

US and UK sign agreement for their AI Safety Institutes to test advanced models jointly

Source copy kept by wheresthe.ai for the story "US and UK sign agreement for their AI Safety Institutes to test advanced models jointly" (<https://wheresthe.ai/d/us-gb-aisi-partnership-mou-2024-04/>).

Source: <https://www.gov.uk/government/news/uk-united-states-announce-partnership-on-science-of-ai-safety>

Text saved 18 Oct 2026 (the page as published).

SHA-256 of the saved text: 4b35201996fe1d315e1a29b143b88fbebfe496de489713d29c191321a37c613d

Press release

UK & United States announce partnership on science of AI safety

The UK and US have signed a Memorandum of Understanding which will see them work together to develop tests for the most advanced AI models.

From:

Department for Science, Innovation and Technology , AI Safety Institute and The Rt Hon Michelle Donelan

Published: 2 April 2024

2022 to 2024 Sunak Conservative government ">

This was published under the 2022 to 2024 Sunak Conservative government

UK and US AI Safety Institutes to work seamlessly with each other, partnering on research, safety evaluations, and guidance for AI safety

Institutes to develop shared capabilities through information-sharing, close cooperation, and expert personnel exchanges

The UK and US have signed (Monday 1 April 2024) a Memorandum of Understanding (MOU) which will see them work together to develop tests for the most advanced artificial intelligence (AI) models, following through on commitments made at the AI Safety Summit last November.

Signed by Technology Secretary Michelle Donelan and US Commerce Secretary Gina Raimondo, the partnership will see both countries working to align their scientific approaches and working closely to accelerate and rapidly iterate robust suites of evaluations for AI models, systems, and agents.

The UK and US AI Safety Institutes have laid out plans to build a common approach to AI safety testing and to share their capabilities to ensure these risks can be tackled effectively. They intend to perform at least one joint testing exercise on a publicly accessible model. They also intend to tap into a collective pool of expertise by exploring personnel exchanges between the Institutes.

The partnership will take effect immediately and is intended to allow both organisations to work seamlessly with one another. AI continues to develop rapidly, and both governments recognise the need to act now to ensure a shared approach to AI safety which can keep pace with the technology's emerging risks. As the countries strengthen their partnership on AI safety, they have also committed to develop similar partnerships with other countries to promote AI safety across the globe.

UK Secretary of State for Science, Innovation, and Technology, Michelle Donelan said:

This agreement represents a landmark moment, as the UK and the United States deepen our enduring special relationship to address the defining technology challenge of our generation.

We have always been clear that ensuring the safe development of AI is a shared global issue. Only by working together can we address the technology's risks head on and harness its enormous potential to help us all live easier and healthier lives.

The work of our two nations in driving forward AI safety will strengthen the foundations we laid at Bletchley Park in November, and I have no doubt that our shared expertise will continue to pave the way for countries tapping into

AI's enormous benefits safely and responsibly.

United States Secretary of Commerce, Gina Raimondo said:

AI is the defining technology of our generation. This partnership is going to accelerate both of our Institutes' work across the full spectrum of risks, whether to our national security or to our broader society. Our partnership makes clear that we aren't running away from these concerns – we're running at them. Because of our collaboration, our Institutes will gain a better understanding of AI systems, conduct more robust evaluations, and issue more rigorous guidance.

By working together, we are furthering the long-lasting special relationship between the U.S. and UK and laying the groundwork to ensure that we're keeping AI safe both now and in the future.

The UK and the United States have always been clear that ensuring the safe development of AI is a shared global issue. Reflecting the importance of ongoing international collaboration, today's announcement will also see both countries sharing vital information about the capabilities and risks associated with AI models and systems, as well as fundamental technical research on AI safety and security. This will work to underpin a common approach to AI safety testing, allowing researchers on both sides of the Atlantic—and around the world—to coalesce around a common scientific foundation.

Share this page

The following links open in a new tab

Share on Facebook

(opens in new tab)

Share on X

(opens in new tab)