ISTARI Perspectives

5 ways generative AI will impact cybersecurity and what to do about them

Abel Archundia
Dr. Michael Zeller

10 • 2023

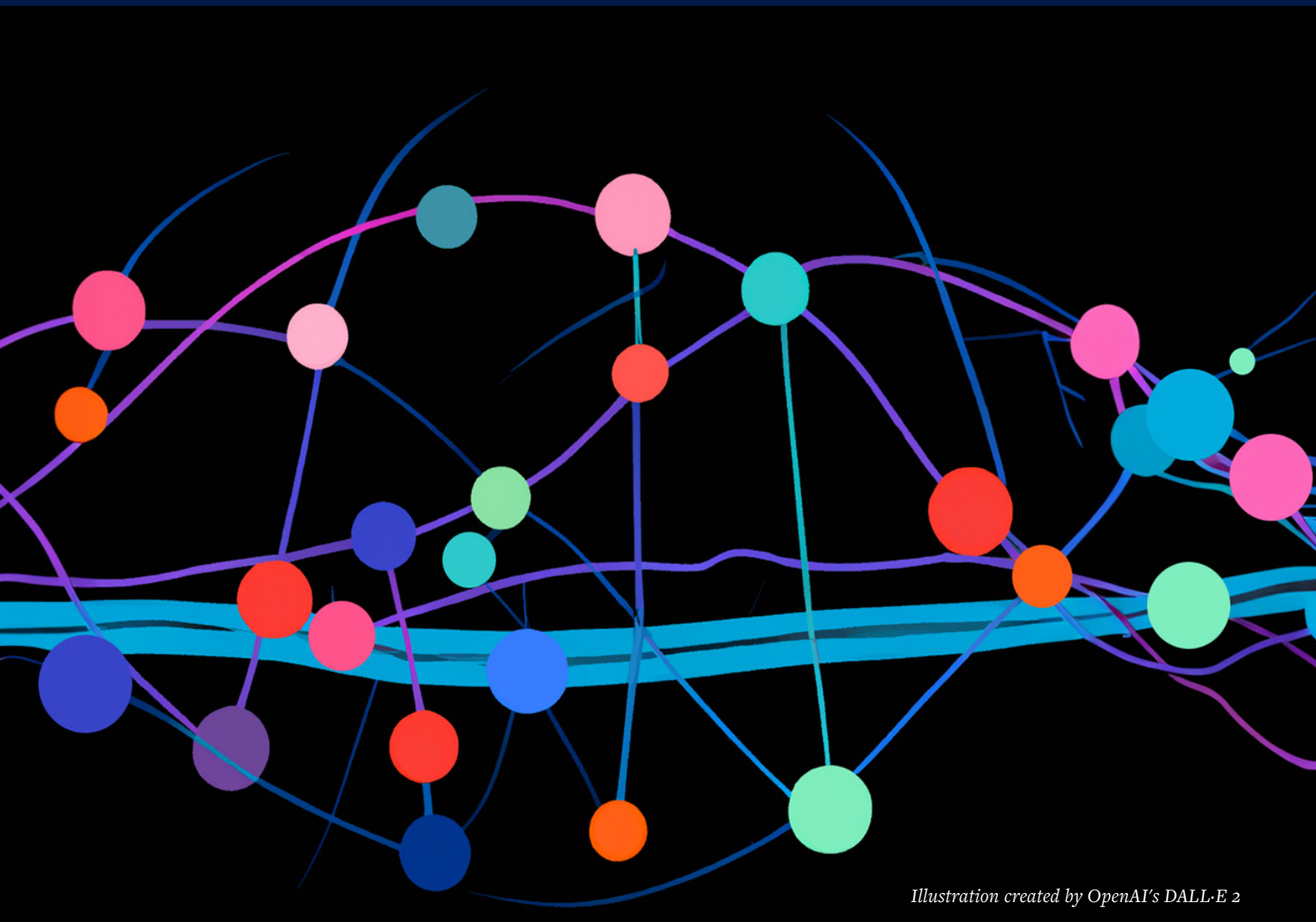

Illustration created by OpenAI's DALL·E 2

5 ways generative AI will impact cybersecurity and what to do about them

When ISTARI hosted its annual CEO forum on cyber resilience in Singapore in February 2023, one CEO asked his peers if they had already used or experimented with ChatGPT. Every CEO in the room raised their hand. This was just four months after ChatGPT was released.

ChatGPT has seen extraordinary growth. In just five days after its release, it cracked the one million-user milestone. By comparison, it took Netflix three and a half years to reach one million users and Spotify five months. Through its mastery of language, ChatGPT has remarkable use cases — from creating a bespoke travel itinerary to brainstorming a birthday present for a spouse to writing software code.

ChatGPT is an example of generative artificial intelligence (AI), which is capable of creating new text, images, audio and videos based on patterns it has learned from existing data. After being pre-trained on large amounts of unlabelled data, GPT foundation models are fine-tuned with supervised learning techniques, as well as reinforcement learning from human feedback.

Other AI and machine learning models power innovation-focused business cases. Research shows companies with a strong digital foundation are better prepared to match advanced digital techniques to business goals. For instance, BCG research reveals that though 83 percent of companies in their Most Innovative Companies ranking have "systematically" invested in AI use cases, only 45 percent have managed to translate it into business impact.

Integrating new technologies into existing business or functional processes is challenging. Common limitations include lack of executive championship, expenses and regulatory complexity. Applying AI must also overcome the hurdles of data-related privacy, identity, access and integrity — the essentials for information security leaders.

The cybersecurity community has started to voice concerns over AI/ML technology. Anyone with an internet-connected computer now can generate code or text that resembles humans, more efficiently and with fewer errors. With the help of AI, bad actors can also tamper with training data sets and convincingly impersonate a trusted identity. Gone are the days of spam emails littered with typos.

The impact of generative AI and other advanced techniques on cybersecurity is profound, but there are positive trends to balance the concerns. Below are five cybersecurity trends in generative AI and three steps cybersecurity leaders can take to address them.

5 cybersecurity trends to watch for generative AI implementation

Identity impersonation (advanced phishing attacks): Generative AI could lead to more convincing phishing emails. For example, by providing the AI with written text examples and some context for an individual, it can recreate style and tone, increasing the risk of successful social engineering or phishing attacks. AI-powered bots can emulate human conversation and interact with individuals online, gather sensitive information and manipulate people into taking malicious actions at scale and with minimal cost. Recent examples are WormGPT and FraudGPT and the AI-generated deepfakes illustrated by Poland's ElevenLabs, which convincingly substituted Leonardo DiCaprio's voice during his famous United Nations-speech with Kim Kardashian's.

Password cracking: Generative AI enables more effective password prediction and cracking, increasing the risk of unauthorised access to sensitive systems and data. Weak or reused passwords become even more vulnerable to exploitation. For example, PassGAN, a machine learning-based AI password cracker, can crack 51 percent of common passwords in less than a minute.

AI-Generated malware: AI can create new, unseen types of malware that can go undetected by security systems. A recent example of AI-powered malware is BlackMamba, an instance of polymorphic malware that constantly evolves and responds to security measures to evade detection. It successfully bypassed industry-leading endpoint detection and response technology.

Data privacy: AI models themselves can be receptive to leaking data if prompts are engineered in a clever way. For example, researchers were able to breach NVIDIA's AI model, NeMo, manipulating the language model into releasing personally identifiable information from the databases in which it was housed. Another risk to data privacy is that employees might provide proprietary company data to language models like ChatGPT in their prompts. But on a positive note, AI can help protect crown jewels by continually monitoring devices and assets and learning from context and threat vectors to cost-efficiently implement zero-trust principles in any environment.

Productivity enablers: Most likely, defenders will embrace AI techniques to help them co-pilot their next best action, enabling teams to multiply their efficiency. AI can assist by cleaning out incidents, events and vulnerabilities requiring intervention from less-severe or repetitive alerts. It also can provide multi-factor checks on identities and optimise the configurations of data safeguards or cloud defence mechanisms with real-time adjustments, similar to how airplane technology helps pilots navigate difficult conditions.

Three things CISOs need to do to protect their organisations:

1. Revise existing policies and develop new guidelines for the use of AI within the company, all while considering the bigger picture. Many of the risks introduced by generative AI often are covered in existing enterprise policies for writing business emails, sharing data with third parties, or using third-party code projects. These policies should be revised with a specific focus on generative AI and proactively stress-tested for the prevalence of tools employees and clients are using regularly. In addition, as companies work on their data sets, CISOs should evaluate if a rules-based mechanism for data clean-up or consolidation can help enforce the organisation's privacy and integrity standards.

2. Ensure continuous monitoring of the integrity of content, brands and digital assets. It's important to deploy technology that surveys the organisation in real time and to use apps to assist in monitoring network intrusions, OS integrity and interactions. This will underpin the ability to scale promising use cases without increasing costs and risk exposure.

3. Stay abreast of the latest AI-related cyber threats and vulnerabilities. Leverage threat intelligence services to understand the tactics and procedures threat actors employ. Mainstream criminals have seen their reach and capability multiplied by open-source AI tooling but their business model — ransom assets for profit — is largely the same because it works. Foiling them is an everyday task and hinges on adopting disciplines and techniques to multiply the capabilities of already stretched teams.

As any technology in early stages of adoption, AI/ML and large language models eventually will become mainstream, like GPS and Wi-Fi. The savvy CISO will view this as a leadership opportunity to help set the agenda for adoption across their enterprises and ensure privacy and model resilience are not compromised as they scale from pilot to everyday use. At stake is laudable business leadership in the digital age.

About the authors:

Abel Archundia

Abel Archundia is Managing Director for Global Advisory and Life Sciences & Industrials at ISTARI. Previously he was Global CIO at Bayer Pharma, and at Novartis' Sandoz Division. He started his career as a consultant at BCG and was later responsible for Dell's business unit in Mexico.

Dr. Michael Zeller

Dr. Michael Zeller is the Head of Artificial Intelligence & Solutions at Temasek and is responsible for the investment company's AI venture, building and accelerating the deployment of AI technologies to create scalable AI products and solutions. Prior to joining Temasek, Dr Zeller was the CEO of Dynam.AI, an AI consultancy and solutions firm. He has over 20 years of experience as an entrepreneur, executive and advisor of technology-centric organisations.