Running LLMs Locally

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

What is Ollama?

1 Open Source

2 Runs LLM

3 Locally
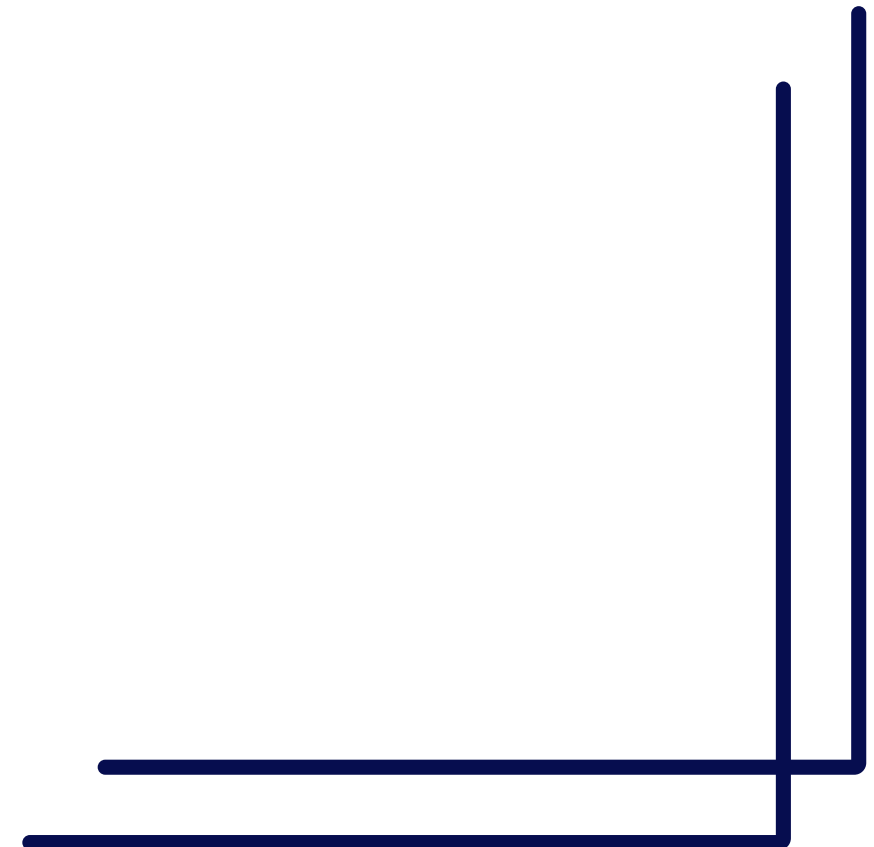
How to Setup?

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) ✗ 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) ✗
```

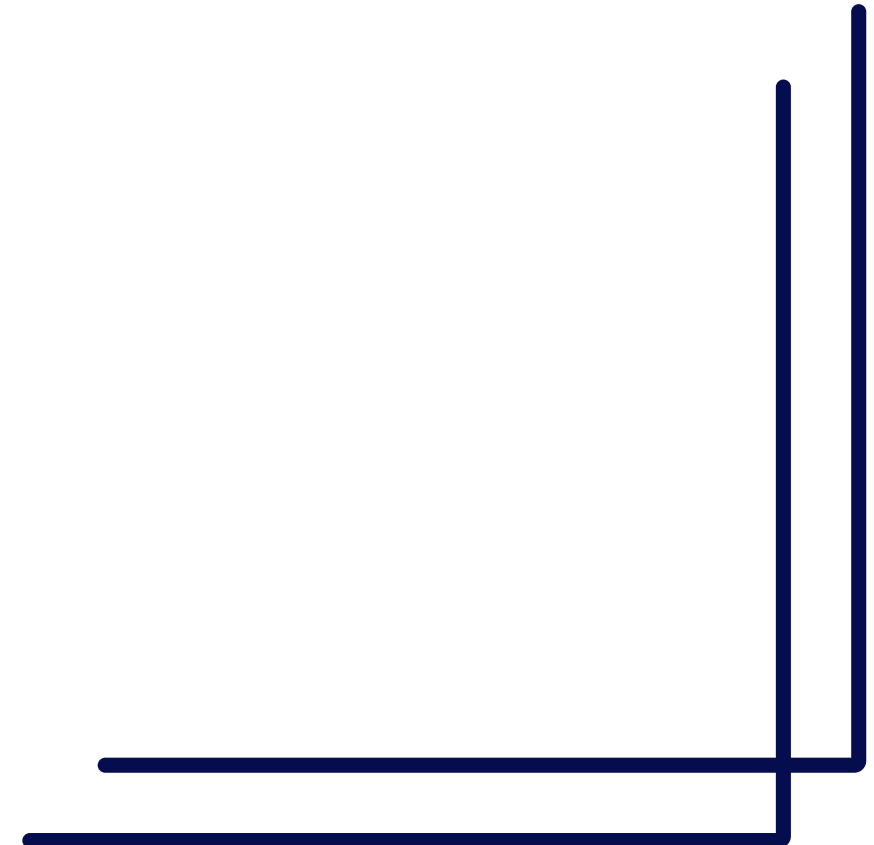# Model Info

ollama run llama3.2 >>> /show info

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

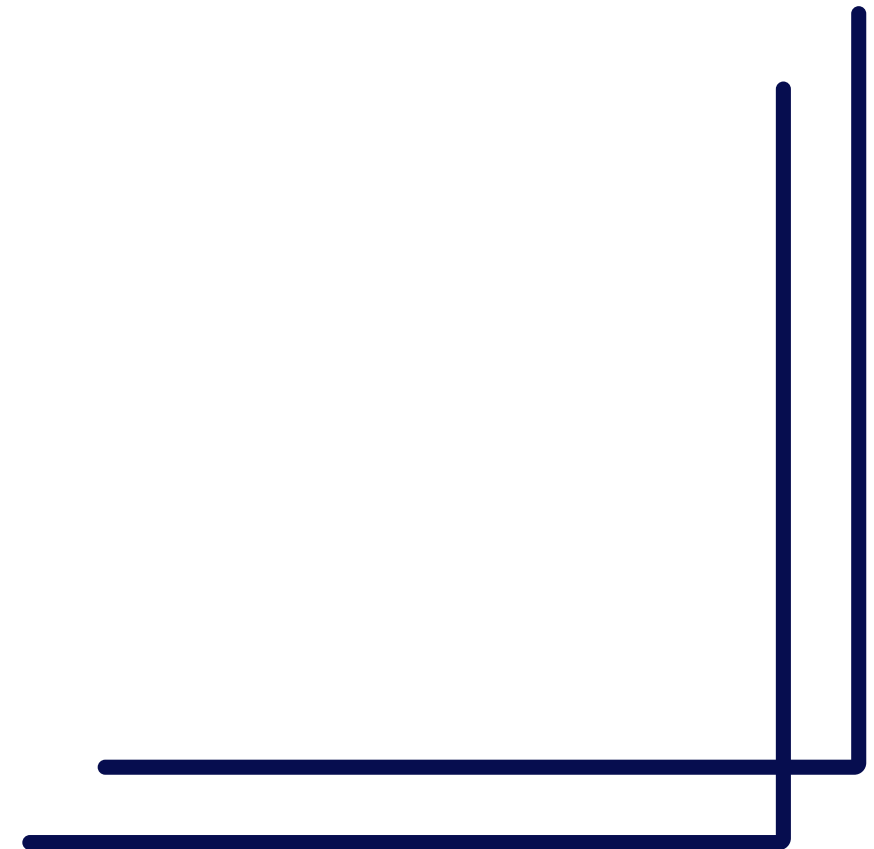# Interact Ollama and the models using GUI

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

LIKE

& SUBSCRIBE