Lindy Cameron at Singapore International Cyber Week

The CEO of the NCSC on reshaping cyber security in the era of generative AI.

Opening remarks

Thank you very much for having me.

I'm thrilled to be back here in Singapore for Singapore International Cyber Week.

Let me start by thanking Singapore's Cyber Security Agency for inviting me here today and allowing me to talk to you about a topic that is as fascinating as it is challenging: how we reshape cyber security in the era of generative AI.

The whole field of **artificial intelligence** is developing at pace – and gripping the public's imagination just as quickly as it's developing.

In two weeks, the UK's Prime Minister will host the first major global summit on AI safety, at Bletchley Park in the UK. It will bring together key countries from all over the world to agree on the safety measures that we will need to evaluate and monitor the most significant AI risks, and to realise its full potential to do good.

This is an important moment for the global technology community to come together and set its intentions to make AI work for our people and societies.

We must seize this opportunity to convince governments and industry that in this rapidly-developing arena, security is a crucial foundation on which AI is built. It is really vital to underpin the broader safety, reliability, predictability and ethics of AI systems.

I'm delighted that the cyber security community in the UK has successfully made the case for secure AI: I challenge my counterparts in this room to go away and do the same before the beginning of November!

Now, I'm not going to pre-empt the conclusions of the summit, but I'm confident that the results will be meaningful, and have real impact on how this technology shapes our lives in the years ahead.

Because the UK government's objective – one that we share with other countries, including Singapore where we sit today – to ensure the safe, secure and reliable development and use of AI.

What are we worried about?

So, what are some of the problems that my organisation, the UK's National Cyber Security Centre, is particularly concerned about in this field?

Well, with new technologies come new challenges, and vulnerabilities exist in both frontier and foundation AI models. Adversarial attacks, such as prompt injections or data poisoning attacks, have the potential to do real harm.

On top of this, we are now seeing the decentralisation and democratisation of what was previously very high-end software engineering, which would before have been confined only to those organisations and nation states that could afford it, and make best use of, those skills.

We are now seeing the rapid propagation of these tools and capabilities – just look at how quickly ChatGPT caught the imagination of the general public in countries all over the world.

And there are some positives to take from that decentralisation – like the widespread use of popular large language models – but we should also be clear that there is the ability for open-source or subverted models to be taken advantage of and used for malicious ends, such as spreading disinformation, something many countries will be grappling with next year as national elections come onto the horizon for many of us.

As if that's not enough to keep us awake at night, we are also worried about the whole range of standard cyber security threats, with which we're all already familiar, and which still apply to AI systems.

What's worth worrying about now? What can wait?

That's a pretty imposing to-do list for my teams back in the UK. So where are we focusing? What's worth worrying about now, and what might be a potential risk later down the line?

Well, I could talk to you about a whole series of what-ifs. There is no shortage of press coverage and hand-wringing about the potential existential threat.

But the truth is, none of us knows exactly how AI's capabilities and uses are going to develop and grow: we only know that they will, and that we will need to keep pace. There's a real test here for all of us as a global community, in how we plan for and respond to that challenge.

What do we care about in AI from a cyber security perspective?

So, how can we begin to meet this moment in AI development?

The first thing to say is that, like our cyber partners in the US, here in Singapore and across the international community, the UK advocates a 'secure by design' approach to software: that vendors must take responsibility for embedding cyber security into their technologies, and their supply chains, from the outset and throughout a system's lifecycle.

This will help society and organisations realise the benefits of AI advances but also help to build wider trust that AI is safe and secure to use.

We know, from experience, that security can often be a secondary consideration when the pace of development is high. Much of the digital architecture we rely on today was never designed with security at its heart. It was built on foundations that are flawed and vulnerable, and that has left us a lot of work to do.

And unless we want to repeat these mistakes, AI developers must, to the best of their abilities, predict possible attacks and identify ways to mitigate them. Failure to do so will risk designing vulnerabilities into future AI systems, building another flawed ecosystem, and unwittingly make the future harder for us all.

Alongside this, there's the question of data security: it's really vital that developers properly recognise the value of their training data, and make sure that it's appropriately protected. That protection must be ongoing, and must keep pace as things change.

This is especially important where sensitive data is involved, particularly when we consider the implications of a breach, or misused of information.

And we also need to consider that AI is shifting our perception of what constitutes intellectual property: model weights are now considered IP, and should be secured as tightly as any other vital proprietary information.

Where to from here?

So, how do we begin to grapple with some of these issues?

At the NCSC, we are committed to working in partnership with our international counterparts, as well as the public and private sectors in the UK to realise the benefits provided by AI and ML systems.

We know AI has the potential to **improve cyber security** by dramatically increasing the timeliness and accuracy of threat detection and response.

The NCSC has previously published guidance on machine learning, to help developers gain a better understanding of the threats and vulnerabilities that some systems face so they can make informed decisions at an organisational level.

And we are continuing to work in close collaboration with international partners to ensure our guidance keeps pace with the development of AI technologies, including for non-technical audiences – the boards that make decisions as well

as the developers who create them. We would like to see the wider cyber security community play a role in communicating the cyber security risks and implications of AI to its developers and users, so people can understand and respond to them in meaningful ways.

The AI Safety Summit being held in the UK in November will bring together countries, leading technology organisations, academia, and civil society to discuss and advise on the actions we need to take, both at a national and international level, on the development of AI, and especially at the frontier of that development.

The Summit is a key moment for continuing the conversation on AI safety and security across the whole global community.

We think it will be an important launchpad for discussions over the months and years to come, **when we will be looking to partners to work together in pursuit of an international, industry-led standard** – as well as increased take up – for cyber security within AI systems.

Concluding remarks

Amid the understandable – sometimes excitable – conversation about the potential existential impacts of AI, we cannot miss the real, practical steps that we need to take to secure AI now and for the future, and they can be taken now.

The work will not be easy – but it will be worth the dramatic benefit that AI will bring to our economies and societies. So long as we remember that cyber security is fundamental to secure AI.

And the discussion is going to be an ongoing one. The conversation about AI security won't start or end with the Summit. As the technology develops and the use of AI proliferates in ways we haven't yet grasped, we must work together to stay ahead of the risks.

And we at the NCSC, we will be there to understand the cyber security threats we face in AI and will work closely with our international partners and industry on

how to increase our collective security.

I look forward to working with all of you to ensure we meet the challenges that we face, together.

Thank you.

PUBLISHED

18 October 2023

DATE OF SPEECH

18 October 2023

LOCATION

Singapore International Cyber Week