

# Google, Microsoft, OpenAI and startup form body to regulate AI development

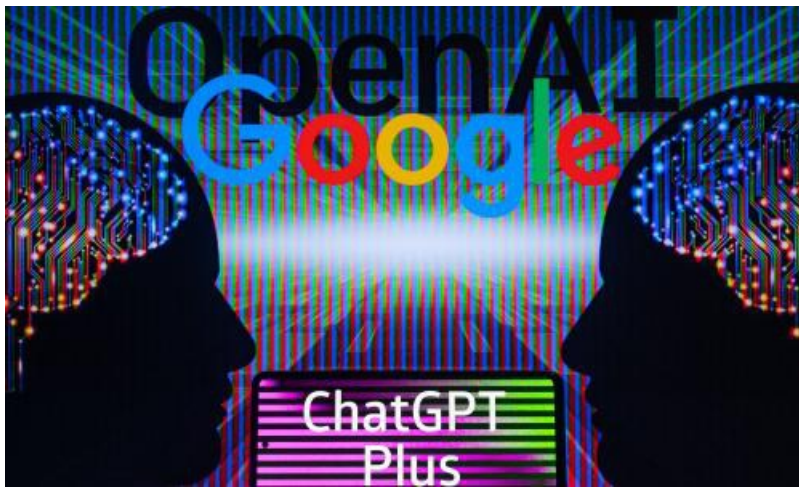
Publication Date: 2023-07-26

Author: Dan Milmo

Section: Technology

Tags: Artificial intelligence (AI), Google, OpenAI, Microsoft, Technology sector, news

Article URL: <https://www.theguardian.com/technology/2023/jul/26/google-microsoft-openai-anthropic-ai-frontier-model-forum>



Four of the most influential companies in artificial intelligence have announced the formation of an industry body to oversee safe development of the most advanced models. The Frontier Model Forum has been formed by the ChatGPT developer OpenAI, Anthropic, Microsoft and Google, the owner of the UK-based DeepMind. The group said it would focus on the “safe and responsible” development of frontier AI models, referring to AI technology even more advanced than the examples available currently. “Companies creating AI technology have a responsibility to ensure that it is safe, secure, and remains under human control,” said Brad Smith, the president of Microsoft. “This initiative is a vital step to bring the tech sector together in advancing AI responsibly and tackling the challenges so that it benefits all of humanity.” The forum’s members said their main objectives were to promote research in AI safety, such as developing standards for evaluating models; encouraging responsible deployment of advanced AI models; discussing trust and safety risks in AI with politicians and academics; and helping develop positive uses for AI such as combating the climate crisis and detecting cancer. They added that membership of the group was open to organisations that develop frontier models, which is defined as “large-scale machine-learning models that exceed the capabilities currently present in the most advanced existing models, and can perform a wide variety of tasks”. The announcement comes as moves to regulate the technology gather pace. On Friday, tech companies – including the founder members of the Frontier Model Forum – agreed to new AI safeguards after a White House meeting with Joe Biden. Commitments from the meeting included watermarking AI content to make it easier to spot misleading material such as deepfakes and allowing independent experts to test AI models. The White House announcement was met with scepticism by some campaigners who said the tech industry had a history of failing to adhere to pledges on self-regulation. Last week’s announcement by Meta that it was releasing an AI model to the public was described by one expert as being “a bit like giving people a template to build a nuclear bomb”. The forum announcement refers to “important contributions” being made to AI safety by bodies including the UK government, which has convened a global summit on AI safety, and the EU, which is introducing an AI act that represents the most serious legislative attempt to regulate the technology. Dr Andrew Rogoyski, of the Institute for People-Centred AI at the University of Surrey, said oversight of artificial intelligence must not fall foul of “regulatory capture”, whereby companies’ concerns dominate the regulatory process. He added: “I have grave concerns that governments have ceded leadership in AI to the private sector, probably irrecoverably. It’s such a powerful technology, with great potential for good and ill, that it needs independent oversight that will represent people, economies and societies which will be impacted by AI in the future.”