Ensuring Trusted AI

BY **DAVID TOH,** Vice-President, The Institute of Internal Auditors Singapore

AI has evolved from a futuristic concept to an integral part of everyday life, shaping everything from shopping suggestions to workplace productivity. Unlocking AI's full potential requires strategic oversight, responsible risk management and an organisational culture ready to embrace innovation. Boards must guide their organisations towards responsible and impactful AI use.

AI is at a watershed. For business leaders who recognise that transformation is unavoidable for their companies to remain competitive and relevant, the key is determining how, when and where to use AI. This extends to how the technology will change the way the business operates and what the company can do to maintain and establish trust among its key stakeholders.

PwC's 27th *Annual Global CEO Survey* in 2024 suggested that a vast majority of companies are already taking steps towards reinvention. About half of the surveyed CEOs expect generative AI (GenAI) to enhance their ability to build trust with stakeholders over the next year, and about 60 per cent expect it to improve product or service quality. Within the next three years, nearly seven in 10 respondents also anticipate that GenAI would increase competition, drive changes to their business models and demand new skills from their workforce.

Overall, CEOs anticipate positive near-term business impacts from GenAI, such as applications that increase revenues through improved product quality and customer trust as well as those that enhance efficiency.

In the evolving landscape, big tech and venture capitalists are pouring investment into GenAI. And employees are expecting to use GenAI tools in their daily workflows. However, organisations have also witnessed how things can go awry, such as leakage of confidential data to public systems, continued struggles with "hallucinations", deepfakes, manipulated content and systemic AI bias.

Most companies are likely to continue to expand their use of AI and explore the possibilities offered. The board should oversee the reasonableness of management's growing interest in the technology, focusing on how this may impact corporate strategy. In particular, the board should pay attention to how risks, especially mission-critical or reputational risks, are managed across the company while pursuing innovation (see box, "Board Oversight is Critical").

Board Oversight is Critical

In their oversight role, boards should recognise strategic opportunities, oversee risks and controls, and keep up with emerging trends. Here are some ways boards can maintain oversight and ensure they understand the risks.

Strategic opportunities

- **Evaluate risk appetite.** AI offers businesses a wide range of real benefits and solutions at all levels. By identifying and analysing patterns and trends in oceans of data that no person or team could make sense of by themselves, AI is able to facilitate decision-making, perform repetitive tasks and drive faster business solutions. Boards should ensure they understand the prioritised business use cases for the company and evaluate whether they are complementary with other strategic business initiatives. Is the company ready, able and equipped for an AI transformation?

- **Third-party vendors.** As companies think through their strategic priorities, they will likely need to engage and work with leading AI technology providers. Most AI systems are secured through a vendor or customised on top of a cloud platform. Boards should understand whom the company plans to work with and the potential concerns, particularly newer vendors who would have access to confidential company data. What processes does the company have to evaluate the possible risks related to third-party providers?

- **Upskill workforce.** AI opportunities would transform the workplace, enabling knowledgeable employees to do far more at greater speed. Companies would need to upskill their workforce to use these tools effectively in such situations. Furthermore, companies would require new talent acquisition skills as part of their talent strategy. Boards must be involved in this process and consider the social implications and long-term talent development.

Risks and controls for trusted AI

- **Responsible AI framework.** AI has the potential to create huge economic benefits but as business applications grow in number and scope, so do the risks – both operational and reputational. AI systems are built with an incredible amount of public and proprietary content. With respect to publicly available large language models, no company can independently validate whether the large quantum of data used is accurate, unbiased, representative, or even relevant to a specific business need. Boards must consider the responsible use and societal impact of using AI technology and how it aligns with the corporate culture.

- **Policies, procedures and controls.** AI-based risks include data security and privacy. Readily available applications are reducing entry barriers for threat actors and data vulnerability to not only theft and loss, but also to perpetuating biases. Furthermore, companies face legal and reputational risks, with potential exposure to intellectual property and proprietary information. In the absence of appropriate controls, inaccurate output and communication of false data are a real threat. Boards have to consider implications related to legal, privacy, security and ethical concerns.

- **Set the tone.** Boards are expected to set the right tone for responsible AI use and ensure that directors and management are clearly aligned on the topic. Directors should understand whether and how management has developed and implemented proper policies, procedures and controls to reduce possible bias and error in algorithms and data. They should also confirm that systems are not manipulated, inadvertently or deliberately. The primary objective must be developing and designing AI systems' practices and outcomes that can be trusted and aligned with corporate values and overall strategies.

Keeping up with emerging regulations

- **Compliance.** The rapid rise of GenAI has spurred regulators to raise concerns about AI practices. As the technology continues to evolve at an unprecedented rate, governments and regulatory bodies are working to establish guidelines to ensure ethical use and mitigate potential risks. Boards have to oversee that management institutes internal mechanisms to keep up with regulatory changes and remain compliant with new initiatives and regulations.

- **Due diligence.** Regulators worldwide are taking their own approaches to set standards in their respective jurisdictions. The European Union adopted the AI Act in June 2023, imposing more stringent requirements and applying copyright laws to GenAI source content. Other regions are also drafting similar legislation, with some focusing on data privacy and others on fairness in AI decision-making processes. In China, regulators have moved to ban the use of deepfakes, a category of GenAI used to create images, video and audio of real people, often for the purpose of manipulation.

- **Connected ecosystem.** In May 2024, Singapore launched a governance framework for GenAI, that expanded on the existing Model Governance Framework covering traditional AI. The new framework aims to set forth a systematic and balanced approach to address GenAI concerns while facilitating innovation. Covering multiple dimensions, the framework aims to support a comprehensive and trusted AI ecosystem, ensuring that AI's decisions are explainable, transparent and fair.

Trusted AI systems

As AI-based applications become more powerful, the burden on companies to reap the benefits and manage the risks of AI is expected to grow dramatically.

To govern an organisation's AI strategy and practice, companies would have to define clearly what the AI models are intended for and how they perform. Building trust in these systems would require management to establish processes to test and evaluate AI systems effectively. Existing functions such as data privacy, IT, data and security governance, risk management and internal audit may have to be engaged at various points to build this trust.

The central feature of AI is that it discovers intricate relationships from within complex data. However, human intervention would have to be mandated to confirm that the system, fed with relevant, appropriate and reliable data, consistently produces results within the expected range of acceptable outcomes.

AI will continue to provide companies with opportunities for greater applications and innovation, additional risks, cost-benefit considerations and regulatory compliance. As companies determine how to utilise the technology responsibly and build trust in their applications, boards must keep a firm hold on the rudder.