Curbing Cyber Risk with AI-Powered Governance

BY **MARK THAM,** Accenture Singapore Country Managing Director

As the promise of artificial intelligence (AI) unfolds, so does its perilous shadow. AI not only empowers but also endangers, amplifying cyber fraud. Collaboration and vigilance are the greatest weapons against the encroaching darkness of AI-enabled deception.

In 2023, videos showing Prime Minister Lee Hsien Loong purportedly promoting a cryptocurrency scheme started to circulate. This was a barefaced scam: audio and video footage of Mr Lee, taken from a real TV interview, had been manipulated using generative AI to present misleading information to viewers and trick them into investing in the scheme.

PM Lee issued a warning on his Facebook page for Singaporeans to stay vigilant against the pitfalls of AI. This incident cast the spotlight on the proliferation of AI-powered scams including deep fakes, which surged ten-fold between 2022 and 2023 and cost Singapore S$334.5 million in the first half of 2023.

Proliferating cyber risks with AI

While AI empowers humans to reach new heights of innovation and productivity that advance social good, it can also be exploited by malicious actors to perpetuate cyber fraud.

Most AI systems today are focused on assisting in tasks and functions, but will soon evolve into agents that can assist and advise us, as well as take decisive actions on our behalf in both the physical and digital worlds.

The rise of large networks of interconnected AI, or "agent ecosystems" can help humans stretch their

creative potential. But in the wrong hands, AI agents could also potentially enable even more sophisticated and complex cyber schemes.

With AI tools, fraudsters can disseminate scams at speed and scale. In addition to phishing and malware, scammers can potentially deploy AI to prey on victims using misinformation, disinformation, identity theft, data privacy breaches, as well as algorithmic manipulation of social media.

Such scams can be so elaborate that even the most digitally savvy consumers and most vigilant companies risk falling prey to them. Companies' cyber security systems, too, may be hard-pressed to detect and abort the multitudes of threats. JP Morgan, for instance, battles a staggering 45 billion hacking attempts daily – despite hiring 62,000 technologists who mostly work on cyber security.

Corporate leaders are aware of the growing urgency to gird up their defences against cyber attacks on their business and customers. That is why 86 per cent of CEOs surveyed by Accenture in 2023 rate cyber trust and resilience for emerging technologies like generative AI as highly relevant for their organisations.

This will be critical not just for corporate and public interests, but also for Singapore's future as a top financial centre and global business hub. Corporate leaders and board directors must work together to collectively harness a powerful double-edged sword like AI.

AI-powered workforce for scam-busting

The resulting losses and reputational damage from all these cyber risks can severely undermine public trust in the use of AI.

But that does not mean corporate leaders should slow down their AI adoption. Scammers are only going to accelerate AI usage to perpetuate fraud. We need to fight fire with fire by harnessing AI effectively to snuff out attempts to exploit this technology for malicious use.

Board directors will need to tap on industry best practices and insights for cyber resilience from the broader ecosystem of regulators, academic institutions, professional services firms and more.

They also need to ensure both leadership teams, as well as their workforce across the organisation, understand AI, and will be continuously trained to identify new and emerging cyber threats and risks. In an augmented ecosystem which includes AI agents, managers will also need to strategically upskill workers to work with these AI co-pilots.

For example, the Monetary Authority of Singapore is exploring ways to fight money laundering and other financial crimes using AI, which can process and analyse terabytes of data at speed to help flag suspicious activities more quickly. AI can, for example, be applied to COSMIC – Collaborative Sharing of Money Laundering/Terrorism Financing Information and Cases – an upcoming digital platform for financial institutions to share information on suspicious customers or transactions.

Meanwhile, corporate leaders are also expanding their AI-powered security arsenals. AI-powered large language models (LLMs) and quantum computing are increasingly being wielded by companies to fight cyber-crime. For instance, fraud review analysts can deploy LLM-based assistants to expedite decision-making on whether a transaction is fraudulent and use LLM to predict the next transaction of a customer in order to pre-emptively assess fraud risks.

Digital trust is key

For AI to be an effective weapon against cyber fraud, organisations need to ensure digital trust is in place. Digital trust is the expectation by individuals that digital technologies and services – and the organisations providing them – will protect all stakeholders' interests and uphold societal expectations and values.

The World Economic Forum offers a "Digital Trust Framework" (see box) which shows how commitment to cyber security, privacy, transparency, redressability, auditability, fairness, interoperability and safety – when taken together and driven by C-suite leaders – can improve both citizen and consumer trust in technology and the companies that create and use new and emerging technologies.

A responsible future built on collaboration

Combating scams and frauds is a massive undertaking that involves multiple stakeholders and cuts across the entire enterprise and ecosystem. A key aspect to its success is stepping up public education about how scams are constantly evolving and the ability to sift out AI-powered scams.

In Singapore, there have been growing concerted efforts to educate the public about the ever-changing tactics of scammers through various media platforms. The National Crime Prevention Council's Scam Alert website, for example, provides the latest updates on a comprehensive list of scams as well as tips and advice on how to avoid fraud.

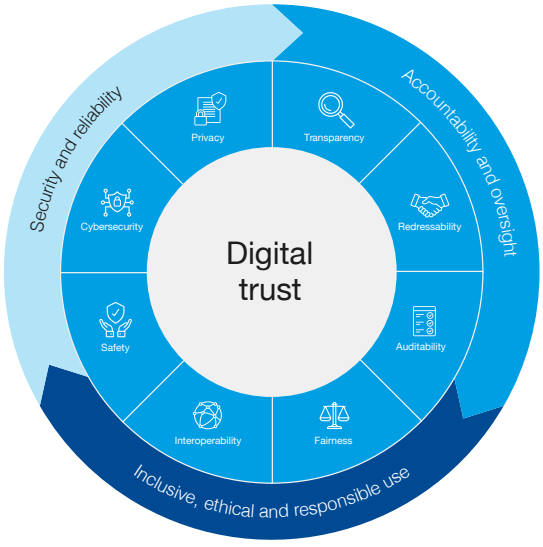
Meanwhile, Singapore regulators are working with financial institutions (FIs) and telecommunication companies (telcos) to increase awareness among the public about digitally-enabled scams, while making clear that consumers have a duty to take precautions.

Under a proposed shared responsibility framework for phishing scams, FIs and telcos have a duty to protect the public. For example, banks must impose a 12-hour cooling off period upon activation of a digital security token during which "high-risk" activities cannot be performed, making it more difficult for scammers to quickly drain their victims' bank accounts, while telcos must implement anti-scam filters over all SMS to block messages with known phishing links.

But the proposed framework also states that consumers are expected to bear the full loss of a phishing scam where the FIs and telcos have both fulfilled their duties.

This puts the onus on consumers alike to take the necessary precautions, while keeping themselves updated on still-emerging regulations to combat scams.

Digital Trust Framework



Source: World Economic Forum

Guardians of the AI-powered future

Cyber resilience starts from the top. Board directors and C-suite leaders must start laying a robust foundation of cyber resilience for their organisations.

To achieve this, they will need to ensure their workforce is constantly trained on evolving cyber risks, equipped to co-pilot with AI to combat scams and ensure digital trust is embedded into their operating systems.

But combating scams and frauds is a shared responsibility for all stakeholders in the ecosystem; hence ensuring consumers and the public are cyber resilient is also crucial. Only then will the future of AI be a secure one for every individual and organisation.