

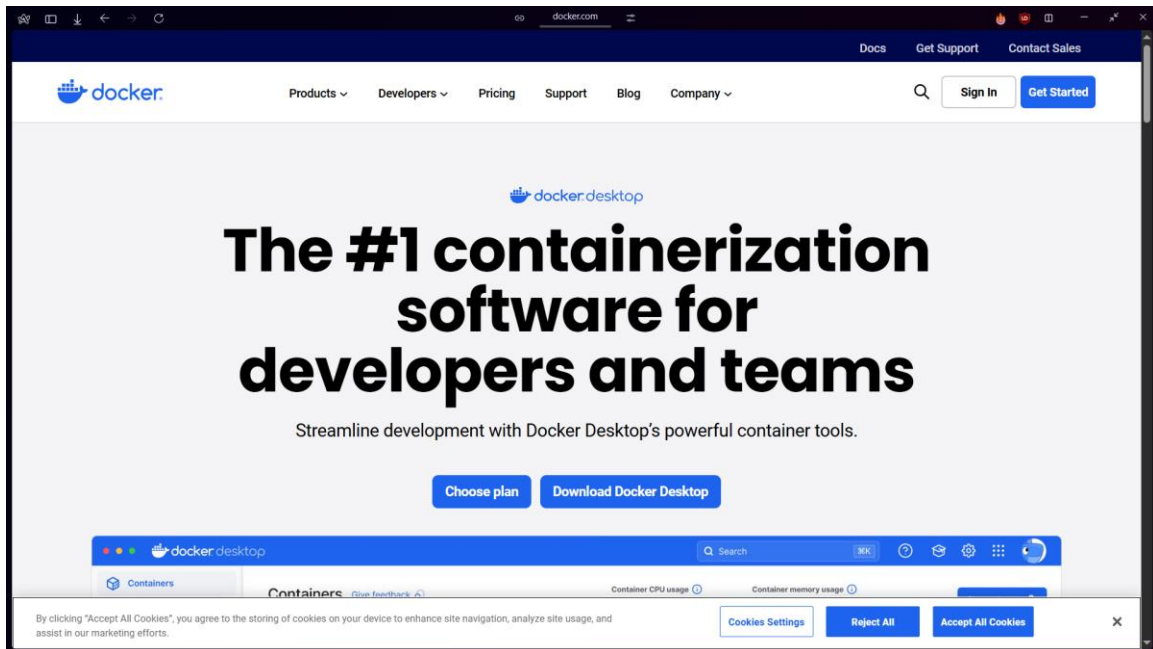
Tutorial: Cómo crear un RAG local con Docker Desktop, Ollama y n8n

Este tutorial te guía paso a paso para montar un sistema RAG (Retrieval-Augmented Generation) local utilizando Docker Desktop, Ollama y n8n. El flujo te permitirá procesar prompts con un modelo como LLaMA 2 y automatizar tareas desde n8n.

1. Instalar Docker

Visita <https://www.docker.com/products/docker-desktop> y descarga Docker Desktop para tu sistema operativo. Sigue las instrucciones del instalador. Una vez instalado, abre una terminal y ejecuta:

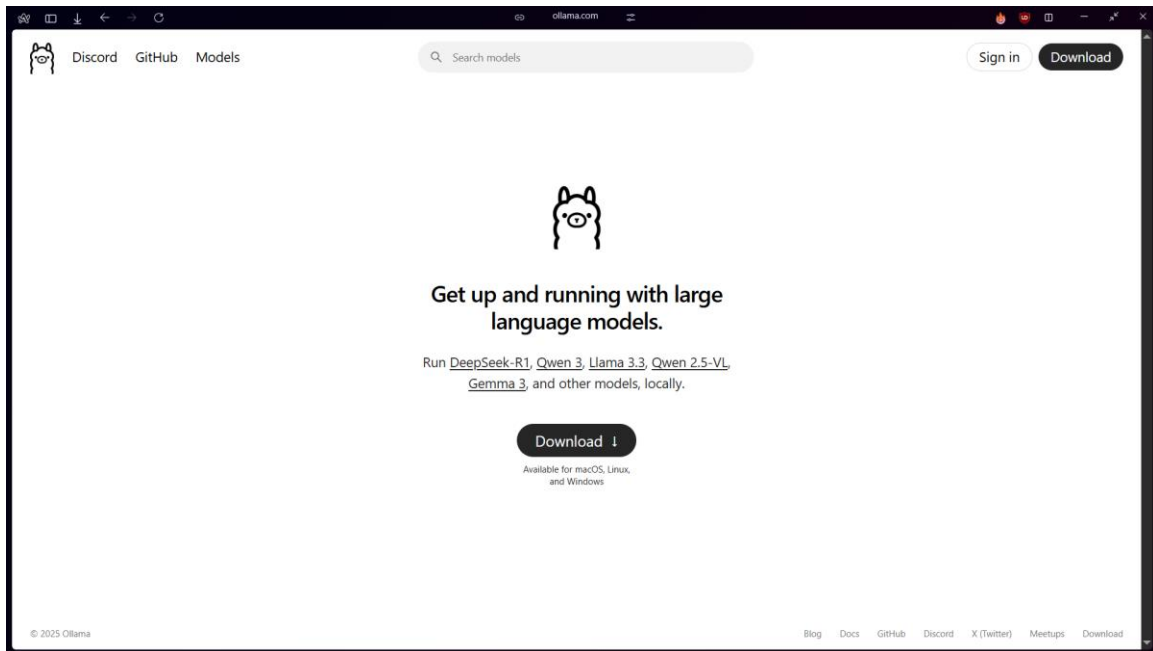
```
docker --version
```



2. Instalar Ollama

Dirígete a <https://ollama.com> y descarga el instalador de acuerdo con tu sistema operativo. Después de la instalación, asegúrate de que funciona escribiendo:

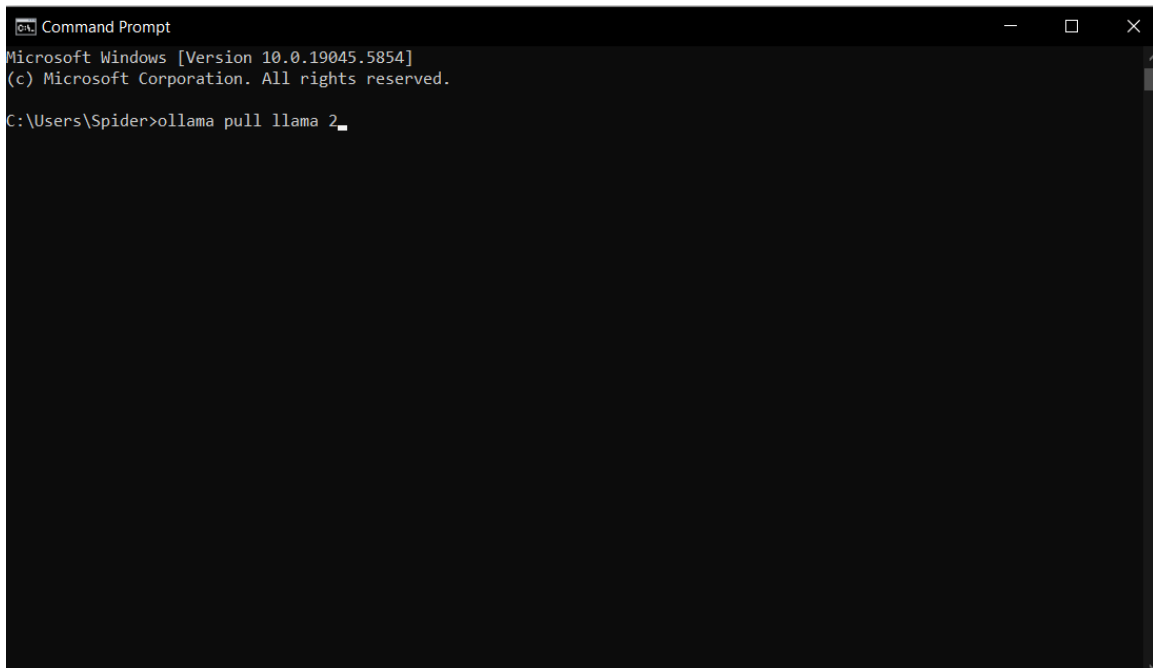
```
ollama --version
```



3. Instalar modelo LLaMA 2

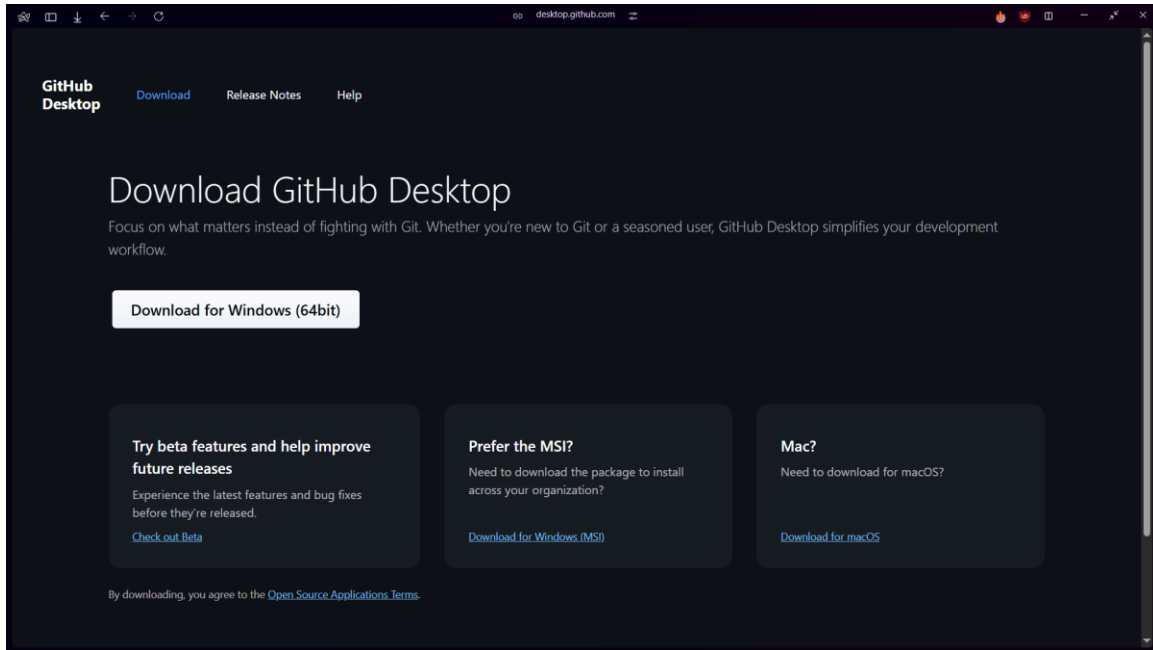
Desde CMD (o Terminal), ejecuta el siguiente comando para descargar el modelo:

`ollama pull llama2`



4. Instalar GitHub Desktop

Descarga GitHub Desktop desde <https://desktop.github.com>. Es una forma sencilla de clonar repositorios sin usar la terminal.



5. Clonar repositorio

Usa GitHub Desktop o el siguiente comando en la terminal:

```
git clone https://github.com/n8n-io/self-hosted-ai-starter-kit.git
```

```
cd self-hosted-ai-starter-kit
```

6. Crear contenedores Docker

Si estás usando una CPU, ejecuta estos comandos:

```
docker compose --profile cpu pull
```

```
docker compose create && docker compose --profile cpu up
```

7. Abrir n8n en local

Abre tu navegador y accede a <http://localhost:5678> para abrir n8n. Desde ahí puedes comenzar a construir tus flujos de trabajo.