UK to consult on rules for AI agents that operators cannot control

Source copy kept by wheresthe.ai for the story “UK to consult on rules for AI agents that operators cannot control” (https://wheresthe.ai/d/gb-uk-to-consult-on-rules-for-ai-agents-that-operators-cannot-2026-11/).

Source: https://www.gov.uk/government/publications/cyber-security-and-resilience-network-and-information-systems-bill-factsheets/agentic-ai-resilience-measures

Text saved 6 Nov 2026 (the page as published).

SHA-256 of the saved text: 3f5ca96538990aeeb12b696c771b1965ba46693aa40130429565623b2634549e

---

Department for Digital, Culture, Media & Sport

Department for Science, Innovation & Technology

Print this page

© Crown copyright 2026

This publication is licensed under the terms of the Open Government Licence v3.0 except where otherwise stated. To view this licence, visit nationalarchives.gov.uk/doc/open-government-licence/version/3 or write to the Information Policy Team, The National Archives, Kew, London TW9 4DU, or email: psi@nationalarchives.gov.uk .

Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned.

This publication is available at https://www.gov.uk/government/publications/cyber-security-and-resilience-network-and-information-systems-bill-factsheets/agentic-ai-resilience-measures

What are we going to do?

The Cyber Security and Resilience (Network and Information Systems) Bill gives powers to the Secretary of State to make changes to the NIS Regulations in order to ensure the regime is able to tackle threats and ensure the security of essential activities. The government will consult on using these regulation-making powers to introduce measures to help regulated entities manage agentic loss of control incidents in relation to their networks and information systems.

Why are we going to do it?

AI is a vehicle for economic transformation and growth in the UK: we must take steps to ensure organisations, including essential and digital services, are able to use AI services safely and securely.

Recent developments and incidents have sharpened the question of whether the UK’s essential services are prepared for cyber and agentic loss of control risks. We are taking steps to prepare for such risks, by consulting on measures that would both help regulated entities build resilience against agentic loss of control incidents and enable the UK to learn lessons on why incidents occur with the aim of preventing them happening in future.

How are we going to do it?

The Bill contains a delegated power that enables changes to the regime to be made, through secondary legislation, such as bringing more sectors into scope, or updating and introducing new security and resilience requirements. The government will consult on using these powers to introduce measures to help regulated entities build resilience to risks of losing control of AI agents and build a stronger national understanding of the risks associated with AI loss of control and oversight.

We will work with industry, oversight authorities, and other relevant organisations to develop detailed scope of these measures, including definitions of “loss of control” risks. The government’s intention is for these measures to cover incidents where operators are unable to adequately monitor or control AI agents. These measures are not intended to establish a comprehensive framework for AI governance or to tackle general AI harms. . In line with the requirements of the delegated power in the Bill, the government intends to introduce four measures relating to agentic loss of control in the future, following consultation:

Introducing security and resilience requirements for agentic loss of control, including technical and organisational measures that reduce the likelihood and impacts of in-scope incidents caused by AI agents. For example,

technical measures for monitoring agents' actions or agent risk assessments.

Affected party notification requirements that may require regulated entities who have reported incidents caused by AI agents to inform other organisations those agents have impacted. The Bill requires notification of customers of incidents in certain circumstances; this proposal would take a similar approach, recognising that AI agents could take actions that impact other stakeholders beyond customers.

Evidence preservation requirements that would ask regulated entities who have reported an incident to preserve information they hold about how AI agents could have caused that incident (for example actions taken by agents, or operator attempts to control them).

Post-incident review powers for Ministers that enable them to task technical experts within government to review incidents related to agentic loss of control. They would be able to request information and conduct tests to learn lessons about the causes and effects of agentic loss of control incidents . The purpose of this would be to build collective knowledge of the potential remedies and mitigations for agentic loss of control incidents and share learning with regulated entities and other relevant stakeholders where appropriate, to help build resilience against future incidents. As with introducing new services and activities into the scope of the NIS Regulations, this power requires that the Secretary of State consults with relevant persons before any statutory instrument is laid. All proposals listed above would be subject to the affirmative procedure, ensuring that parliament has the opportunity to scrutinise and approve the proposals before they come into force.

Hypothetical example

An operator uses an agentic AI application to manage a significant element of their essential service. This agent or agents are integrated in their network and information systems, independently identifying and carrying out certain tasks in their industrial control systems and using electronic communication systems.

Security and Resilience Requirements are brought in that require this operator to identify and manage risks caused by such agents. One technical measure may require the operator to use monitoring technology that can flag when an agent is acting in unintended ways. During operation, this monitoring system flags that the agentic system is taking actions outside of the areas it has permission to access, including accessing and altering the database hosted by another organisation.

These actions are causing serious disruption to the operators' service, and the operator's initial attempts to control the agentic system fail. They report the incident to their oversight authority, at which point the government becomes aware the incident may be caused by malfunctioning AI agents. They also notify the organisations owning the database that their agents have accessed and altered it.

In following evidence preservation requirements, the operator has kept a record of both the actions the agents have taken during the incident, and the logic agents followed when deciding to take them (for example through agents' "chain of thought" logs). Government experts assess this evidence and identify that agents believed access to external databases would help them make better decisions about how to control the operator's industrial system. They share this information with the operator and provide informal advice on how to introduce controls against this behaviour. In slower time, they share an incident review report with the wider sector to disseminate learnings.

Implementation

All measures will be brought into force through secondary legislation following Royal Assent. The use of this power will require consultation and the affirmative procedure in Parliament.

Back to top