

Meta releases big, new open-source AI large language model

Open-sourced by accident -- or was it? -- back in March, Meta has now officially opened up Llama-v2, its newest large language model.



Written by **Steven Vaughan-Nichols**, Senior Contributing Editor on July 19, 2023



Getty Images/Eugene Mymrin

Meta, better known to most of us as Facebook, has released a commercial version of Llama-v2, its open-source large language model (LLM) that uses artificial intelligence (AI) to generate text, images, and code.

The first version of the Large Language Model Meta AI (Llama), was publicly announced in February and was restricted to approved researchers and organizations. However, it was soon leaked online in early March for anyone to download and use.

Meta incidentally filed take-down orders to sites such as GitHub and open-source AI group Hugging Face to corral the purloined program. Eventually, faced with the code being easily available across the web, Meta gave up trying to order the tide to go back. Instead, it embraced the release.

Also: [Want to build your own AI chatbot? Say hello to open-source HuggingChat](#)

Both versions of Llama have been trained on Common Crawl, GitHub, Wikipedia, Project Gutenberg, ArXiv, Stack Exchange, and other open test websites. While [Microsoft and OpenAI's ChatGPT](#) got the headlines, many open-source developers turned to Llama.

/ newsletters

ZDNET Tech Today

ZDNET's Tech Today newsletter is a daily briefing of the newest, most talked about stories, five days a week.

subscribe

see all →

Besides having access to Llama, Meta also shared its weights. The other major LLMs haven't. With weights, the parameters learned by a model during training, it's much easier to create and run custom AI programs. The other big LLMs, such as GPT, are usually only accessible through application programming interfaces (API).

While [AI is built on open-source foundations](#), Llama is the first major open-source LLM. Its pre-trained models have been trained on 2 trillion tokens, and have to double Llama 1's context length. Its fine-tuned models have been trained on over 1 million human annotations. Its model size parameters range from 7 to 70 billion parameters.

Also: [Microsoft announces Azure AI trio at Inspire 2023](#)

So, that's why open-source developers welcomed Llama, but [why did Meta open Llama-v2 up](#)? According to its researchers, "While many companies have opted to build AI behind closed doors, we are releasing Llama 2 openly to encourage responsible AI innovation. Based on our experience, an open approach draws upon the collective wisdom, diversity, and ingenuity of the AI practitioner community to realize the benefits of this technology. Collaboration will make these models better and safer."

But, is Llama-v2 actually open source? While Meta says the right things, the [Llama 2 Community License Agreement](#) has not been approved by the gold standard group of open-source licensing, the Open Source Initiative (OSI).

With that in mind, Meta has made [Llama 2 available free of charge for research and commercial use](#). Meta is also including model weights and starting code for the pre-trained model and conversational fine-tuned versions. This gives developers a major step forward in putting Llama-powered applications to use.

Also: [5 ways to explore the use of generative AI at work](#)

Strictly open-source or not, Llama 2 is certainly open enough for most practical purposes. And, as Amanda Brock, CEO of [OpenUK](#), put it, it's "Not an OSI approved license but a [significant release of Open Technology](#) ... This is a step to moving AI from the hands of the few to the many, democratizing technology and building trust in its use and

future through transparency. No, it's not perfect, and yes, there is more work to be done, but this bold move sets the tone for AI-open innovation with a responsible but light-touch principles-based approach to regulating the use and development of AI."

Meta isn't just playing nice with the open-source community. Meta also declared Microsoft is its preferred Llama 2 partner. So, Llama 2 is available in the Azure AI model catalog, enabling developers using Microsoft Azure to build with it. The LMA is also optimized to run locally on Windows, giving developers a seamless workflow as they bring generative AI experiences to customers across different platforms. Llama 2 is available through Amazon Web Services (AWS), Hugging Face, and other providers.

Also: Microsoft's Inspire 2023: 10 AI and partnership announcements to know

Besides opening up the code, in this release, Meta is also trying to make Llama safer, more well-behaved, and less prone to hallucinations than the other models.

Meta has done this by Red-Teaming (Security Testing) Exercises designed to clean up safety weaknesses. Meta has also released a Developer's Guide for Safe and Responsible Use to help developers understand and apply the best practices for developing and responsible model testing. Finally, Meta's provided a Llama Acceptable Use Policy to prohibit certain use cases to help ensure that these models are being used fairly and responsibly.

The end result? Meta hopes to catch up and surpass OpenAI. Who knows, it may be able to do it. As a Google AI engineer recently wrote, "The uncomfortable truth is, we aren't positioned to win this [Generative AI] arms race, and neither is OpenAI. While we've been squabbling, a third faction has been quietly eating our lunch." That third group? The open-source community.

/ featured

Is Windows 10 too popular for its own good?

5 ways to find the best place to start your career

This is how generative AI will change the gig economy for the better



[show comments](#) ↓

we equip you to harness the power of disruptive innovation, at work and at home.

[topics](#)

[galleries](#)

[videos](#)

[do not sell or share my personal information](#)

[about ZDNET](#)

[meet the team](#)

[sitemap](#)

[reprint policy](#)

[join | log in](#)

[newsletters](#)

[site assistance](#)

[licensing](#)