How AI is Being Weaponised and Cyber Security

This is the second article in our 3-part AI safety series.

Our first article in this series provided information on the transformative insights and outcomes of the AI Safety Summit held at Bletchley Park.

This second article will discuss AI's potential to be used by hackers to carry out devastating cyber-attacks, relevant guidance and practical steps businesses should take to protect themselves in this AI landscape.

What types of AI-enabled cyber-attacks are there?

Prompt injection attacks

A prompt injection attack occurs when a hacker with malevolent intent uses carefully crafted prompts to bypass filters included in large language models to cause harm.

Data poisoning attacks

A data poisoning attack occurs when cyber criminals seize open-source generative AI technology and feed into the model malicious training data to manipulate the data output.

Deepfake attacks

These are a form of synthetic media in which audio or video content is manipulated to make them indistinguishable from real media.

What are the legal consequences of cyber security breaches?

Successful cyber security attacks can have serious consequences for companies that suffer them. They might

Tel: 0344 967 0793 | Email: info@LA-Law.com | www.lesteraldridge.com

result in companies paying heavy fines, regulatory sanctions and investigations and potentially criminal liability. The associated business disruption and the loss of data, intellectual property, and confidential information can also be highly damaging and expensive, both in terms of financial losses and remedial actions. Companies that suffer cyber security breaches may be open to adverse media attention and the immeasurable damage of lost consumer trust and confidence. These outcomes are not easily managed.

What guidance is there on AI and cybersecurity?

On 27 November 2023, the National Cyber Security Centre (NCSC) announced new global cyber security guidelines entitled "Guidelines for Secure AI System Development". The guidelines are aimed at providers of AI systems who have built their own models, or are using models hosted by an organisation or using external application programming interfaces (APIs). Four stages are covering the end-to-end lifecycle of an AI system, including suggested behaviours to enhance the effectiveness of cyber security measures, as follows:

Secure design

This section covers understanding risks and threat modelling and specific topics and trade-offs to consider in system and model design.

Secure development

This section contains guidelines for the development stage of the AI system development life cycle, including supply chain security, documentation, and asset and technical debt management.

Secure deployment

This section contains guidelines for the deployment stage of the AI system development life cycle, including protecting infrastructure and models from compromise, threat or loss, developing incident management processes, and responsible release.

Secure operation and maintenance

This section provides guidelines on actions particularly relevant once a system has been deployed, including logging and monitoring, update management and information sharing.

What should UK businesses do?

Businesses wishing to use or develop AI technology should follow the Guidelines for Secure AI System

Tel: 0344 967 0793 | Email: info@LA-Law.com | www.lesteraldridge.com

Development in conjunction with established cyber security, risk management, and incident response best practices.

Businesses are strongly recommended to arrange appropriate cyber security insurance. Without cyber security insurance, business owners may be left footing the bill for damages that arise from a breach.

The National Cybersecurity Alliance (NCA) reports that one in five small businesses will experience a data breach, with over half of those affected having to close their doors because they cannot recover from the financial consequences.

Our next article in this AI safety series shall cover the risks relating to approved and frequent unapproved use of AI by employees within a company and the consequential need for AI policies to be implemented.

How we can help

For further information on the legal implications of AI and how it may impact your business, please do not hesitate to contact partner Dean Drew at dean.drew@LA-law.com or 0330 0539 759.

Tel: 0344 967 0793 | Email: info@LA-Law.com | www.lesteraldridge.com