

‘Bletchley made me more optimistic’: how experts reacted to AI summit

Publication Date: 2023-11-03

Author: Dan Milmo

Section: Technology

Tags: Artificial intelligence (AI), Elon Musk, Computing, Rishi Sunak, features

Article URL: <https://www.theguardian.com/technology/2023/nov/03/bletchley-made-me-more-optimistic-how-experts-reacted-to-ai-summit>



Bletchley Park, a milestone in Alan Turing’s journey to technological immortality, heard warnings this week that the coming wave of artificial intelligence systems could threaten humanity. But for one of the world’s leading tech investors, holding back AI development will be just as damaging in terms of deaths in car crashes, pandemics and poorly targeted munitions that could have been prevented by the technology. “We believe any deceleration of AI will cost lives. Deaths that were preventable by the AI that was prevented from existing is a form of murder,” wrote Marc Andreessen, an early investor in Facebook, Pinterest and Twitter, in a blogpost last month titled *The Techno-Optimist Manifesto*. When it comes to AI, Andreessen is not the only techno-optimist out there, despite the pessimistic view of the technology dominating the agenda in the run-up to last week’s AI safety summit at Bletchley. Elon Musk described AI at the gathering as “one of the biggest threats to humanity”, but in a conversation with Rishi Sunak on Thursday he also cast the technology in utopian terms. AI is “a magic genie” that grants you limitless wishes and will usher in an “age of abundance”, he said. This is the more benign view of artificial general intelligence, a theoretical system that can carry out an array of tasks at human or above-human levels of intelligence. Musk has said his AI venture, xAI, will attempt to build a system that is “maximally curious” about humanity. Others believe the existential fears – based on highly advanced AI systems evading human control and making decisions that threaten humanity – are overplayed. Nick Clegg, the former deputy prime minister turned powerful tech executive, this compared the fuss over AI to the 80s “moral panic” over video games. One of his colleagues – Yann LeCun, a respected figure in the field – has described concerns over AI wiping out humanity as “preposterous”. Generally, however, the optimism is more nuanced than Andreessen’s impassioned take. Clegg, who attended the summit as president of global affairs at Mark Zuckerberg’s Meta, acknowledged this week that there are short-term threats that need to be addressed, such as the possibility that AI-generated disinformation will affect elections in the US, India, Indonesia, Mexico and the UK next year. “The newspapers love running pictures of scary robots with glaring red eyes and saying they’re going to take over tomorrow. But actually there’s a lot of homework that needs to be done on more proximate challenges, which I worry may play second fiddle to some of those more speculative risks.” This caveated optimism was typical among attendees at the AI safety summit. Demis Hassabis, the chief executive of Google DeepMind, a world-leading AI research lab, has warned that dealing with the risks from artificial intelligence must be treated as seriously as the climate crisis. But he is also an unabashed optimist about the positive work that AI can do in fields such as health and medicine, as shown by his unit’s achievements in mapping protein structures via the AlphaFold program. Mustafa Suleyman, a co-founder of DeepMind

and now the CEO of the tech firm Inflection AI, says the technology will bring a “huge boost” to the economy and society, but the industry also needs to eliminate potential harms. “Can we do it? It’s a steep challenge, a narrow path, but with enough focus, I believe we absolutely can.” Even the most alarmed experts have become more optimistic over the past week. Max Tegmark, the MIT professor behind an influential letter calling for a six-month pause in giant AI experiments, was at Bletchley. He says: “[The summit] has actually made me more optimistic. It really has superseded my expectations.” Some pessimists remain unconvinced by the summit’s achievements, such as an agreement on government involvement in testing new AI systems. Noortje Marres, a professor and expert in digital sociology at the University of Warwick, says the summit’s communique did not refer to any mechanism for involving citizens and affected groups in the governance of AI. “In other words, AI is not only harmful, it is profoundly undemocratic,” she says. Andreessen, who describes AI as “our alchemy, our Philosopher’s Stone”, perhaps finds himself more of an outlier after events at Bletchley this week. If there is optimism, governments attending the summit would like it to be qualified. Prof Dame Muffy Calder, the head of the college of science and engineering at the University of Glasgow, says risks such as disinformation and biases in the data sets that systems are trained upon need to be addressed. “I am an optimist. But let’s develop those application areas where we have already seen proven benefit, like health and medicine. And let’s focus on how to tackle the current risks, such as disinformation, right now. These are the foundations upon which we can build the future of AI.”