

UK renames its AI Safety Institute the AI Security Institute

Source copy kept by wheresthe.ai for the story "UK renames its AI Safety Institute the AI Security Institute" (<https://wheresthe.ai/d/gb-ai-security-institute-renamed-2025-02/>).

Source: <https://www.gov.uk/government/news/tackling-ai-security-risks-to-unleash-growth-and-deliver-plan-for-change>

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

SHA-256 of the saved text: f434f9dbd41df7406c5c99060b479a5a3b478c3f9329601179e80f7ddb6cb82d

Press release

Tackling AI security risks to unleash growth and deliver Plan for Change

UK's AI Safety Institute becomes 'UK AI Security Institute'.

From:

Department for Science, Innovation and Technology , AI Security Institute and The Rt Hon Peter Kyle MP

Published: 14 February 2025

2024 to 2026 Starmer Labour government ">

This was published under the 2024 to 2026 Starmer Labour government

UK's AI Safety Institute becomes 'UK AI Security Institute' - strengthening protections against the risks AI poses to national security and crime

Institute bolstered by new criminal misuse team, partnering with the Home Office, to research a range of crime and security issues which could harm UK citizens

New agreement reached with AI giant Anthropic on AI opportunities to help grow the economy as part of our Plan for Change

Safeguarding Britain's national security - a key pillar of the government's Plan for Change - and protecting citizens from crime - will become founding principles of the UK's approach to the responsible development of artificial intelligence from today (Friday 14 February), as the Technology Secretary sets out his vision for a revitalised AI Security Institute in Munich.

Speaking at the Munich Security Conference and just days after the conclusion of the AI Action Summit in Paris, Peter Kyle has today recast the AI Safety Institute the ' AI Security Institute'. This new name will reflect its focus on serious AI risks with security implications, such as how the technology can be used to develop chemical and biological weapons, how it can be used to carry out cyber-attacks, and enable crimes such as fraud and child sexual abuse.

The Institute will also partner across government, including with the Defence Science and Technology Laboratory , the Ministry of Defence's science and technology organisation, to assess the risks posed by frontier AI .

As part of this update, the Institute will also launch a new criminal misuse team which will work jointly with the Home Office to conduct research on a range of crime and security issues which threaten to harm British citizens.

One such area of focus will be the use of AI to make child sexual abuse images, with this new team exploring methods to help to prevent abusers from harnessing the technology to carry out their appalling crimes. This will support work announced earlier this month to make it illegal to own AI tools which have been optimised to make images of child sexual abuse.

This means the focus of the Institute will be clearer than ever. It will not focus on bias or freedom of speech, but on advancing our understanding of the most serious risks posed by the technology to build up a scientific basis of evidence which will help policymakers to keep the country safe as AI develops. To achieve this, the Institute will work alongside wider government, the Laboratory for AI Security Research (LASR), and the national security community; including building on the expertise of the National Cyber Security Centre (NCSC), the UK's national technical authority for cyber security, including AI .

The announcement comes just weeks after the government set out its new blueprint for AI to deliver a decade of national renewal, harnessing the technology to deliver on the Plan for Change. A revitalised AI Security Institute will ensure we boost public confidence in AI and drive its uptake across the economy so we can unleash the economic growth that will put more money in people's pockets.

Secretary of State for Science, Innovation, and Technology, Peter Kyle said:

The changes I'm announcing today represent the logical next step in how we approach responsible AI development – helping us to unleash AI and grow the economy as part of our Plan for Change.

The work of the AI Security Institute won't change, but this renewed focus will ensure our citizens – and those of our allies - are protected from those who would look to use AI against our institutions, democratic values, and way of life.

The main job of any government is ensuring its citizens are safe and protected, and I'm confident the expertise our Institute will be able to bring to bear will ensure the UK is in a stronger position than ever to tackle the threat of those who would look to use this technology against us.

As the AI Security Institute bolsters its security focus, the Technology Secretary is also taking the wraps off a new agreement which has been struck between the UK and AI company Anthropic.

This partnership is the work of the UK's new Sovereign AI unit, and will see both sides working closely together to realise the technology's opportunities, with a continued focus on the responsible development and deployment of AI systems.

This will include sharing insights on how AI can transform public services and improve the lives of citizens, as well as using this transformative technology to drive new scientific breakthroughs. The UK will also look to secure further agreements with leading AI companies as a key step towards turbocharging productivity and speaking fresh economic growth – a key pillar of the government's Plan for Change.

Chair of the AI Security Institute Ian Hogarth said:

The Institute's focus from the start has been on security and we've built a team of scientists focused on evaluating serious risks to the public.

Our new criminal misuse team and deepening partnership with the national security community mark the next stage of tackling those risks.

Dario Amodei, CEO and co-founder of Anthropic said:

AI has the potential to transform how governments serve their citizens. We look forward to exploring how Anthropic's AI assistant Claude could help UK government agencies enhance public services, with the goal of discovering new ways to make vital information and services more efficient and accessible to UK residents.

We will continue to work closely with the UK AI Security Institute to research and evaluate AI capabilities in order to ensure secure deployment.

Today's reset for the AI Security Institute comes just weeks after the UK government kickstarted the year by setting out a new blueprint for AI to spark a decade of national renewal.

Thanks to the work of the Institute, the UK now stands ready to fully realise the benefits of the technology while bolstering our national security as we continue to harness the age of AI .

Further Information:

The agreement between the UK and Anthropic on AI opportunities .

DSIT media enquiries

Email press@dsit.gov.uk

Monday to Friday, 8:30am to 6pm 020 7215 3000

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)