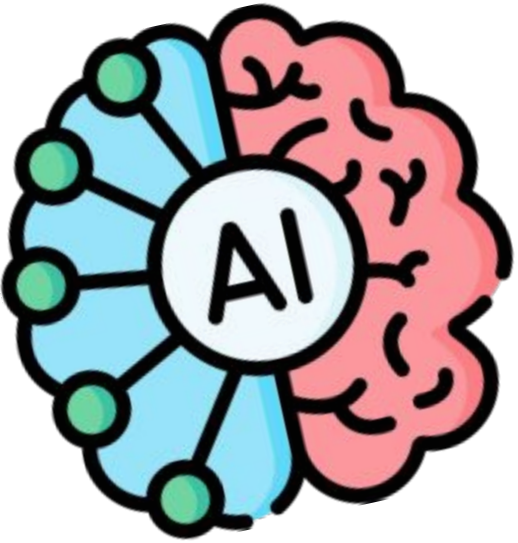


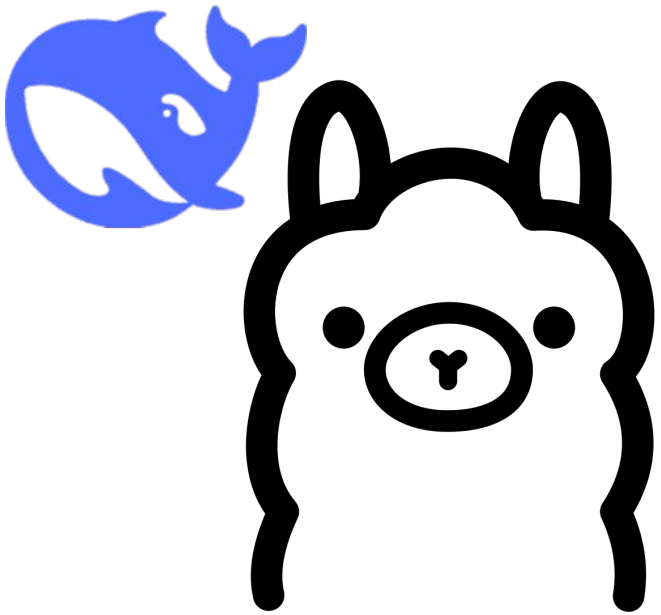
El cerebro

Las herramientas

El cerebro

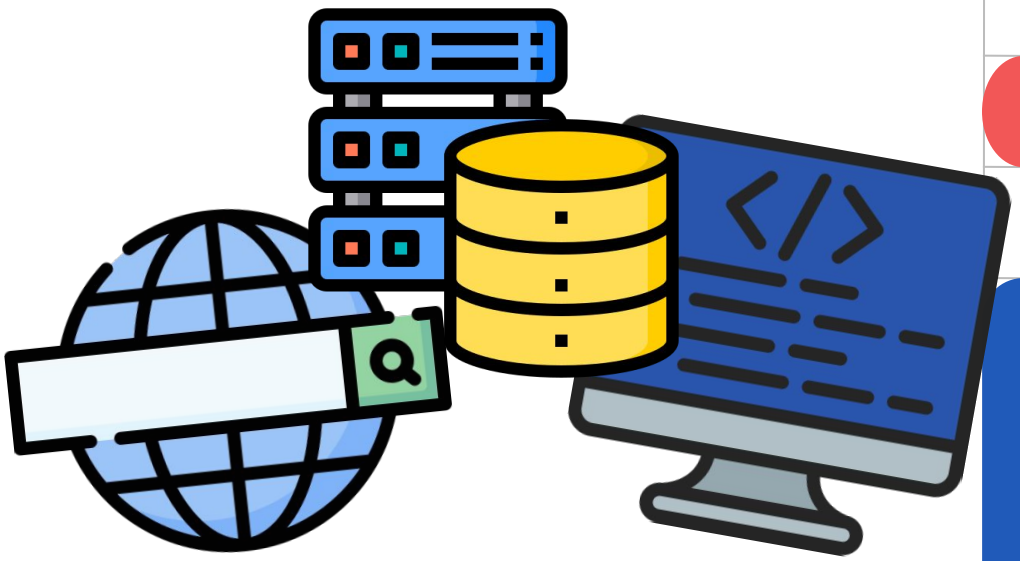
- un modelo de lenguaje
- interpreta las instrucciones
- se encarga del razonamiento y la planificación
- genera la respuesta final



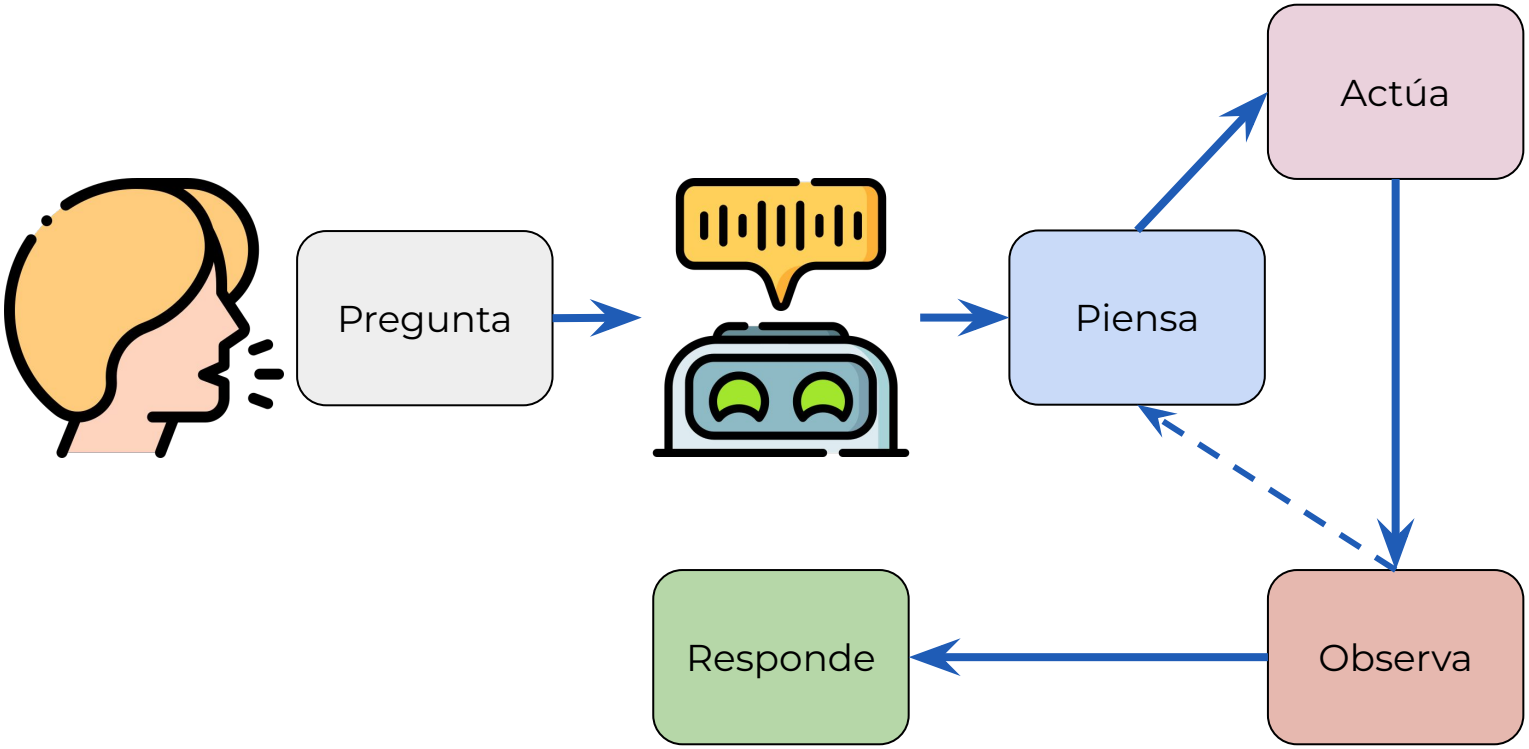


Las herramientas

- son los **cacharros** con los que permitimos que juegue el LLMs
- una herramienta tiene que tener un **objetivo claro y bien definido**
- nos ayudan a sobrepasar las limitaciones del LLM



Pensar > Actuar > Observar



De la teoría a la práctica

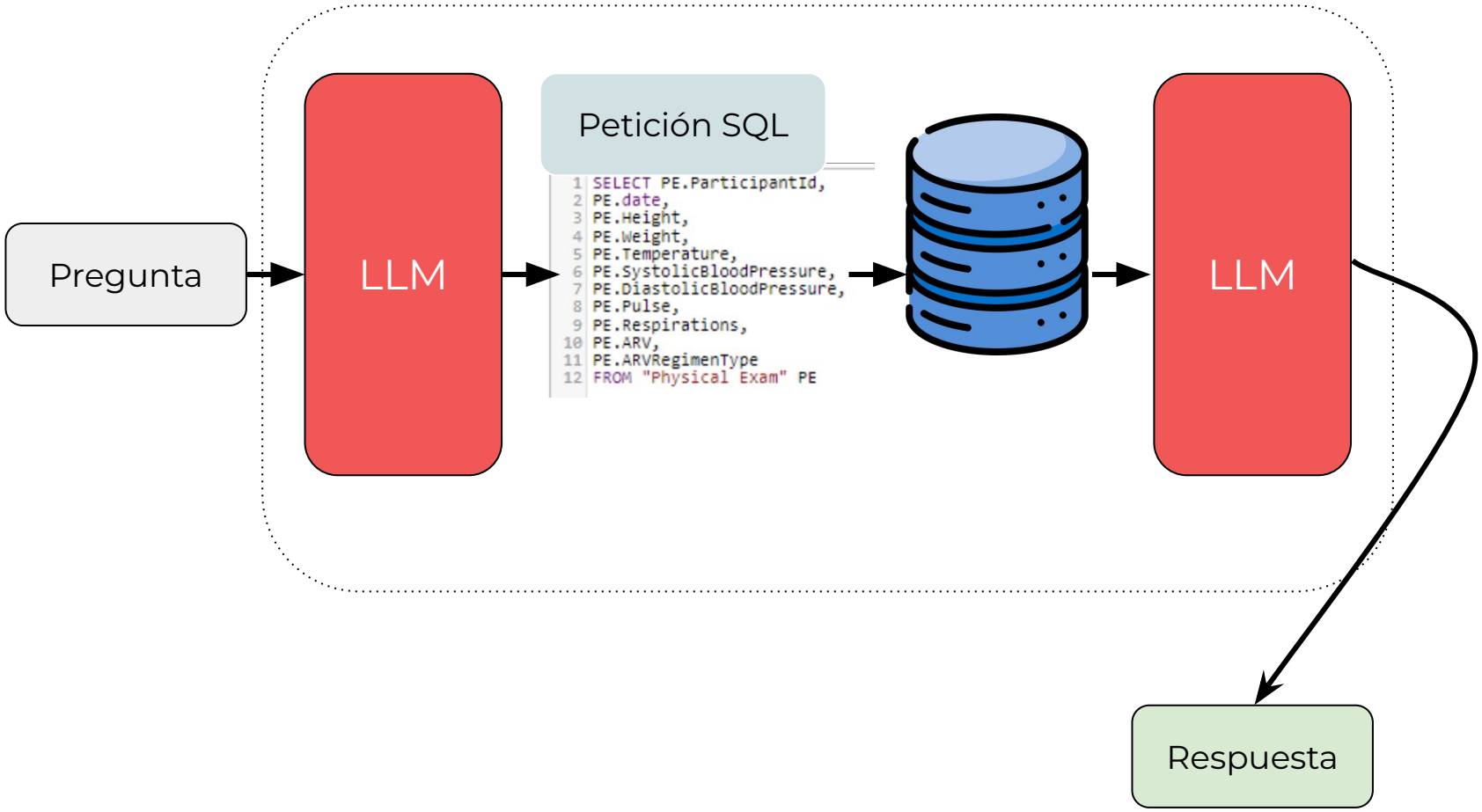
Más de 69% de los developers usan SQL

Imagina cuántos usuarios interactúan con bases de datos SQL sin saberlo

¿Te imaginas pedir info a tu BBDD sin usar SQL?

Y no hablo de pasarte a MongoDB :)

¿Cómo funciona un AgenteSQL?



Momento DEMO

¿Qué se esconde debajo de la capucha?

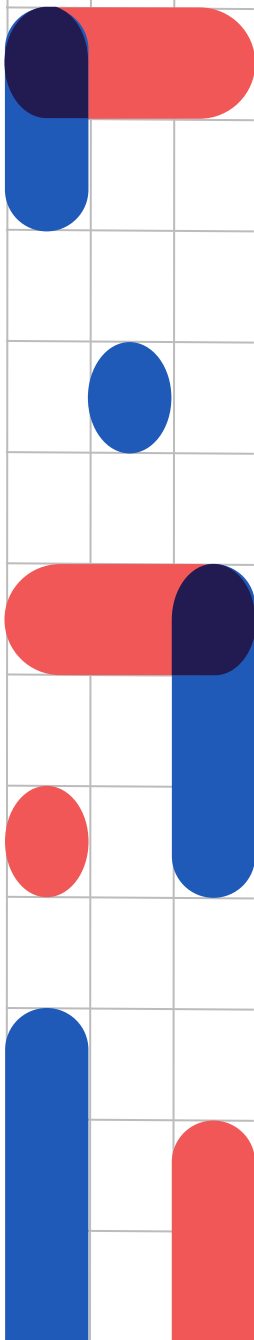
REACT: SYNERGIZING REASONING AND ACTING IN LANGUAGE MODELS

Shunyu Yao[*,1], Jeffrey Zhao[2], Dian Yu[2], Nan Du[2], Izhak Shafran[2], Karthik Narasimhan[1], Yuan Cao[2]

[1]Department of Computer Science, Princeton University
[2]Google Research, Brain team
[1]{shunyuy,karthikn}@princeton.edu
[2]{jeffreyzhao,dianyu,dunan,izhak,yuancao}@google.com

Idea principal de ReACT

- ReACT permite al modelo intercalar pasos de **razonamiento** y **acción**, creando un proceso de resolución de problemas más dinámico e interpretable
  - **Razonar para actuar**: El modelo puede descomponer objetivos complejos en subobjetivos más simples, planificar secuencias de acciones y ajustar sus planes en función de la información obtenida del entorno
  - **Actuar para razonar**: El modelo puede interactuar con la base de datos para obtener información adicional que pueda utilizar en su proceso de razonamiento



Cosas importantes a tener en cuenta

Transparencia

Siempre tenemos que saber qué código se ejecuta y tener nuestras copias de seguridad por si el modelo la lía

Privacidad

Pensamos 100 veces antes de usar una API externa sobre bases de datos sensibles

Sesgos

La respuesta final puede ser afectada por los sesgos del LLM que estamos usando

Agentes IA: el presente de la interacción con sistemas y datos

PyData Granada

Mariia Chizhikova

NaturalmenteIA

@mariia.en.ia