Llama 2: why is Meta releasing open-source AI model and are there any risks?

Publication Date: 2023-07-20

Author: Dan Milmo

Section: Technology

Tags: Artificial intelligence (AI), Meta, Computing, Nick Clegg, Google, Microsoft, ChatGPT, explainers

Article URL: https://www.theguardian.com/technology/2023/jul/19/why-is-meta-releasing-llama-2-open-source-ai-model-mark-zuckerberg

Mark Zuckerberg's Meta has this week released an open-source version of an artificial intelligence model, Llama 2, for public use. The large language model (LLM), which can be used to create a ChatGPT-like chatbot, is available to startups, established businesses and lone operators. But why is Meta doing this and what are the potential risks involved? What does an open-source LLM do? LLMs underpin AI tools such as chatbots. They are trained on vast datasets that enable them to mimic human language and even computer coding. If an LLM is made open-source that means its content is made freely available for people to access, use and tweak to their own purpose. Llama 2 is being released in three versions, including one that can be built into an AI chatbot. The idea is that startups or established businesses can access Llama 2 models and tinker with them to create their own products including, potentially, rivals to ChatGPT or Google's Bard chatbot – although by Meta's own admission Llama 2 is not quite at the level of GPT-4, the LLM behind OpenAI's ChatGPT. Why is Meta releasing it for public use? Nick Clegg, Meta's president of global affairs, told BBC Radio 4's Today programme on Wednesday that making LLMs open-source would make them "safer and better" by inviting outside scrutiny. "With the … wisdom of crowds you actually make these systems safer and better and, crucially, you take them out of the … clammy hands of the big tech companies which currently are the only companies that have either the computing power or the vast reservoirs of data to build these models in the first place." There is also a possibility that by giving all comers the chance to launch a rival to ChatGPT, Bard or Microsoft's Bing chatbot, Meta is potentially diluting the competitive edge of tech peers such as Google. Meta has admitted in research published alongside Llama 2 that it "lags behind" GPT-4, but it is a free competitor to OpenAI nonetheless. Microsoft is a key financial backer of OpenAI but is nonetheless supporting the launch of Llama 2. The LLM is available for download via the Microsoft Azure, Amazon Web Services and Hugging Face platforms. Are there concerns about open-source AI? Tech professionals including Elon Musk, a co-founder of OpenAI, have expressed concerns about an AI arms race. Open-sourcing makes a powerful tool in this technology available to all. Dame Wendy Hall, regius professor of computer science at the University of Southampton, told the Today programme there were questions over whether the tech industry could be trusted to self-regulate LLMs, with the problem looming even larger for open-source models. "It's a bit like giving people a template to build a nuclear bomb," she said. Dr Andrew Rogoyski, of the Institute for People-Centred AI at the University of Surrey, said open-source models were difficult to regulate. "You can't really regulate open source. You can regulate the repositories, like Github or Hugging Face, under local legislation," he said. "You can issue licence terms on the software that, if abused, could make the abusing company liable under various forms of legal

redress. However, being open source means anyone can get their hands on it, so it doesn't stop the wrong people grabbing the software, nor does it stop anyone from misusing it." Does Meta have safeguards in place? If you apply to download Llama 2 you are required to agree to an "acceptable use" policy that includes not using the LLMs to encourage or plan "violence or terrorism" or to generate disinformation. However, LLMs such as that behind ChatGPT are prone to producing false information and can be coaxed into overriding safety guardrails to produce dangerous content. The Llama 2 release is also accompanied by a responsible use guide for developers.