

# Who is attending Sunak's AI safety summit – and what will they discuss?

Publication Date: 2023-10-30

Author: Dan Milmo

Section: Technology

Tags: Artificial intelligence (AI), Computing, Rishi Sunak, explainers

Article URL: <https://www.theguardian.com/technology/2023/oct/30/who-is-attending-sunaks-ai-safety-summit-and-what-will-they-discuss>



Global leaders, tech executives and experts – including Elon Musk – are gathering on Wednesday and Thursday at Bletchley Park, the home of second world war codebreakers, for a landmark summit on safety in artificial intelligence. In a speech last week Rishi Sunak said AI – the term for computer systems that can perform tasks typically associated with intelligent beings – brought opportunities but also significant risks, such as making it easier for rogue actors to make chemical or biological weapons. Here we answer your questions about the summit. What will the summit discuss? The AI safety summit will look at frontier AI systems, which the government describes as “highly capable” models that can perform a wide variety of tasks matching or exceeding the performances of the most advanced AI available today. An example of frontier AI, according to a government document released last week, is the “large language model” technology that underpins AI tools such as the ChatGPT chatbot and its Google-made rival, Bard. But the main concern is the future: the power of models being released next year and in years to come. How can they be tested and monitored to ensure they do not cause harm? Who will be attending? Sunak, the UK prime minister, and the technology secretary, Michelle Donelan, will be there, along with Kamala Harris, the US vice-president, the European Commission president, Ursula von der Leyen, and the Italian prime minister, Giorgia Meloni. However, other world leaders, including the US president, Joe Biden, the French president, Emmanuel Macron, and the Canadian prime minister, Justin Trudeau, have decided not to make the trip. Downing Street said on Monday the rejections did not count as a “snub” to the prime minister, adding that he was pleased at the levels of attendance from governments, the industry and civil society. The UK still does not know whether anyone from the Chinese government will attend, having made a public push for senior officials to do so. The tech industry will be represented by executives from companies including Google’s AI unit Google DeepMind, the ChatGPT developer OpenAI, and Mark Zuckerberg’s Meta. Meta’s president of global affairs, the former UK deputy prime minister Nick Clegg, will be attending. Experts attending include two of the three “godfathers” of modern AI, Geoffrey Hinton and Yoshua Bengio, who are concerned about the pace of development in AI and believe the risk of extinction from the technology is on a par with the threat from pandemics and nuclear war. However, their fellow “godfather”, Yann LeCun – now chief AI scientist at Meta – has described fears that AI could wipe out humanity as “preposterous”. What is on the agenda? Day one of the summit will look at risks such as national security threats and the existential-level threat of systems evading human control. However, in a nod to those who have warned that long-term threats are overshadowing immediate problems such as AI-made deepfakes, there will be a discussion of “issues including election disruption, erosion of social trust, and exacerbating global inequalities”. There will also be some

discussion of the positive side of AI, such as its potential use in education. Separately, Harris will deliver a speech setting out the Biden administration's approach to AI in more detail. British officials insist they do not see this as a distraction from the summit, with Harris and Sunak due to meet for dinner on Wednesday night. The White House outlined its regulatory stance on AI on Monday by publishing an executive order that included requiring companies to share safety test results with the US government before releasing their AI models to the public. The second day will involve Sunak convening a smaller group of foreign governments, companies and experts to discuss what concrete steps can be taken to address AI safety risks. Sunak has already said he will call for an AI equivalent of the Intergovernmental Panel on Climate Change, which would produce an annual report on developments in the technology and associated risks. What is the summit likely to achieve? It will not produce a formal regulatory body on AI. But Sunak hopes it will produce a consensus on the risks posed by unrestricted AI development and the best way to mitigate them. For instance, officials are attempting to thrash out a communique on the nature of AI risks, with an initial draft reportedly referring to AI's potential for causing "catastrophic harm". Officials hope they will be able to secure an agreement from one or more AI developers that they will slow down their development of frontier AI at the very least. They believe that having all the major AI companies represented at the same forum may increase the pressure on them to act in tandem. Sunak intends for this to be the first in a series of regular international AI summits, following the template set by G7, G20 and Cop conferences. If he is voted out of office next year, he may not get to attend another, but if they do continue, they could be one of his most lasting legacies.