

‘Very scary’: Mark Zuckerberg’s pledge to build advanced AI alarms experts

Publication Date: 2024-01-19

Author: Dan Milmo

Section: Technology

Tags: Artificial intelligence (AI), Mark Zuckerberg, Facebook, Meta, Technology sector, news

Article URL: <https://www.theguardian.com/technology/2024/jan/19/mark-zuckerberg-artificial-general-intelligence-system-alarms-experts-meta-open-source>



Mark Zuckerberg has been accused of taking an irresponsible approach to artificial intelligence after committing to building a powerful AI system on a par with human levels of intelligence. The Facebook founder has also raised the prospect of making it freely available to the public. The Meta chief executive has said the company will attempt to build an artificial general intelligence (AGI) system and make it open source, meaning it will be accessible to developers outside the company. The system should be made “as widely available as we responsibly can”, he added. In a Facebook post, Zuckerberg said it was clear that the next generation of tech services “requires building full general intelligence”. AGI is not a strictly defined term but it commonly refers to a theoretical AI system that can carry out an array of tasks at a level of intelligence that matches or exceeds humans. The potential emergence of AGI has alarmed experts and politicians around the world who fear such a system, or a combination of multiple AGI systems, could evade human control and threaten humanity. Zuckerberg confirmed that Meta would consider making its AGI open source, or freely available to developers and the public to use and adapt, like the company’s Llama 2 AI model. “This technology is so important and the opportunities are so great that we should open source and make it as widely available as we responsibly can, so everyone can benefit,” he said. Dame Wendy Hall, a professor of computer science at the University of Southampton and a member of the UN’s advisory body on AI, said the prospect of open source AGI was “really very scary” and that it was irresponsible of Zuckerberg to consider it. “The thought of open source AGI being released before we have worked out how to regulate these very powerful AI systems is really very scary,” she said. “In the wrong hands technology like this could do a great deal of harm. It is so irresponsible for a company to suggest it.” Hall added: “Thankfully I think it’s still many years away before such an aspiration [AGI] can be achieved in any meaningful way, so we have time to put the regulation systems in place. But it is a matter of public safety that we progress this work with some urgency.” Meta was among the tech firms at a global AI safety summit in the UK last year that committed to allowing governments to vet artificial intelligence tools before and after their release. Another UK-based expert said open sourcing an AGI system was not a decision that should be left in the hands of a tech company. “There are deep and complex arguments about the merits of open-sourcing current AI models, pushing that into the realm of AGI could be world-saving or catastrophic,” said Dr Andrew Rogoyski, a director of the Institute for People-Centred AI at the University of Surrey. “These decisions need to be taken by international consensus, not in the boardroom of a tech giant.” In an interview on Thursday with the tech news website the Verge, Zuckerberg said he would lean towards open sourcing for as long as it made sense to and was the “safe and responsible thing to do”. Meta’s decision last year to open source

Llama 2 drew criticism from some experts, including Hall, who said it was “a bit like giving people a template to build a nuclear bomb”. The California-based OpenAI, the developer of ChatGPT, is also building AGI, which it defines as “AI systems that are generally smarter than humans”, while Demis Hassabis, the head of Google’s AI division Google DeepMind, has said AGI could be less than a decade away. OpenAI’s chief executive, Sam Altman, sounded a note of caution at the World Economic Forum in Davos, Switzerland, this week, when he said further advances in AI would be impossible without a breakthrough in energy provision such as nuclear fusion. Zuckerberg did not give a timeframe for developing an AGI system, although he referred to Meta having built an “absolutely massive amount of infrastructure” to create new AI systems. The company has ordered a significant stockpile of AI processing chips. Zuckerberg added that work on a sequel to Llama 2 was also under way.