

Introducing a new AI chapter in the NCSC research problem book

A revolution is underway at the intersection of AI and cyber security.

The latest frontier AI capabilities are re-writing what is possible for both cyber attackers and defenders, with [recent publications from the NCSC highlighting](#) the opportunities and risks presented by AI.

The Five Eyes cyber security agencies have warned that [AI is accelerating the pace and scale of cyber threats](#), while the [NCSC's Cyber Shield](#) initiative sets out an ambitious vision for harnessing AI to strengthen cyber defence.

Turning this vision into reality will require more than advances in AI capability alone; it will require answers to difficult questions around system design, assurance, security and trust.

To help focus innovation across this new AI-enabled world, we've re-written our [NCSC research problem book to include an entirely new chapter 'AI in cyber security problems'](#).

For years, AI and machine learning have helped defenders analyse data, identify threats and automate routine tasks. The new generation of AI systems (particularly generative AI and agentic AI) can plan, reason, use tools, and take actions with increasing autonomy. This is game-changing, making some things much easier to automate but in the process, creating new problems to overcome.

Many of the assumptions that underpin today's cyber security best practice were developed before this kind of AI existed. Concepts such as vulnerability management, incident response and risk management will need to evolve to account for both AI-enabled systems and AI-enabled attackers. This evolution may require a re-design of processes from the ground up to fully leverage the potential benefits to defenders.

Our [new problem book chapter on 'AI in cyber security'](#) identifies seven areas spanning three broad themes:

- the use of AI to strengthen cyber defence
- the security and assurance of AI systems themselves
- the use of data-driven methods to improve cyber resilience at scale

A common thread runs throughout them all: human-AI collaboration. As AI systems become more capable, we need to understand where autonomy can provide genuine value, where human judgement remains essential, and how meaningful oversight can be maintained without losing the benefits of machine-speed operations.

AI has the potential to fundamentally reshape cyber security. Whether that future is secure, resilient and trustworthy will depend on the choices we make today and the research we undertake tomorrow. The new problem book should focus attention on some of the most important challenges ahead.

Addressing these challenges will require a collective effort across the cyber security and AI communities. We welcome feedback on the challenges identified in the problem book, suggestions for additional research questions, and examples of ongoing or planned research that could help address them. Insights from across academia, industry and government will help ensure that future cyber security approaches remain effective, resilient and trustworthy in an AI-enabled world.

If you would like to contribute to the discussion, please complete [our feedback form](#).

Paul J, Technical Director for Cyber AI Research, NCSC

Brad A, Principal Data Science & AI Researcher, Industry 100 Secondee & Alan Turing Institute Visiting Collaborator

WRITTEN BY



Paul J

Technical Director for Cyber AI Research, NCSC



Brad A

Principal Data Science & AI Researcher, Industry 100 Secondee & Alan Turing Institute Visiting Collaborator

PUBLISHED

20 October 2026

WRITTEN FOR

Cyber security professionals

Large organisations

Public sector

Small & medium sized organisations

PART OF BLOG

Events and initiatives