

UK government scales AI use for cyber defence and prepares for autonomous agent incidents

Source copy kept by wheresthe.ai for the story "UK government scales AI use for cyber defence and prepares for autonomous agent incidents" (<https://wheresthe.ai/d/gb-uk-government-scales-ai-use-for-cyber-defence-and-prepare-2026-10/>).

Source: <https://www.gov.uk/government/news/britain-steps-up-national-defences-against-growing-ai-risks>

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

SHA-256 of the saved text: d8f81309c332f6541a74d53b5bb6e7fbb10e32b26d67977d81b0a2be3cb0a0a9

Press release

Britain steps up national defences against growing AI risks

Britain will be better protected from the risks posed by increasingly powerful AI by a suite of measures.

From:

Cabinet Office , AI Security Institute , National Cyber Security Centre and Department for Digital, Culture, Media and Sport

Published: 12 October 2026

Government to put frontier AI to work strengthening cyber defences, building on pilots that have already identified and addressed over 1,000 vulnerabilities.

Stronger protections against emerging AI threats to be considered through the Cyber Security and Resilience Bill regime, such as how serious AI incidents and near misses can be identified and reported.

New work underway to prepare Britain for serious incidents involved autonomous AI agents.

Rigorous evaluations of the latest generation of models underway through the AI Security Institute.

Britain will be better protected from the risks posed by increasingly powerful AI by a suite of measures, with the government urgently putting frontier AI to work strengthening cyber defences, examining stronger legal protections and preparing for serious incidents involving autonomous AI systems.

Announced today (Monday 12 October), the move builds on two pilots that have already helped identify and fix over a thousand potential cyber vulnerabilities across government systems, and detect potential cyber threats. Extending the programme will help protect the public services people rely on every day to reduce risk in the short-term, and will support longer-term adoption across the public sector.

Speaking at Imperial College London's science and tech incubator, Base One, AI Minister Kanishka Narayan argued for a whole-society approach to resilience in the face of emerging AI risks, so that this technology continues to serve British people's interests and boost our prosperity, as well as our security.

The Minister also warned that as frontier AI systems become increasingly powerful and capable of acting autonomously, the potential consequences of failures or malicious use could become far more serious. He will argue that Britain must act now to strengthen its defences and prepare for future serious incidents.

The Minister also confirmed that scientists at Britain's AI Security Institute are now conducting rigorous evaluations of the latest generation of advanced AI models, following work with leading AI developers and the US Government.

As AI becomes more capable and used more widely, the government is taking further steps to ensure the UK is prepared for the potential risks and opportunities it can bring - so British people remain protected as the transformational technology continues to evolve, and so the UK can harness its potential to drive good growth and reindustrialisation.

AI Minister Kanishka Narayan said:

I want Britain to become one of the hardest countries in the world to attack using artificial intelligence.

We will use every power at our disposal. We will legislate wherever necessary. And we won't look to law alone. We will look to British ingenuity: our scientific expertise, technological capabilities and institutions to understand these increasingly powerful systems, defend against emerging threats and respond when things go wrong.

And to the companies developing the world's most powerful AI systems, my message is: Listen when your scientists warn of serious dangers. Investigate unexpected behaviour. And if you cannot adequately control the risks, do not release the system.

Cyber Minister, Liz Lloyd said:

As cyber incidents continue to disrupt globally, organisations large and small, private and public, are all being confronted with an undeniable reality - it's now business-critical to ensure you're on the front-foot when it comes to cyber security.

Through scaling up projects piloting AI to protect public services and progressing a Cyber Bill to boost our defences, government is putting cyber security and AI preparedness at the heart of the UK's digital future.

Alongside these measures, the government will explore strengthening parts of the Cyber Security and Resilience Bill and the wider regime, to ensure the UK's essential services such as hospitals and utilities are better protected from emerging AI related risk. This includes AI loss of control risk, where autonomous AI systems evade human oversight.

Work is also underway with the National Cyber Security Centre and the AI Security Institute to look at how incidents involving AI agents - which can perform complex tasks with little human supervision - should be responded to. A major incident involving AI agents could play out quicker and wider than other major incidents that governments are designed to respond to, and rising to this challenge will require the development of new expertise and capabilities.

In addition to this, the government will look at how near-miss incidents can be identified and reported as well as ensuring lessons can be learned quickly to reduce the risk of future harm. The detail of any measures will be worked through with parliamentarians, as well as technical experts and the operators who would need to deliver them.

The emerging field of AI verification will also be an important part of efforts to reduce serious AI risks. It helps check that claims about how AI systems have been developed, tested and assessed stands up to scrutiny. With world leading expertise and some of the sector's leading organisations based in the UK, Britain is well placed to play a leading role in this work. More details on support for UK AI verification efforts will be set out in the coming months.

These measures are the first steps on a longer journey Britain needs to take to seize AI's potential as a driver for good growth and better public services, while at the same time being resilient to its potential risks. Wider work is being spearheaded by the Prime Minister's AI Taskforce. In the coming months the Taskforce will focus on the areas most impacted by AI; from national security to economic growth and the future of our public services and workforce.

As the Prime Minister set out at the UN General Assembly in September, this government will do whatever is necessary to protect British people from AI risks as they arise. And next year, AI will be at the heart of the UK's G20 Presidency.

Notes to editors:

Ministers wrote an open letter to businesses earlier this year urging them to take action to strengthen their cyber defences against AI cyber threats. This includes a) taking cyber security seriously at the top of organisations, b) getting the basics right with Cyber Essentials and c) following NCSC advice and signing up to their Early Warning Service.

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)