GCSCC AI Cybersecurity Conference Outcomes Report: Securing the Cyber Future, '*Cyber Resilience in the Age of AI and Geopolitical Uncertainty*'

October 2025

Background

Artificial Intelligence (AI) is no longer a distant promise. It is rapidly being integrated into business operations and everyday life. However, as AI capabilities accelerate, so does the complexity of the cyber threat landscape. Rapid technological and geopolitical changes are converging to create challenges for how we as nations, businesses and citizens protect our core values, rights and digital integrity. We are at an unprecedented turning point for cybersecurity that demands new methods for building an intelligent and resilient digital future.

To help meet this challenge, the Global Cyber Security Capacity Centre (GCSCC), brought together cybersecurity experts from industry, government, academia and the international community for the 2025 AI Cybersecurity Conference in Oxford, UK from 29 September-1 October. Under the theme 'Securing the Cyber Future: Cyber Resilience in the Age of AI and Geopolitical Uncertainty', speakers discussed a range of topics related to AI and cybersecurity governance. Side events showcased projects of the United Kingdom's Laboratory for AI Security Research (LASR) initiative.

The following document summarises the main discussion points.

Managing AI Security Risks at a Global Level

The GCSCC AI Cybersecurity Conference 2025 main day created an opportunity for stakeholders within the AI and cybersecurity communities to connect and discuss collaborative strategies for managing AI cybersecurity risks and opportunities at a global level. Taken together, the sessions facilitated a rich dialogue on AI cybersecurity risks, novel approaches and areas for future collaboration.

Across four sessions, speakers explored global trends for AI cybersecurity and the challenges these trends are producing for different nations in the international environment:

Diverging Approaches to AI and Cybersecurity in Global Governance

- **Global coherence:** Fragmented global debates are hindering the advancement of a consistent and coordinated international approach to AI cybersecurity issues.
- **Regulatory divides:** Growing regulatory divides between major economic centres including the EU, the US, and China bring challenges related to interoperability among different systems or regions.
- **Underdeveloped international standards:** International standards are considered vital for advancing collective security and trust in AI, yet they remain at the early stages of development and adoption.
- **Inclusive governance frameworks:** Non-inclusive or unilateral global AI governance efforts risk marginalising the perspectives of developing countries and producing frameworks that do not consider the challenges faced by these countries in the current digital environment.
- **Disparate development patterns:** Emerging patterns of AI adoption are not following traditional north-south development trends and assessments of AI cybersecurity risks should consider the cross-border implications of these patterns on global systems.

Inequalities of the AI Supply Chain

- **Market concentration:** Extreme concentration of market power in the hands of a select few organisations in two countries introduces new dependencies that some view as threatening to economic and national resilience.
- **Gaps between innovator sand adopters**: Concentrated innovation in a few regions has raised concerns that the introduction of AI is reinforcing global inequalities in digital supply chains.
- **Restrictive capacities**: Limited AI capabilities in some emerging markets are constraining the capacities of these markets to contribute to AI innovation.
- **Independent AI ecosystems**: Efforts are being pursued in nations with the necessary AI capabilities and capacities (e.g. Singapore, India and Brazil) to develop independent AI ecosystems and disrupt existing concentrated market dynamics.

The Evolution of Cybercrime with AI

- **Shifting cybercrime dynamics:** AI is reshaping the nature of cybercrime, including through enhancing harms (i.e. deepfakes, phishing, sextortion) and the scale and frequency of attacks, with global implications.

- **Law enforcement capabilities:** Law enforcement entities around the world are struggling with scalability and coordination, and are generally perceived to not have limited capabilities or capacities to combat cybercrime in the AI era.

Shaping the Future of Global AI Cybersecurity

- **Security culture:** Neither AI safety or cybersecurity risks have not been the driving force of AI development, as organisations prefer to focus on innovation and winning the competition race without adopting security by design principles.
- **Siloed communities:** AI and cybersecurity experts continue to largely operate independently from each other and there are concerns that the separation between these communities may undermine broader cyber resilience efforts.
- **Fragmented international negotiation processes:** multilateral AI and cybersecurity negotiations are occurring independently of each other and risk being misaligned.

Pathways to Advancing Global AI Cybersecurity

Through the sessions participants raised potential next steps for advancing global AI cybersecurity. These include:

- **Foster global dialogue and negotiations:** Inclusive international dialogues and decision-making processes on AI cybersecurity issues that involve broad global participation and consensus building will improve coherence on AI cybersecurity issues and governance approaches.
- **Align international processes:** Connecting international AI and cybersecurity processes will help to facilitate mutually reinforcing processes that complement, rather than retract from each other.
- **Break with binary thinking:** Moving away from the false choices between regulation and innovation, and individualistic and collectivist approaches, will create opportunities for more dynamic and durable AI cybersecurity frameworks.
- **Advance collective security mindset:** Incorporating collective security perspectives into AI safety and security debates and activities, rather than narrow national or clustered approaches, will inform a more holistic understanding of the AI cybersecurity risks and defences.
- **Bridge Silos:** Continuing to break down divides between the AI and cybersecurity community will facilitate greater cross-pollination between these subject areas and ensure AI innovation and security progress harmoniously and concurrently.
- **Design interoperable solutions:** Collaboration between governments, academics, and industry is required to ensure AI cybersecurity solutions are interoperable and transferable between nations, and to avoid the creation of a two-tiered system of 'haves' and have nots'.

- **Build with security foundations:** Security by design remains a foundational principle for technology development that must be carried forward into the development of resilient AI systems.
- **Establish global best practice:** Agreed guidelines of global best practices will improve awareness of risks and assist AI adopters incorporate effective cybersecurity controls.
- **Review rights frameworks:** Greater research into the intersection of AI systems, digital rights and cybersecurity will provide insights into the disruptions new developments are making within this nexus, and how these disruptions can be managed.
- **Strengthen law enforcement networks:** Advanced collaboration between industry and government is crucial for enhancing law enforcement capabilities to meet the scale and complexity of AI enabled cybercrimes.
- **Diffuse innovation:** Undertaking steps to increase the international breadth of AI innovation outside existing countries and organisations will reduce market concentration and associated risks.
- **Lift global education:** Enhancing international collaboration to strengthen AI education and democratise AI knowledge, including through partnerships with industry, will drive greater awareness of and capabilities to address AI cybersecurity risk and opportunities.

AI Cybersecurity Research Insights

Prior to the main day, the conference showcased latest research from LASR and partner institutions in Oxford and beyond. The aim was to advance research development through critical feedback from the wider community and to strengthen collaboration across AI and cybersecurity networks and to promote interdisciplinary engagement.

The day began with a panel discussion examining the **complexity of the AI supply chain**, integrating both technical and policy dimensions. Participants discussed the growing adoption of AI across supply chains, particularly within critical infrastructure sectors. They highlighted that rapid integration, often driven by competitive pressures and constrained by ageing legacy systems, introduces new vulnerabilities. Furthermore, attackers are increasingly exploiting weaknesses in Internet of Things (IoT) devices and synthetic authentication mechanisms. To address these risks, the panel recommended embedding security-by-design principles, establishing baseline security standards, and prioritising the modernisation of legacy systems. Enhancing transparency through initiatives such as AI bills of materials was also seen as a way to strengthen accountability across the supply chain.

The afternoon sessions featured a series of research showcases highlighting new models, emerging threats, and opportunities in the UK's AI security landscape. Researchers underscored the growing complexity of AI-enabled cyber threats, including Subversive Alignment Injection (SAI) and jailbreak prompts that undermine model guardrails. They observed that attackers are

increasingly exploiting AI-generated code, synthetic identities, and adaptive adversarial techniques to evade security controls. Innovative research also explored the use of red teaming large language models (LLMs) to expose bias, though challenges remain around standardisation and cost barriers. The rise of agentic AI further expands attack surfaces, highlighting the need for trust graphs, open identity standards, and zero-trust architectures to mitigate prompt injection and model-context vulnerabilities. Overall, a key takeaway was that AI is amplifying both the speed and sophistication of cyberattacks, underscoring the urgency of developing proactive, AI-specific defence strategies.

The event concluded with and update on the LASR Opportunity Call, a programme that provides the support to build, validate and test new AI security capabilities through funding, mentorship, and access to tooling. A series of presentations focused on work currently being carried out by several selected SMEs in the UK under this programme.

Measuring National AI Readiness

Following the conference main day, attendees were invited to a tabletop exercise on the GCSCC's National AI Cybersecurity Readiness Metric (the Metric). This new tool will enable countries to benchmark their existing readiness and provide an evidence base for future AI cybersecurity decision making.

The event began with government representatives from Mongolia and Cyprus speaking on their experiences undertaking the first trials of the tool in 2025. In a subsequent tabletop exercise, participants were presented with two AI driven cybersecurity scenarios and asked to utilise the Metric to inform their strategic decision-making processes. Through this exercise participants were able to familiarise themselves with the tool and test its applicability, while feeding their insights back into the research and the refinement process of the Metric.

Contact Details

For additional information please contact:

Carolin Weisser Harris, Lead International Operations, Global Cyber Security Capacity Centre

Email: carolin.weisser@cs.ox.ac.uk