

# **365 days of learning for an Artificial Intelligence Engineer**

**LLM local**

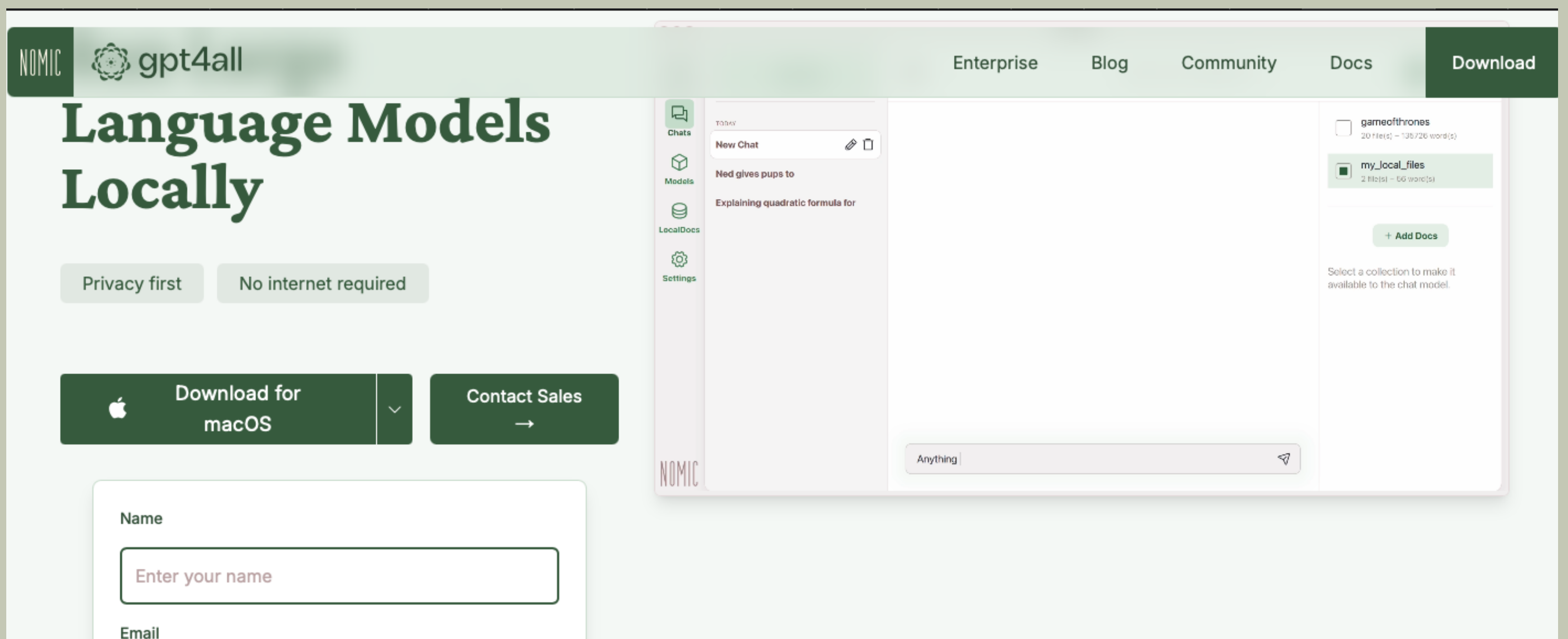


**Edgar Rios Linares**

# GPT4All

Go homepage gpt4all  
Download for your SO  
Install in your computer

<https://www.nomic.ai/gpt4all>



# Open GPT4All

Inicio

Chats

Modelos

Docs Locales

Config.

GPT4All v3.6.1



Inicio



Chats



Modelos



Docs Locales



Config.



Suscribirse al boletín

## Bienvenido a GPT4All

La aplicación de chat LLM que prioriza la privacidad



Iniciar chat

Chatear con cualquier LLM



DocumentosLocales

Chatear con tus archivos locales



Buscar modelos

Explorar y descargar modelos



### Latest News

GPT4All v3.6.1 was released on December 20th and fixes issues with the stop generation and copy conversation buttons which were broken in v3.6.0.

---

GPT4All v3.6.0 was released on December 19th. Changes include:

- **Reasoner v1:**
  - Built-in javascript code interpreter tool.
  - Custom curated model that utilizes the code interpreter to break down, analyze, perform, and verify complex reasoning tasks.
- **Templates:** Automatically substitute chat templates that are not compatible with Jinja2Cpp in GGUFs.
- **Fixes:**
  - Remote model template to allow for XML in messages.
  - Jinja2Cpp bug that broke system message detection in chat templates.
  - LocalDocs sources displaying in unconsolidated form after v3.5.0.



Notas de la versión



Documentación



Discord



X (Twitter)

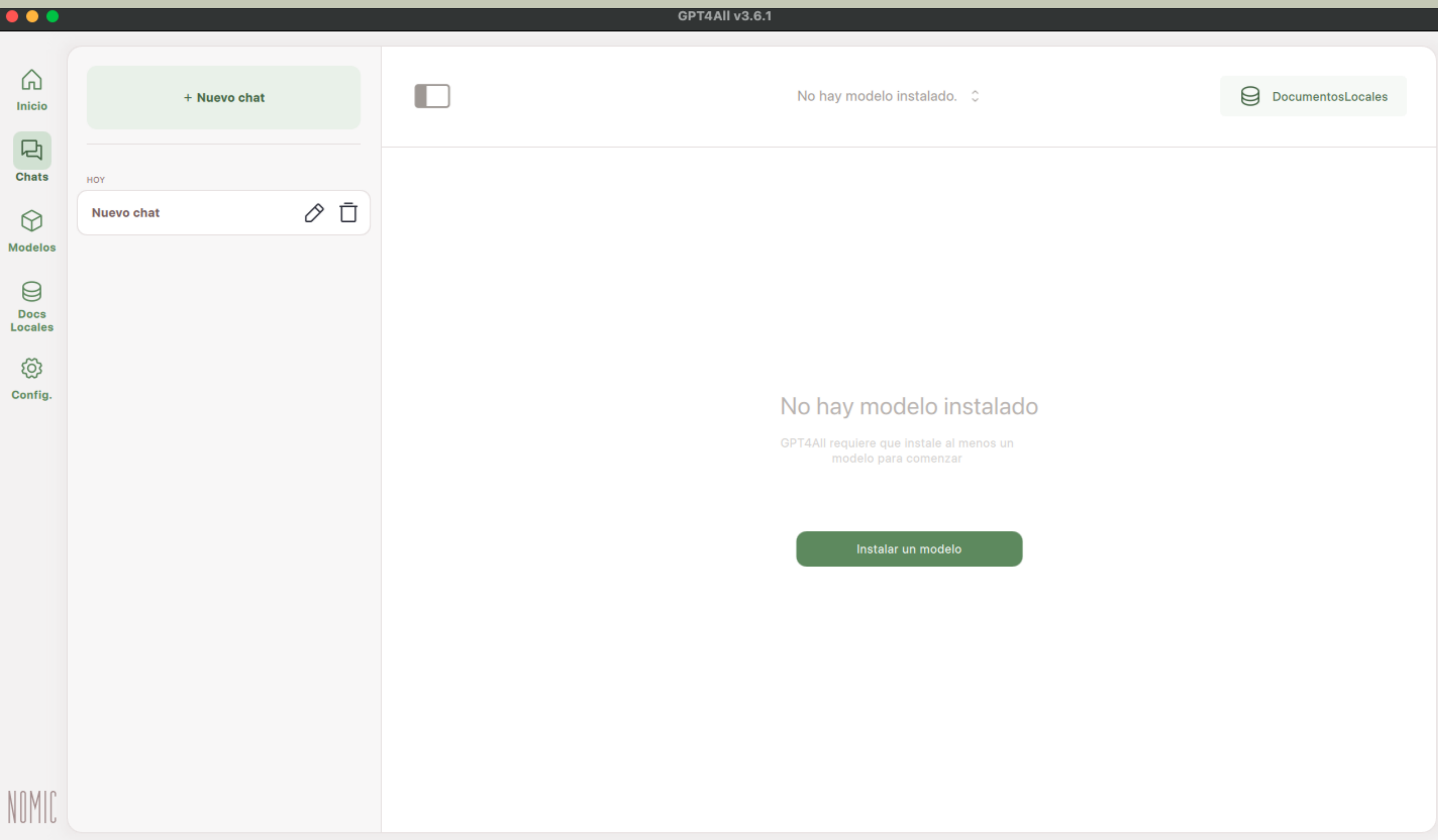


Github



nomic.ai

# Install a Model of LLM



# Install Qwen Model

Inicio

Chats

Modelos

Docs Locales

Config.

GPT4All v3.6.1

← Modelos existentes

Explorar modelos

GPT4All

HuggingFace

These models have been specifically configured for use in GPT4All. The first few models on the list are known to work the best, but you should only attempt to use models that will fit in your available memory.

All

Reasoning

Reasoner v1

- Based on [Qwen2.5-Coder 7B](#)
- Uses built-in javascript code interpreter
- Use for complex reasoning tasks that can be aided by computation analysis
- License: [Apache License Version 2.0](#)
- #reasoning

Tamaño del archivo	RAM requerida	Parámetros	Cuantificación	Tipo
4.13 GB	8 GB	8 billion	q4_0	qwen2

Descargar

Llama 3 8B Instruct

- Fast responses
- Chat based model
- Accepts system prompts in Llama 3 format
- Trained by Meta
- License: [Meta Llama 3 Community License](#)

Descargar

NOMIC

# Reasoner Model

Inicio

Chats

Modelos

Docs Locales

Config.

GPT4All v3.6.1

← Modelos existentes

Explorar modelos

GPT4All

HuggingFace

These models have been specifically configured for use in GPT4All. The first few models on the list are known to work the best, but you should only attempt to use models that will fit in your available memory.

All

Reasoning

Reasoner v1

Cancelar

35.91 MB/s

Tamaño del archivo	RAM requerida	Parámetros	Cuantificación	Tipo
4.13 GB	8 GB	8 billion	q4_0	qwen2

Llama 3 8B Instruct

NOMIC

# Check installation

Inicio

Chats

Modelos

Docs Locales

Config.

GPT4All v3.6.1

← Modelos existentes

## Explorar modelos

GPT4All

HuggingFace

These models have been specifically configured for use in GPT4All. The first few models on the list are known to work the best, but you should only attempt to use models that will fit in your available memory.

All

Reasoning

Reasoner v1

- Based on [Qwen2.5-Coder 7B](#)
- Uses built-in javascript code interpreter
- Use for complex reasoning tasks that can be aided by computation analysis
- License: [Apache License Version 2.0](#)
- #reasoning

Tamaño del archivo

4.13 GB

RAM requerida

8 GB

Parámetros

8 billion

Cuantificación

q4\_0

Tipo

qwen2

Eliminar

Llama 3 8B Instruct

- Fast responses
- Chat based model
- Accepts system prompts in Llama 3 format
- Trained by Meta
- License: [Meta Llama 3 Community License](#)

Descargar

NOMIC

Go to chat

GPT4All v3.6.1

Inicio

Chats

Modelos

Docs Locales

Config.

NOMIC

+ Nuevo chat

HOY

Nuevo chat

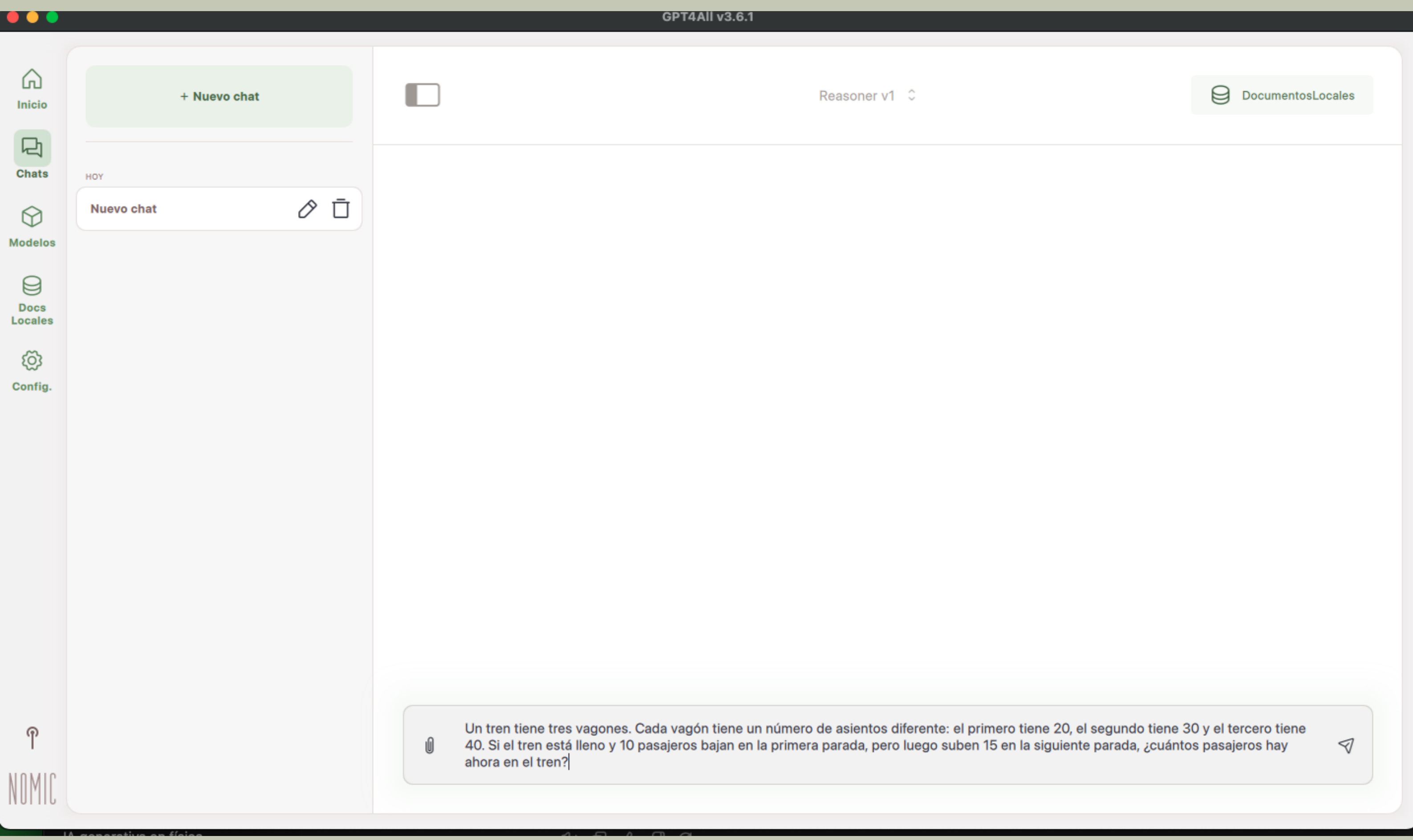
Reasoner v1

DocumentosLocales

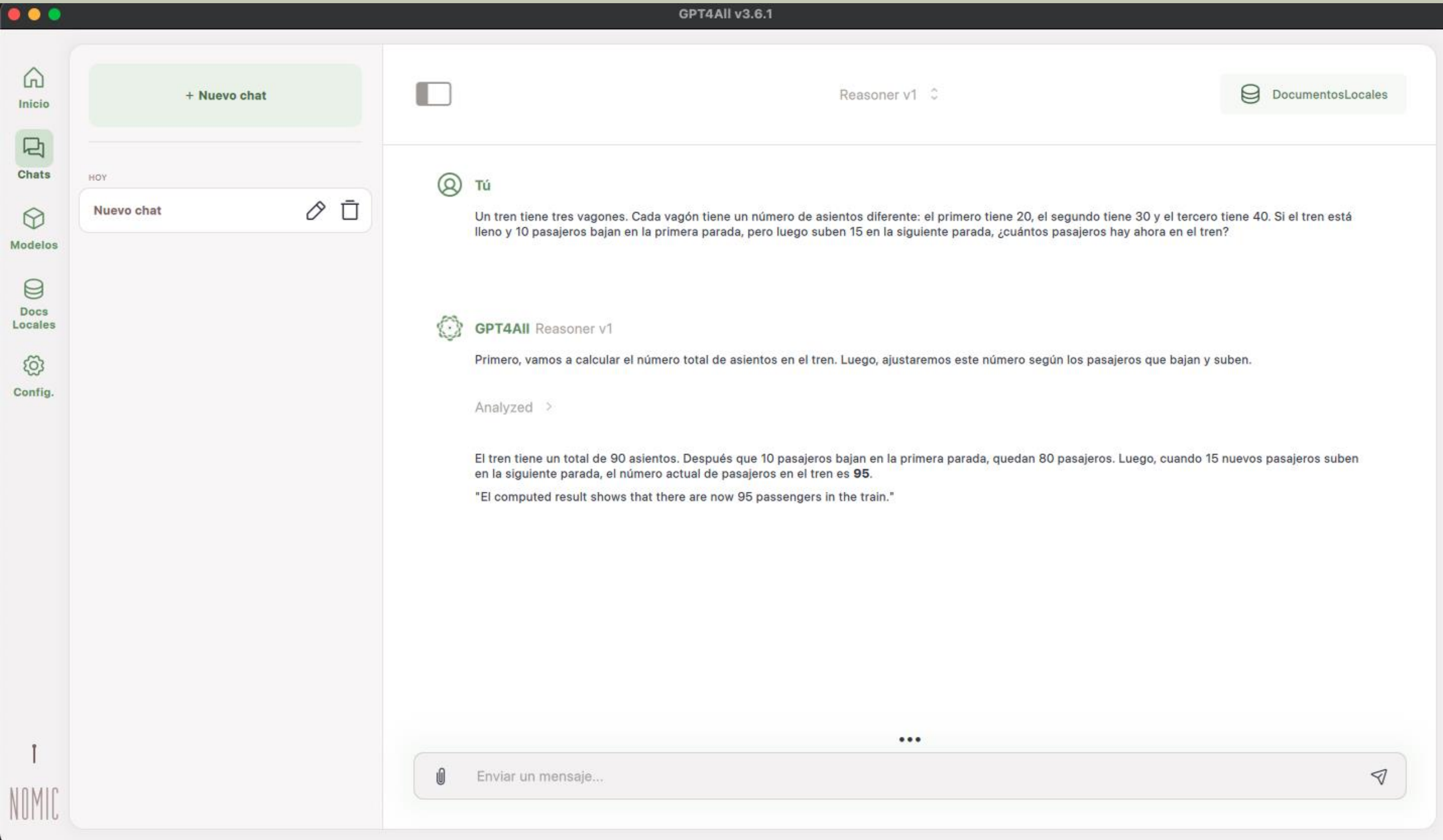
Enviar un mensaje...



# Test a problem of reasoning



# Validate result



# Install Llama Model

**GPT4All v3.6.1**



## Inicio



## Chats



## Modelo



Docs

## Locales



### Config

← Modelos existentes

## Explorar modelos

GPT4All

HuggingFace

These models have been specifically configured for use in GPT4All. The first few models on the list are known to work the best, but you should only attempt to use models that will fit in your available memory.

All

### Reasoning

## Llama 3 8B Instruct

- Fast responses
- Chat based model
- Accepts system prompts in Llama 3 format
- Trained by Meta
- License: [Meta Llama 3 Community License](#)

Cancelar

30.32 MB/s

Tamaño del archivo	RAM requerida	Parámetros	Cuantificación	Tipo
4.34 GB	8 GB	8 billion	q4_0	LLaMA3



OMI

# Upload documents

Inicio

Chats

Modelos

Docs Locales

Config.

GPT4All v3.6.1

← Colecciones existentes

Agregar colección de documentos

Agregue una carpeta que contenga archivos de texto plano, PDFs o Markdown. Configure extensiones adicionales en Configuración.

Nombre

LLM Implementation

Carpeta

is/Downloads/ml\_ai\_roadmap/llm\_implementation

Explorar

Crear colección

NOMIC

# Check upload docs

Inicio

Chats

Modelos

Docs Locales

Config.

GPT4All v3.6.1

Documentos Locales

Chatea con tus archivos locales

+ Agregar colección

LLM Implementation

/Users/edgarrioslinares/Downloads/ml\_ai\_roadmap/llm\_implementation

13 archivos – 28244 palabra(s) **nomic-embed-text-v1.5** 0 of 379 embeddings

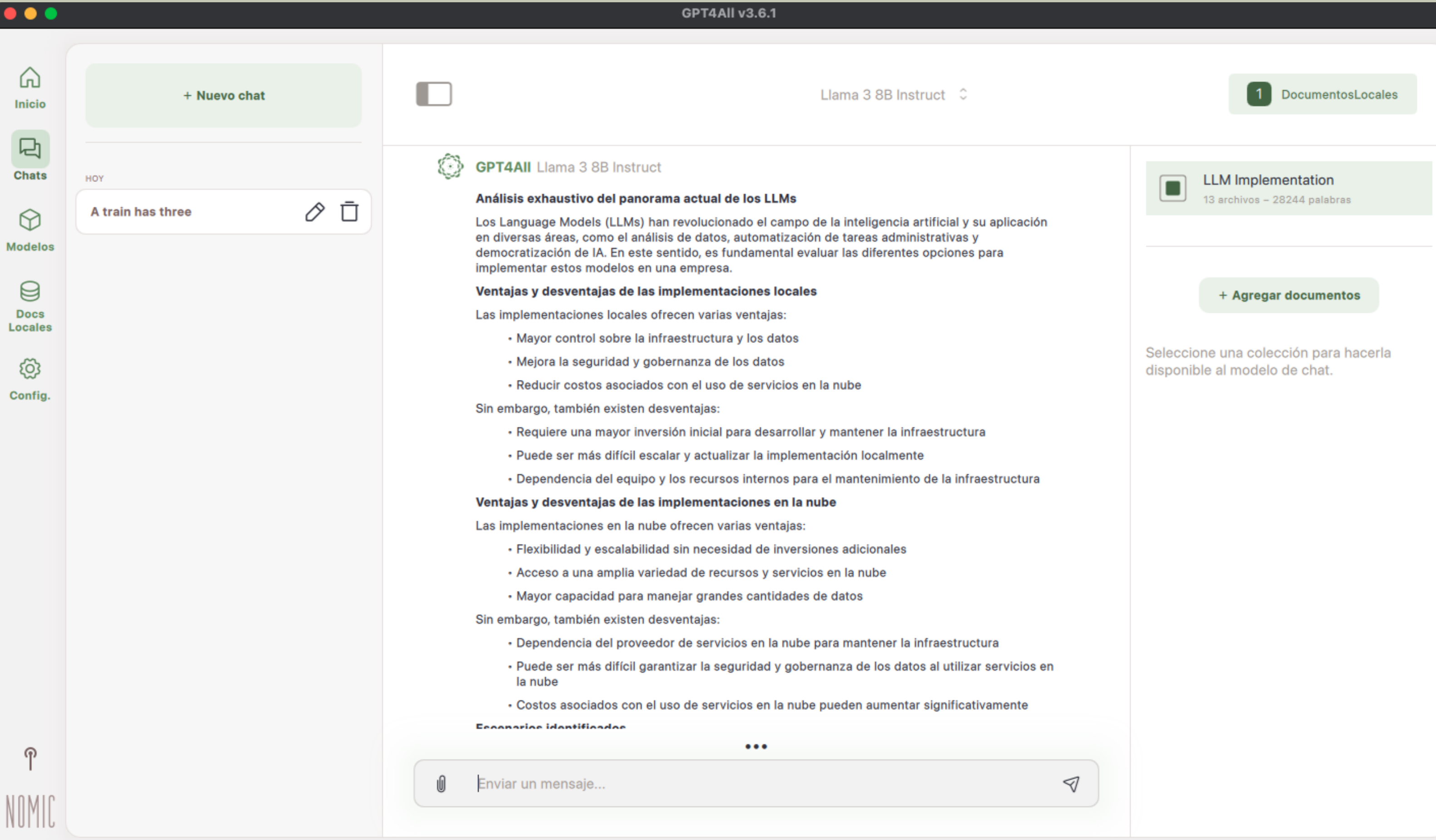
Eliminar

INCRUSTANDO

Incrustación en progreso 0%

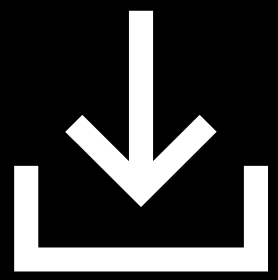
NOMIC

# Test a prompt

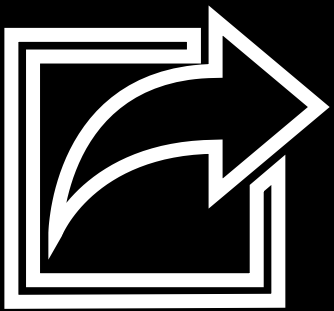




# ¿Was it helpful to you??



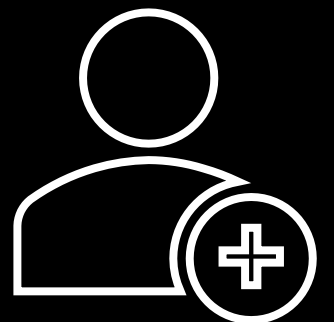
i Save it!



iShare it!



Edgar Rios Linares



iFollow me for more!

 LinkedIn —> <https://www.linkedin.com/in/erlinares/>

 Follow us on X: <https://x.com/erlinares><sup>[L]</sup><sub>[SEP]</sub>

 GitHub: [https://github.com/erlinares/365\\_AI\\_Journey/](https://github.com/erlinares/365_AI_Journey/)

 Discord: <https://discord.gg/5fFM2zh8>