TRENDING

HPE Discover 2023

Getting the most out of your data centre

Join the ITPro Network

How to redu

When you purchase through links on our site, we may earn an affiliate commission. Here's how it works.

Home > Technology > Artificial Intelligence

Meta's Llama 2 is the 'first' free ChatGPT competitor, but experts rebut open source claims

By Rory Bathgate published 1 day ago

The model is a feather in Meta's democratized AI cap, and could become widely embraced by businesses









(Image credit: Getty Images)

Meta has unveiled its new AI large language model, Llama 2, which can be used freely for both commercial and research purposes and operates on a similar level to ChatGPT.

Llama 2 has been released as multiple models, with a 7-billion (7B), 13billion (13B), and 70-billion (70B) parameter sizes and fine-tuned chat variants of each.

Developers have already made Llama 2 available via API, and the model is likely to become even more widely used as it is embraced by businesses.

In benchmarks comparing the model to other open-source models, Llama 2 performed at a consistently higher level particularly when it came to math and benchmarks that probed the depth of its knowledge base.



A daily dose of IT news, reviews, features and insights, straight to your inbox!

Your Email Address

- Contact me with news and offers from other
- Receive email from us on behalf of our trusted partners or sponsors

SIGN ME UP

By submitting your information you agree to the Terms & Conditions and Privacy Policy and are aged 16 or over.

Advertisement

RECOMMENDED VIDEOS FOR YOU...

Intel | How to create a secure foundation to ke...

How to create a secure foundation to keep you business safe. Tech Ra...



Advertisement

Meta has also compared Llama 2 70B to GPT-3.5, which is used as ChatGPT's primary model, and achieved similar results across all areas except for math problems. It also almost completely outperformed Google's PaLM but was eclipsed by GPT-4 and PaLM 2.

Llama 2 70B particularly excelled at Multi-task Language Understanding, a 'zero-shot' benchmark intended to stress test a model's pre-training knowledge by presenting it with tasks that fall outside of the data on which it was trained.

Despite being less powerful than ChatGPT in some instances, it may prove more beneficial to companies looking to <u>adopt AI at zero cost</u>, although applicants whose products and services have more than 700 million monthly active users must approach Meta directly for permission to use the model.

The policy has drawn the ire of some commentators, who have argued that it limits access by certain competitors.

Lucas Beyer @giffmana · Follow Ah yes, the well known "except you, FAANG" clause that's so common in *open source* licenses like GPL, MIT, BSD, Apache2, ... See more

However, a number of high-ranking individuals from companies and institutions throughout the AI ecosystem signed a statement in support of Meta's approach to AI:

"We support an open innovation approach to AI," read the statement.

"Responsible and open innovation gives us all a stake in the Al development process, bringing visibility, scrutiny, and trust to these technologies. Opening today's Llama models will let everyone benefit from this technology."

The open technology non-profit group OpenUK has also registered its support for Llama 2, but acknowledged that the access limitations on the model prevent it from being fully classified as open source.

"Supporting this step by Meta to formally open up an LLM, offering open development and open innovation in this space, is a no-brainer for OpenUK and had unanimous support from our Board," said Amanda Brock, CEO at OpenUK.

"The democratization of AI and opening innovation around it to enable collaboration across our open communities is an essential step in the

MOST READ

- 1 Synology DiskStation
 DS1823xs+ review: A powerful
 eight-bay NAS at a reasonable
 price
- 2 Citrix discloses critical NetScaler Gateway vulnerability
- 3 How distribution can deliver valuable customer data insights
- 4 Wikipedia co-founder warns that the USA could run away with Al development
- 5 New Microsoft Al Cloud Partner Program entwines Al with the cloud

Advertisement

future of this most impactful of technologies. Whilst we recognize that this license is not an official 'open source software' one, it provides certainty for those that want to take advantage of Llama 2. This is a fantastic step in the right direction for open AI communities."

Llama 2 was trained between January and July 2023, and Meta stated it used an offline dataset. It also trained a 34-billion parameter-size model for testing, which features in its Llama 2 academic paper, but will not be released.

Meta has committed to releasing further versions of tuned Llama 2 models "as [it improves] model safety with community feedback"

The model is being distributed by Microsoft as part of an existing partnership with Meta, and will be available via Azure Al and optimized to run on Windows systems.

It will also be made available through the platforms of AWS and Hugging Face to suit a wide range of use cases throughout the market and open-source community.

At the time of launch, Meta listed partners and supporters including Accenture, Jio, IBM, DropBox, Nvidia, Intel, and AMD.

Qualcomm and Meta also announced that they will collaborate on bringing Llama 2 to mobile devices in 2024, to run Al workloads directly on devices for a localized, private experience.

Llama 2's precursor LLaMA was also open source, available on request and intended primarily for research purposes.

A week after it was announced, the model was leaked on 4chan and consequently widely shared, with some <u>calling for it to be stored on Bitcoin</u> for perpetual access.

Researchers had warned that the leaked model could be used for malicious purposes, such as <u>social engineering</u>.

ITPro Newsletter

A daily dose of IT news, reviews, features and insights, straight to your inbox!

Your Email Address	SIGN ME UP
☐ Contact me with news and offers from other Future brands	
☐ Receive email from us on behalf of our trusted partners or sponsors	
By submitting your information you agree to the Terms & Conditions and Privacy Policy and are aged 16 or over.	



Rory Bathgate

Rory Bathgate is a staff writer at ITPro covering the latest news on UK networking and data protection, privacy and compliance. He can sometimes be found on the ITPro Podcast, swapping a keyboard for a microphone to discuss the latest in tech trends.

In his free time, Rory enjoys photography, video editing and graphic design alongside good science fiction. After graduating from the University of Kent with BA in English and American Literature, Rory

took an MA in Eighteenth-Century Studies at King's College London. He joined ITPro in 2022 as a graduate, after four years in student journalism. You can contact Rory at rory.bathgate@futurenet.com or on LinkedIn.

MORE ABOUT ARTIFICIAL INTELLIGENCE



Wikipedia cofounder warns that the USA could run away with Al development ▶



8x8 unveils new partner ecosystem to advance Al experiences ▶

LATEST



Synology
DiskStation
DS1823xs+ review:
A powerful eightbay NAS at a
reasonable
price ▶

SEE MORE LATEST ▶

TOPICS

META

SPONSORED FINANCIAL CONTENT



ISA Millionaires Reveal Their Secrets & Strategies. Risk Of Loss.

Hargreaves Lansdown



Trusted by traders for 40 years. 72% of retail CFD accounts lose money CITY INDEX



Track your shares, dividends & more in one place with Sharesight.

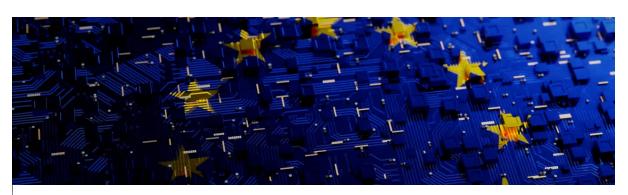
Sharesight Investment Tracker



Investors' Chronicle: Investing in today's climate

Sponsored by Allianz

MOST POPULAR



EU Cyber Resilience Act a 'death knell' for open source software, critics warn

By Ross Kelly 18 July 2023



Arrow Electronics unveils new managed services portfolioBy Daniel Todd 18 July 2023



Slackware celebrates 30 years in the Linux distribution worldBy Richard Speed 17 July 2023



Report: Female tech workers disproportionately affected by industry layoffs

By Ross Kelly 17 July 2023



CIOs can drive more value than other C-suite executives, says expert

By Rory Bathgate 17 July 2023



Weekly cyber attacks reach two-year high amid ransomware resurgence

By Ross Kelly 17 July 2023



US says National Cybersecurity Strategy will focus on market resilience and private partnerships

By Rory Bathgate 14 July 2023



AlmaLinux follows Oracle in ditching RHEL compatibility

By Richard Speed 14 July 2023



JumpCloud reveals nation-state hackers breached internal systems, following customer speculation

By Rory Bathgate 14 July 2023



TD Synnex becomes exclusive UK distributor for Synology BeeDrive

By Daniel Todd 14 July 2023



8x8 unveils new partner ecosystem to advance AI experiences

By Daniel Todd 13 July 2023

RESOURCES



The gratitude gap

Posted



Top six networking trends for 2023

Posted



The near and far future of ransomware business models

Posted



Unified consoles create a seamless multi-cloud managemer experience

Posted

VIEW MORE WHITEPAPERS

Terms and conditions Contact Future's experts Privacy policy Cookies policy

Accessibility statement Careers About Us Contact Us

GDPR consent

[©] Future Publishing Limited Quay House, The Ambury, Bath BA1 1UA. All rights reserved. England and Wales company registration number 2008885.