



Running LLMs Locally

1. CLI - **Ollama**
 2. GUI - **Msty**
 3. Ollama **Python Library**
- 





What is Ollama?

1

Open Source

2

Runs LLM

3

Locally



How to Setup?

<https://ollama.com/>

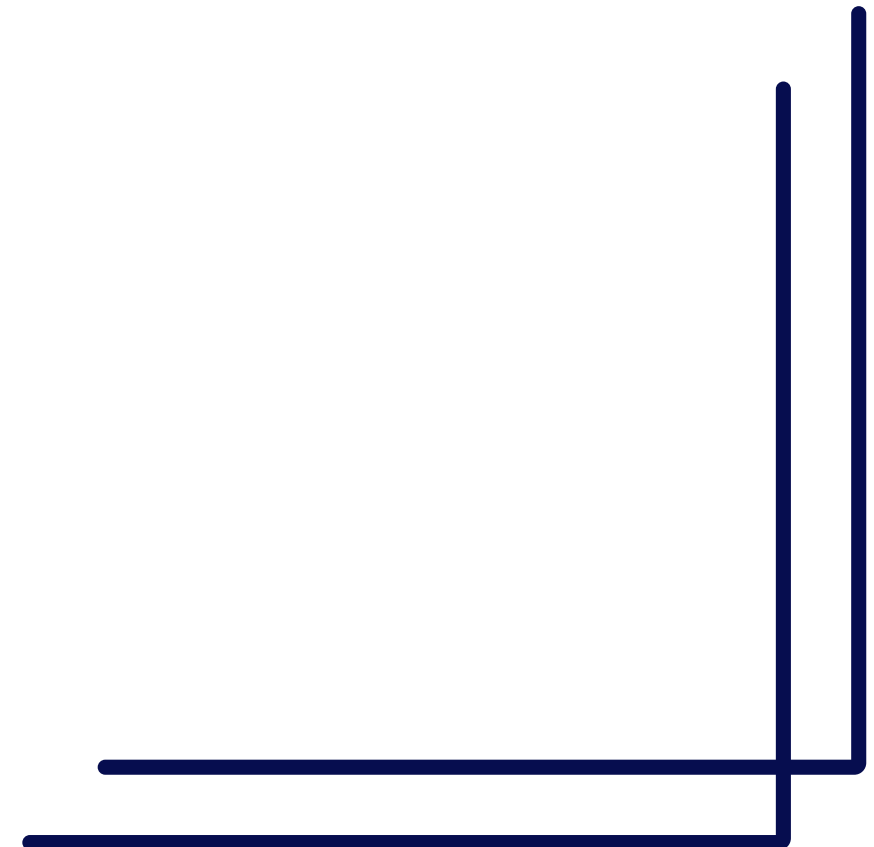


Get up and running with large language models.

Run Llama 3.3, DeepSeek-R1, Qwen 3, Mistral, Gemma 3, and other models, locally.

Download ↓

Available for macOS,
Linux, and Windows



Interact Ollama and the models using CLI

• → **Ollama** git:(main) x **ollama** list

NAME	ID	SIZE	MODIFIED
gemma3:1b	8648f39daa8f	815 MB	52 minutes ago
mxbai-embed-large:latest	468836162de7	669 MB	55 minutes ago

○ → **Ollama** git:(main) x

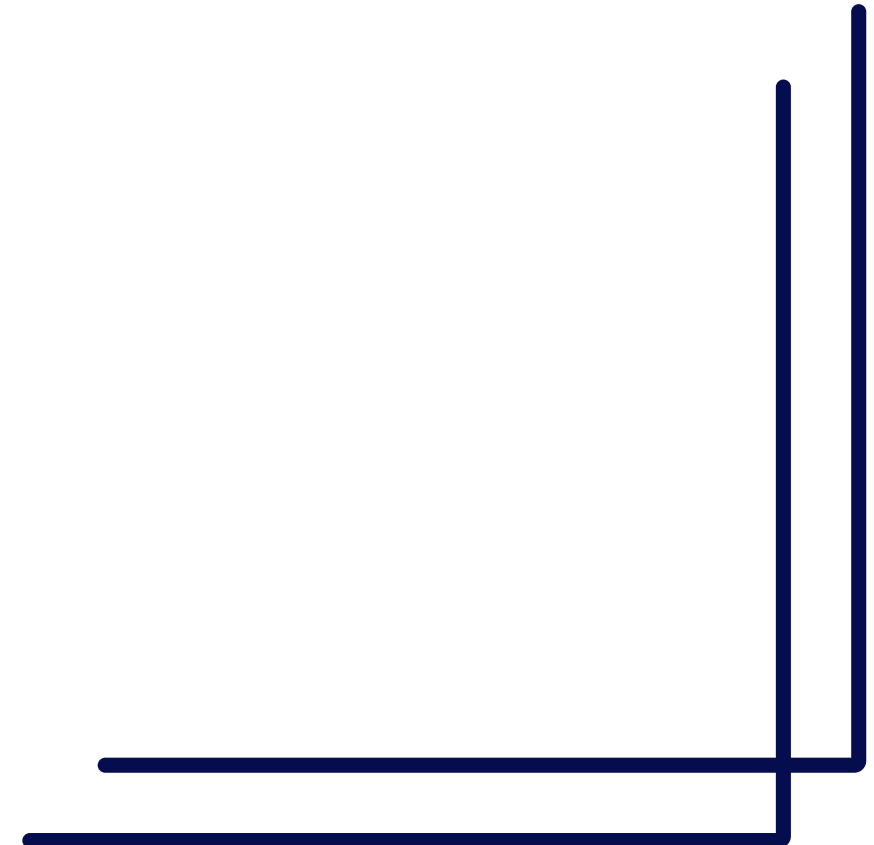




Model Info

ollama run llama3.2 >>> /show info

```
>>> /show info
Model
  architecture      llama
  parameters        3.2B
  context length     131072
  embedding length   3072
  quantization       Q4_K_M
```



Interact Ollama and the models using GUI



[Compare Us](#) ▾

[Pricing](#)

[Resources](#) ▾

 New in the latest release: Qwen3, GPT 4.1, o3, and o4-mini Models, RTL Chat Layout, Extended Code Highlighting, and more! →

The easiest way to use local and online AI models

Without Msty: painful setup, endless configurations, confusing UI, Docker, command prompt, multiple subscriptions, multiple apps, chat paradigm copycats, no privacy, no control.

With Msty: one app, one-click setup, no Docker, no terminal, offline and private, unique and powerful features.



Using local AI models for free like the new [Reasoning model from Deepseek](#) is just a step away!

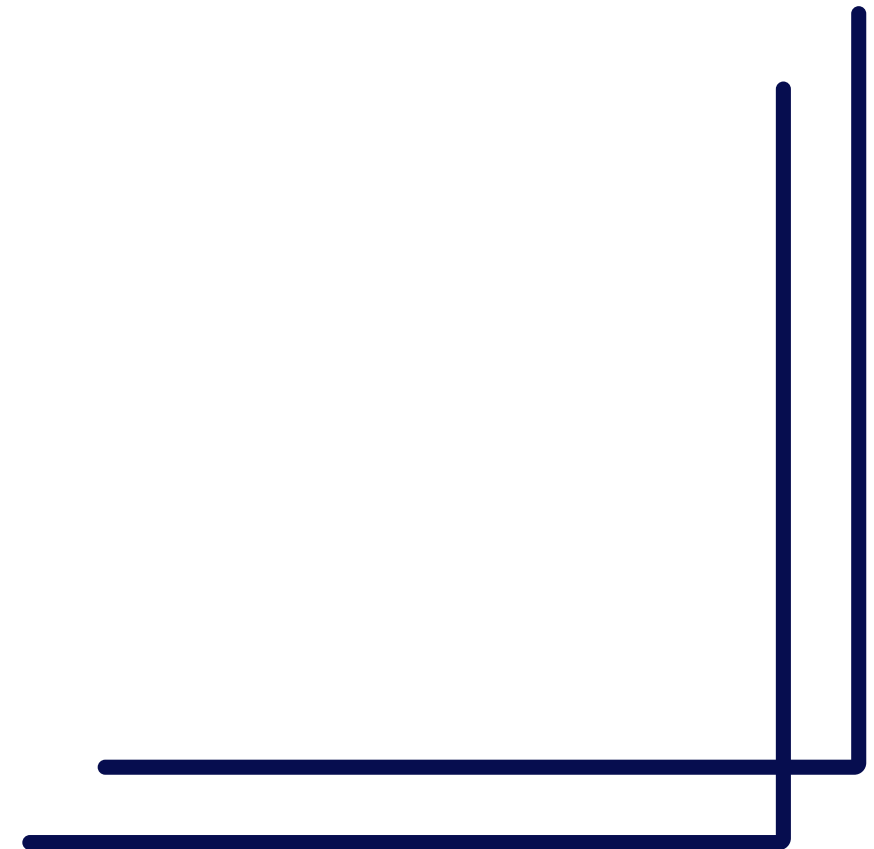
Download Msty ▾





Interact Ollama and the models using Python Library

```
import ollama
```



Use Cases for Local LLMs

Government & Defence — Sensitive Data Processing

Lawyers & Legal Firms — Confidential Case Work

Health & Pharma — Private Medical Data

Private Notes — Understand your own thoughts

