

AI-focused tech firms locked in 'race to the bottom', warns MIT professor

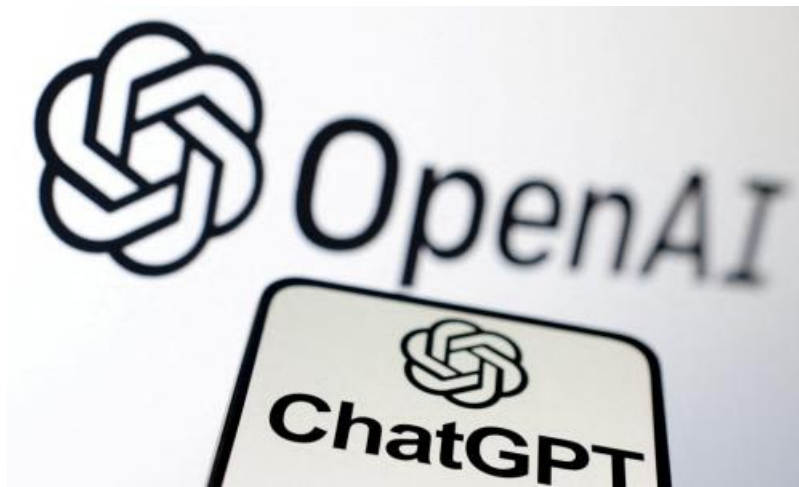
Publication Date: 2023-09-21

Author: Dan Milmo

Section: Technology

Tags: Technology, Artificial intelligence (AI), Computing, ChatGPT, Meta, Google, news

Article URL: <https://www.theguardian.com/technology/2023/sep/21/ai-focused-tech-firms-locked-race-bottom-warns-mit-professor-max-tegmark>



The scientist behind a landmark letter calling for a pause in developing powerful artificial intelligence systems has said tech executives did not halt their work because they are locked in a “race to the bottom”. Max Tegmark, a co-founder of the Future of Life Institute, organised an open letter in March calling for a six-month pause in developing giant AI systems. Despite support from more than 30,000 signatories, including Elon Musk and the Apple co-founder Steve Wozniak, the document failed to secure a hiatus in developing the most ambitious systems. Speaking to the Guardian six months on, Tegmark said he had not expected the letter to stop tech companies working towards AI models more powerful than GPT-4, the large language model that powers ChatGPT, because competition has become so intense. “I felt that privately a lot of corporate leaders I talked to wanted [a pause] but they were trapped in this race to the bottom against each other. So no company can pause alone,” he said. The letter warned of an “out-of-control race” to develop minds that no one could “understand, predict, or reliably control”, and urged governments to intervene if a moratorium on developing systems more powerful than GPT-4 could not be agreed between leading AI companies such as Google, ChatGPT owner OpenAI and Microsoft. It asked: “Should we develop nonhuman minds that might eventually outnumber, outsmart, obsolete and replace us? Should we risk loss of control of our civilisation?” Tegmark, a professor of physics and AI researcher at the Massachusetts Institute of Technology, said he viewed the letter as a success. “The letter has had more impact than I thought it would,” he said, pointing to a political awakening on AI that has included US Senate hearings with tech executives and the UK government convening a global summit on AI safety in November. Expressing alarm about AI had gone from being taboo to becoming a mainstream view since the letter’s publication, Tegmark said. The letter from his thinktank was followed in May by a statement from the Center for AI Safety, backed by hundreds of tech executives and academics, declaring that AI should be considered a societal risk on a par with pandemics and nuclear war. “I felt there was a lot of pent-up anxiety around going full steam ahead with AI, that people around the world were afraid of expressing for fear of coming across as scare-mongering luddites. The letter legitimised talking about it; the letter made it socially acceptable. “So you’re getting people like [letter signatory] Yuval Noah Harari saying it, you’ve started to get politicians asking tough questions,” said Tegmark, whose thinktank researches existential threats and potential benefits from cutting-edge technology. Fears around AI development range from the immediate, such as the ability to generate deepfake videos and mass-produce disinformation, to the existential risk posed by super-intelligent AIs that evade human control or make irreversible and highly consequential decisions. Tegmark warned against describing the development of digital “god-like general intelligence” as a long-term threat, citing some AI practitioners

who believe it could happen within a few years. The Swedish-American scientist said November's UK AI safety summit, to be held at Bletchley Park, was a "wonderful thing". His thinktank has said the summit should target three achievements: establishing a common understanding of the severity of risks posed by AI; recognising that a unified global response is needed; and embracing the need for urgent government intervention. He added that a hiatus in development was still needed until global agreed safety standards were met. "Making models more powerful than what we have now, that has to be put on pause until they can meet agreed-upon safety standards." He added: "Agreeing on what the safety standards are will naturally cause the pause." Tegmark also urged governments to take action on open-source AI models that can be accessed and adapted by members of the public. Mark Zuckerberg's Meta recently released an open-source large language model, called Llama 2, and was warned by one UK expert that such a move was akin to "giving people a template to build a nuclear bomb". "Dangerous technology should not be open source, regardless of whether it is bio-weapons or software," Tegmark said.