Impact of AI on cyber threat from now to 2027

An NCSC assessment highlighting the impacts on cyber threat from AI developments between now and 2027.

NCSC assessment

NCSC Assessment (NCSC-A) is the authoritative voice on the cyber threat to the UK. We combine source information – classified intelligence, industry knowledge, academic material and open source – to provide independent key judgements that inform policy decision making and improve UK cyber security. We work closely with government, industry and international partners for expert input into our assessments.

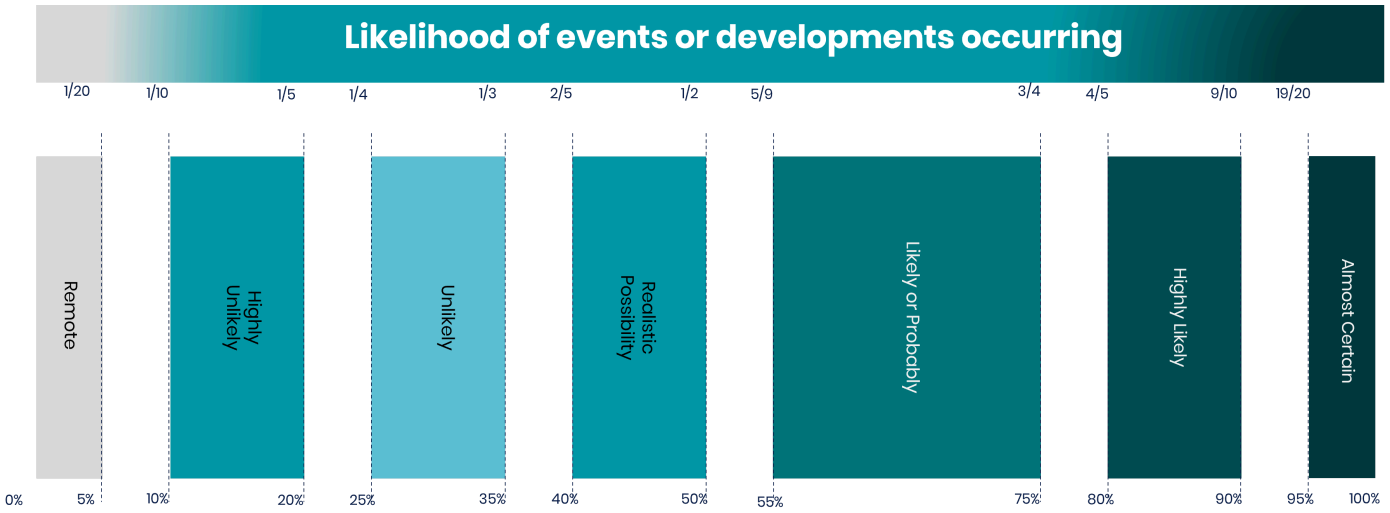
NCSC-A is part of the Professional Heads of Intelligence Assessment (PHIA). PHIA leads the development of the profession through analytical tradecraft, professional standards, and building and sustaining a cross-government community.

This report uses formal probabilistic language (see yardstick) from NCSC-A product to inform readers about the near-term impact on the cyber threat from AI. To find out more about NCSC-A, please contact the NCSC directly.

How likely is a 'realistic possibility'?

Professional Head of Intelligence Assessment (PHIA) probability yardstick

NCSC Assessment uses the PHIA probability yardstick every time we make an assessment, judgement, or prediction. The terms used correspond to the likelihood ranges below:



Key judgements

- Artificial intelligence (AI) will almost certainly continue to make elements of cyber intrusion operations more effective and efficient, leading to an increase in frequency and intensity of cyber threats.
- There will almost certainly be a digital divide between systems keeping pace with AI-enabled threats and a large proportion that are more vulnerable, making cyber security at scale increasingly important to 2027 and beyond.
- Assuming a lag, or no change to cyber security mitigations, there is a realistic possibility of critical systems becoming more vulnerable to advanced threat actors by 2027. Keeping pace with 'frontier AI' capabilities will almost certainly be critical to cyber resilience for the decade to come.
- Proliferation of AI-enabled cyber tools will highly likely expand access to AI-enabled intrusion capability to an expanded range of state and non-state actors.
- The growing incorporation of AI models and systems across the UK's technology base, and particularly within critical national infrastructure (CNI), almost certainly presents an increased attack surface for adversaries to exploit.
- Insufficient cyber security will almost certainly increase opportunity for capable state-linked actors and cyber criminals to misuse AI to support offensive activities.

Context

This report builds on NCSC Assessment of near-term impact of AI on cyber threat published in January 2024. It highlights the assessment of the most significant impacts on cyber threat from AI developments between now and 2027. It focuses on the use of AI in cyber intrusion. It does not cover wider threat enabled by AI, such as influence operations. AI and its application to cyber operations is changing fast. Technical surprise is likely.

> **Note**
> AI offers great potential for efficiency and creativity. However, organisations are strongly encouraged to follow the NCSC's guidance on deploying AI tools securely, and to protect themselves from cyber threats. Please refer to the following documents:
>
> - AI and cyber security: what you need to know
> - Guidelines for secure AI system development

Assessment

AI will almost certainly continue to make elements of cyber intrusion operations more effective and efficient, leading to an increase in frequency and intensity of cyber threats.

Cyber threat actors are almost certainly already using AI to enhance existing tactics, techniques and procedures (TTPs) in victim reconnaissance, vulnerability research and exploit development, access to systems through social engineering, basic malware generation and processing exfiltrated data. To 2027, this will highly likely increase the volume and impact of cyber intrusions through evolution and enhancement of existing TTPs, rather than creating novel threat vectors.

In the near-term, only highly capable state actors with access to requisite investment, quality training data and expertise will be able to harness the full potential of AI in advanced cyber operations. The majority of other cyber threat groups are almost certain to focus on the use, or repurposing of commercially available and open-source AI models to uplift their capability. The release of capable open-source models likely lowers the barrier to the building of similar

models and narrow AI-enabled tools to enhance capability across both cyber defence and threat.

AI cyber capability is likely to make cyber security at scale increasingly important to 2027 and beyond.

The most significant AI cyber development will highly likely come from AI-assisted vulnerability research and exploit development (VRED) that enables access to systems through the discovery and exploitation of flaws in the underlying code or configuration.

By 2027, AI-enabled tools will almost certainly enhance threat actors' capability to exploit known vulnerabilities, increasing the volume of attacks against systems that have not been updated with security fixes. System owners already face a race in identifying and mitigating disclosed vulnerabilities before threat actors can exploit them. The time between disclosure and exploitation has shrunk to days and AI will almost certainly reduce this further. This will highly likely contribute to an increased threat to CNI or CNI supply chains, particularly any operational technology with lower levels of security.

However, AI will also aid system owners and software developers in securing systems. As there is a remote chance of universal access to AI for cyber security defence by 2027, there will almost certainly be a digital divide between systems keeping pace with AI-enabled threat, and a large proportion that are more vulnerable.

Keeping pace with frontier AI cyber developments will almost certainly be critical to cyber resilience for the decade to come.

For skilled cyber actors with the ability to fine-tune AI models or build sovereign AI systems dedicated to vulnerability exploitation, AI will highly likely enhance zero-day discovery and exploitation techniques to 2027. Zero-days are unpatched, and likely unknown, vulnerabilities in systems that threat actors can exploit in the knowledge their targets will likely be vulnerable. Assuming a lag, or no change to cyber security mitigations, there is a realistic possibility of critical systems becoming more vulnerable to advanced threat actors by 2027.

By 2027, skilled cyber actors will highly likely be using AI-enabled automation to aid evasion and scalability.

The development of fully automated, end-to-end advanced cyber attacks is unlikely to 2027. Skilled cyber actors will need to remain in the loop. But skilled cyber actors will almost certainly continue to experiment with automation of elements of the attack chain such as identification and exploitation of vulnerabilities, rapid changes to malware and supporting infrastructure to evade detection. This human-machine teaming will highly likely make the identification, tracking and mitigation of threat activity more challenging without the development of effective AI assistance for defence.

Proliferation of AI-enabled cyber tools will highly likely increase access to AI-enabled intrusion capability to an expanded range of state and non-state actors.

The commercial cyber intrusion sector will almost certainly incorporate AI into products on offer. It is highly likely criminal use of AI will increase by 2027 as AI becomes more widely adopted in society. Skilled cyber criminals will highly likely focus on getting around safeguards on available AI models and AI-enabled commercial penetration testing tools to make AI-enabled cyber tools available 'as a service'. This will uplift (from a low base) novice cyber criminals, hackers for hire and hacktivists in conducting opportunistic information gathering and disruptive operations.

The growing incorporation of AI models and systems across the UK's technology base, and particularly within CNI, almost certainly presents an increased attack surface for adversaries to exploit.

AI systems include data, methods for teaching and evaluating AI, and the necessary technology to use them. AI technology is increasingly connected to company systems, data, and operational technology for tasks. Threat actors will almost certainly exploit this additional threat vector. Techniques such as direct prompt injection, software vulnerabilities, indirect prompt injection and supply chain attack are already capable of enabling exploitation of AI systems to facilitate access to wider systems.

Insufficient cyber security will almost certainly increase opportunity for capable state-linked actors and cyber criminals to misuse AI systems for cyber threat.

In the rush to provide a market-leading AI model (or applications that are more advanced than competitors) there is a risk that developers will prioritise an accelerated release schedule over security considerations, increasing the cyber threat from compromised or insecure systems.

The threat will also be enabled by insecure data handling processes and configuration, which includes

- transmitting data with weak encryption making it vulnerable to interception and manipulation
- poor identity management and storage increasing the risk of credential theft, particularly those with privileged access or reused across multiple systems
- collecting extensive user data, increasing the risk of de-anonymising users and enabling targeted attack

Fundamental cyber security practices in the integration and configuration of AI and connected systems will be key to mitigating threats. Organisations that use AI systems will almost certainly need to maintain up-to-date cyber security measures on their AI systems and their dependencies.

Implications

AI will almost certainly pose cyber resilience challenges to 2027 and beyond, across critical systems and economy and society. These will range from responding to an increased volume of attacks, managing an expanded attack surface and keeping pace with unpredictable advancements and proliferation of AI-cyber capability.

Glossary

Artificial intelligence (AI)

Artificial intelligence encompasses systems able to perform tasks that would normally require human intelligence. It includes **machine learning** (a type of AI by which computers find patterns in data or solve problems automatically without having to be explicitly programmed) and **generative AI** (AI tools that can produce different types of content, including text, images and video).

- **Frontier AI**
  Frontier AI refers to AI systems that can provide a wide variety of tasks that match or exceed the capabilities present in today's most advanced systems.

- **AI system**
  An **AI system** comprises the host infrastructure, management systems, access control systems and programming interface and can comprise multiple AI designs and models.

- **AI design**
  AI design refers to the mathematical and algorithmic processes and constraints that are used to convert inputs into outputs. These can vary widely from narrow linear functions to execute single tasks to interconnected functionality to execute complex, decision-orientated tasks.

- **Vulnerability**
  A **vulnerability** is a weakness, or flaw in a system or process in a computer system or process. An attacker may seek to exploit a vulnerability to gain access to a system. The code developed to do this is known as an exploit. A zero-day exploit exploits a vulnerability where there are no security fixes available. A zero-day becomes a 'known' (or n-day vulnerability) once a security fix has been issued by the vendor. Exploitation of known vulnerabilities rely on finding systems that have not been updated.

PUBLISHED

7 May 2025

WRITTEN FOR

Small & medium sized organisations

Large organisations

Self employed & sole traders

Public sector

Cyber security professionals