

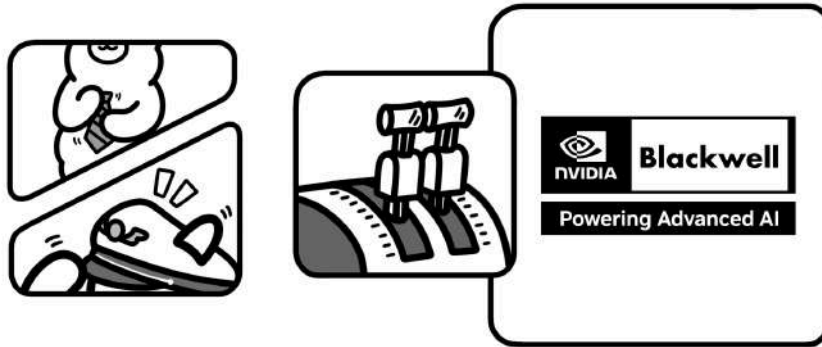


Gemma 4

Ollama partnered with Google DeepMind for the launch of Gemma 4. The Gemma 4 models are designed to deliver frontier-level performance at each size. They are well-suited for reasoning, agentic workflows, coding, and multimodal understanding.

All Gemma 4 models are available to run entirely locally on Ollama.

Ollama's cloud now runs Gemma 4 in partnership with NVIDIA on their latest Blackwell GPUs, available through a subscription for developers and third-party integrations.



Try Gemma 4

Get started with Gemma 4 using your favorite tools:

OpenClaw 🦄

```
ollama launch openclaw --model gemma4:31b-cloud
```

Claude Code

```
ollama launch claude --model gemma4:31b-cloud
```

More integrations [available](#).

Model sizes

MODEL SIZES

Gemma 4 is available in 4 sizes, 2B, 4B, 26B and 31B.

Cloud 31B

```
ollama run gemma4:31b-cloud
```

2B

```
ollama run gemma4:e2b
```

4B

```
ollama run gemma4:e4b
```

26B

```
ollama run gemma4:26b
```

31B

```
ollama run gemma4:31b
```

Learn more about the model from the [Ollama model page](#).

[Download Ollama](#)

If you have any feedback, please directly reply to this email or join Ollama's [Discord channel](#).

❤️ Ollama

You are receiving this email because you opted-in to receive updates from Ollama
Ollama, 744 High Street, Palo Alto, CA 94301

[Unsubscribe](#)