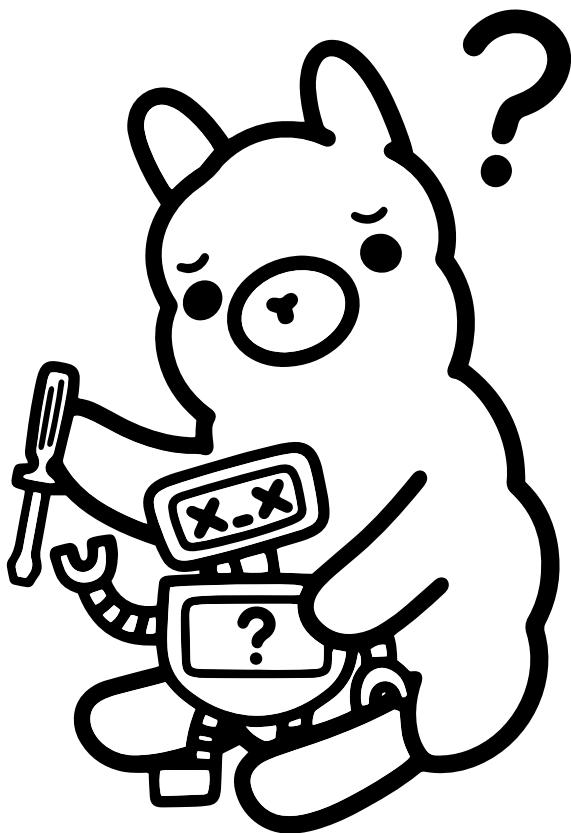


Ollama

Ollama es una plataforma de modelos de lenguaje grande (LLM) que te permite ejecutar y gestionar modelos de lenguaje localmente.

Se maneja a través de la línea de comandos (CLI) pero puedes integrarlo con diferentes interfaces gráficas y servidores web para facilitar su uso.



Instalación

Para instalar Ollama, sigue estos pasos:

1. Abre la terminal en Linux.
2. Instala curl si no lo tienes instalado.
3. Ejecuta el siguiente comando, descargará un script y lo ejecutará:

```
curl -fsSL https://ollama.com/install.sh | sh
```

4. Una vez instalado, puedes verificar la instalación ejecutando:

```
ollama version
```

Uso básico

Puedes descargar modelos preentrenados con `ollama pull <modelo>`, por ejemplo:

```
ollama pull qwen2.5:0.5b
```

Descargará el modelo *qwen2.5* de 0.5 billones de parámetros.

Para listar los modelos disponibles localmente, usa:

```
ollama list
```

Si quieres ejecutar un modelo y hacerle preguntas, usa:

```
ollama run <modelo>
```

Acerca de los modelos

Los modelos tienen diferentes capacidades y tamaños. Siguiendo el ejemplo anterior, en la [página de Ollama](#) podemos encontrar diferentes versiones de *qwen2.5*:

- qwen2.5:0.5b - 0.5 billones de parámetros
- qwen2.5:1.5b - 1.5 billones de parámetros
- qwen2.5:3b - 3 billones de parámetros`

Cuantos más parámetros, mayor capacidad de comprensión y generación de texto, pero también requerirá más recursos del sistema.

Usos de ollama

Puedes usar Ollama para diversas tareas:

- Generación de texto
- Resumen de documentos
- Traducción de idiomas
- Respuesta a preguntas
- Integración con aplicaciones personalizadas

El futuro es usar IA en local, no dependas de servicios en la nube, o acabarás siendo dependiente de ellos, estarás regalando tus datos, tu privacidad y tu libertad mientras les pagas por ello.