

Countries at a UK summit pledge to tackle AI's potentially 'catastrophic' risks

November 1 2023, by KELVIN CHAN and JILL LAWLESS



US Vice President Kamala Harris delivers a policy speech on the Biden-Harris Administration's vision for the future of Artificial Intelligence (AI), at the US Embassy in London, Wednesday, Nov. 1, 2023. Harris is on a two day visit to England to attend the AI Summit at Bletchley Park. Credit: AP Photo/Kin Cheung

Delegates from 28 nations, including the U.S. and China, agreed Wednesday to work together to contain the potentially "catastrophic" risks posed by galloping advances in artificial intelligence.

The first international AI Safety Summit, held at a former codebreaking spy base near London, focused on cutting-edge "frontier" AI that some scientists warn could pose a risk to humanity's very existence.

British Prime Minister Rishi Sunak said the declaration was "a landmark achievement that sees the world's greatest AI powers agree on the urgency behind understanding the risks of AI—helping ensure the long-term future of our children and grandchildren."

But U.S. Vice President Kamala Harris urged Britain and other countries to go further and faster, stressing the transformations AI is already bringing and the need to hold tech companies accountable—including through legislation.

In a speech at the U.S. Embassy, Harris said the world needs to start acting now to address "the full spectrum" of AI risks, not just existential threats such as massive cyberattacks or AI-formulated bioweapons.

"There are additional threats that also demand our action, threats that are currently causing harm and to many people also feel existential," she said, citing a senior citizen kicked off his health care plan because of a faulty AI algorithm or a woman threatened by an abusive partner with deep fake photos.



Tesla and SpaceX's CEO Elon Musk attends the first plenary session on of the AI Safety Summit at Bletchley Park, on Wednesday, Nov. 1, 2023 in Bletchley, England. Digital officials, tech company bosses and researchers are converging Wednesday at a former codebreaking spy base near London to discuss and better understand the extreme risks posed by cutting-edge artificial intelligence. Credit: Leon Neal/Pool Photo via AP

The AI Safety Summit is a labor of love for Sunak, a tech-loving former banker who wants the U.K. to be a hub for computing innovation and has framed the summit as the start of a global conversation about the safe development of AI.

Harris is due to attend the summit on Thursday, joining government officials from more than two dozen countries including Canada, France,

Germany, India, Japan, Saudi Arabia—and China, invited over the protests of some members of Sunak's governing Conservative Party.

Getting the nations to sign the agreement, dubbed the Bletchley Declaration, was an achievement, even if it is light on details and does not propose a way to regulate the development of AI. The countries pledged to work toward "shared agreement and responsibility" about AI risks, and hold a series of further meetings. South Korea will hold a mini virtual AI summit in six months, followed by an in-person one in France a year from now.

China's Vice Minister of Science and Technology, Wu Zhaohui, said AI technology is "uncertain, unexplainable and lacks transparency."



A delegate takes a selfie with Tesla and SpaceX's CEO Elon Musk during the first plenary session on of the AI Safety Summit at Bletchley Park, on Wednesday, Nov. 1, 2023 in Bletchley, England. Digital officials, tech company bosses and researchers are converging Wednesday at a former codebreaking spy base near London to discuss and better understand the extreme risks posed by cutting-edge artificial intelligence. Credit: Toby Melville/Pool Photo via AP

"It brings risks and challenges in ethics, safety, privacy and fairness. Its complexity is emerging," he said, noting that Chinese President Xi Jinping last month launched the country's Global Initiative for AI Governance.

"We call for global collaboration to share knowledge and make AI technologies available to the public under open source terms," he said.

Tesla CEO Elon Musk is also scheduled to discuss AI with Sunak in a conversation to be streamed on Thursday night. The tech billionaire was among those who signed a statement earlier this year raising the alarm about the perils that AI poses to humanity.

European Commission President Ursula von der Leyen, United Nations Secretary-General Antonio Guterres and executives from U.S. artificial intelligence companies such as Anthropic, Google's DeepMind and OpenAI and influential computer scientists like Yoshua Bengio, one of the "godfathers" of AI, are also attending the meeting at Bletchley Park, a former top secret base for World War II codebreakers that's seen as a birthplace of modern computing.



Britain's Michelle Donelan, Secretary of State for Science, Innovation and Technology, 6th right front row, with Digital Ministers who are attending the AI Safety Summit in Bletchley Park, Milton Keynes, England, Wednesday, Nov. 1, 2023. Credit: AP Photo/Alastair Grant

Attendees said the closed-door meeting's format has been fostering healthy debate. Informal networking sessions are helping to build trust, said Mustafa Suleyman, CEO of Inflection AI.

Meanwhile, at formal discussions "people have been able to make very clear statements, and that's where you see significant disagreements, both between countries of the north and south (and) countries that are more in favor of open source and less in favor of open source," Suleyman told reporters.

Open source AI systems allow researchers and experts to quickly discover problems and address them. But the downside is that once an open source system has been released, "anybody can use it and tune it for malicious purposes," Bengio said on the sidelines of the meeting.

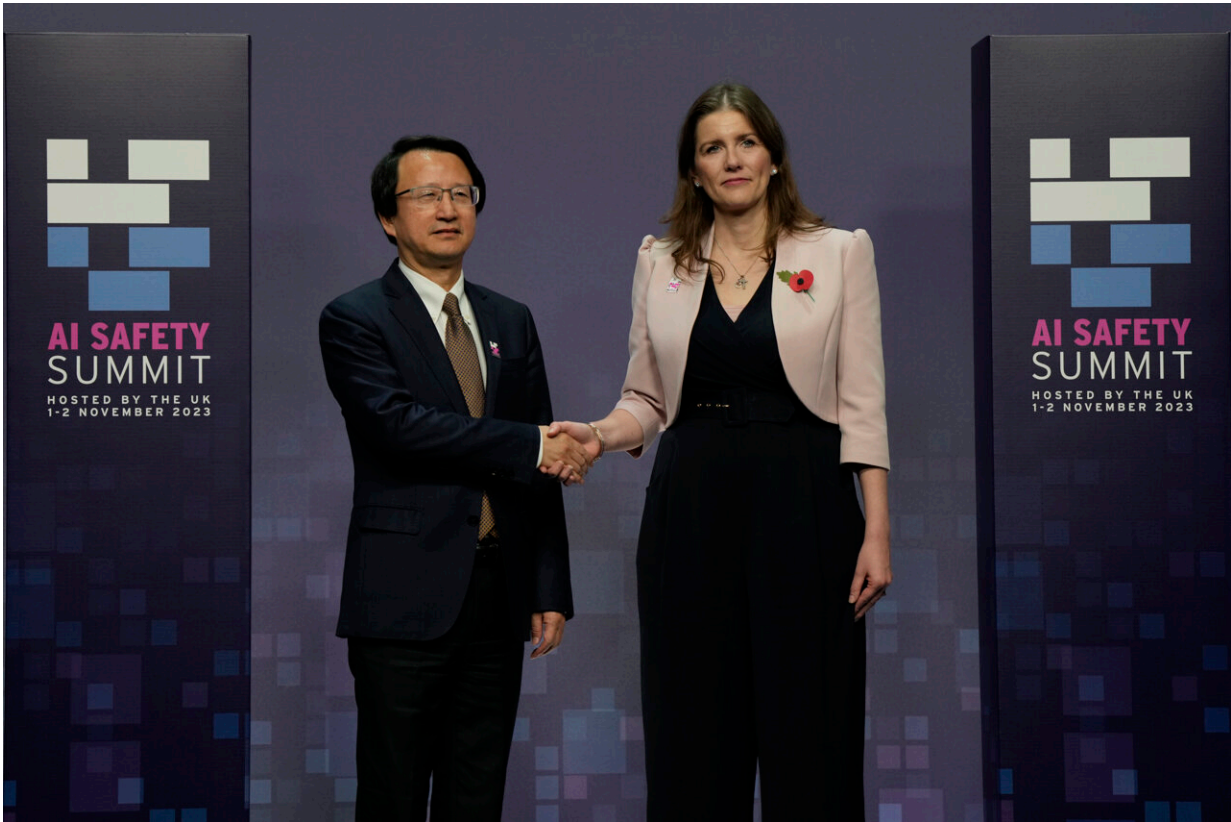
"There's this incompatibility between open source and security. So how do we deal with that?"



US Vice President Kamala Harris, with husband Second Gentleman Douglas Emhoff, arrives at Stansted Airport for her visit to the UK to attend the AI safety summit at Bletchley Park, in Stanstead, England, Tuesday, Oct. 31, 2023. Credit: Joe Giddens/PA via AP



Britain's Michelle Donelan, Secretary of State for Science, Innovation and Technology, right, shakes hands with Indonesia's Nezar Patria, Deputy Minister of Communications and Informatics as he arrives at the AI Safety Summit in Bletchley Park, Milton Keynes, England, Wednesday, Nov. 1, 2023. Credit: AP Photo/Alastair Grant



Britain's Michelle Donelan, Secretary of State for Science, Innovation and Technology, right, and Wu Zhaohui, Chinese Vice Minister of Science and Technology, shake hands prior to the AI Safety Summit in Bletchley Park, Milton Keynes, England, Wednesday, Nov. 1, 2023. Credit: AP Photo/Alastair Grant

Only governments, not companies, can keep people safe from AI's dangers, Sunak said last week. However, he also urged against rushing to regulate AI technology, saying it needs to be fully understood first.

In contrast, Harris stressed the need to address the here and now, including "societal harms that are already happening such as bias, discrimination and the proliferation of misinformation."

She pointed to President Joe Biden's [executive order](#) this week, setting out AI safeguards, as evidence the U.S. is leading by example in developing rules for artificial intelligence that work in the public interest.

Harris also encouraged other countries to sign up to a U.S.-backed pledge to stick to "responsible and ethical" use of AI for military aims.

"President Biden and I believe that all leaders ... have a moral, ethical and social duty to make sure that AI is adopted and advanced in a way that protects the public from potential harm and ensures that everyone is able to enjoy its benefits," she said.

© 2023 The Associated Press. All rights reserved. This material may not be published, broadcast, rewritten or redistributed without permission.

Citation: Countries at a UK summit pledge to tackle AI's potentially 'catastrophic' risks (2023, November 1) retrieved 10 May 2024 from <https://techxplore.com/news/2023-11-countries-uk-summit-pledge-tackle.html>

| |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only. |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|