ACCOUNTABILITY

Cracking the code: Tackling the complexities of AI governance

BY **SREEREMA BANOO**

There have been renewed calls for improved artificial intelligence (AI) governance and legislative oversight to address the ethical application of the technology including limiting bias.

Countries like China, the US and those in the European Union have come up with AI regulatory frameworks. In the latest series of initiatives to shape the development of AI, the US, UK and several other countries including Singapore, Australia, Germany and Chile signed on to a new framework in November 2023 to create AI systems that are "secure by design".

In Malaysia, the Ministry of Science, Technology and Innovation is developing a code of ethics and governance — to be ready by this year — that would form the basis of AI regulation for the country. The government has outlined the principles of responsible AI in the Malaysia AI Roadmap 2021-2025. These include fairness; reliability; safety and control; privacy and security; inclusiveness; the pursuit of human benefits and happiness; accountability; and transparency.

Those familiar with AI governance say the overriding goal should be protecting the stakeholders, beginning with consumers and individuals. Asia School of Business CEO, president and dean Sanjay Sarma says issues such as customer privacy, liability, and safety and security need to be addressed.

"A second stakeholder category is creators. We are well and truly into the creator economy through platforms such as YouTube and TikTok. Perhaps we need ways for AI to acknowledge where its generative capacity is sourced from," he says, acknowledging that though this may prove difficult, at the very least creators should be able to mark their work with a "do not mine" request that reputable AI engines will have to respect.

KPMG Malaysia head of technology consulting Alvin Gan says principles that are incorporated in any legislation relating to AI should address key concerns about bias and transparency.

"Unchecked bias can lead to unintended consequences, adversely affecting customers, society and even global communities. Use cases for AI are increasingly based on sensitive personal information, which has raised much public concern about how unfair societal bias, developer bias and model bias could impact decisions and ultimately lead to discrimination against consumers," he says.

Some examples of bias in recent years include sexism as well as other algorithmic biases such as ageism, racism and classism.

For AI solutions to be fully exploited and transformative, trust is imperative, Gan says. According to the recent KPMG global study, *Trust in Artificial Intelligence*, three out of five people across 17 countries surveyed are wary about trusting in AI systems, reporting either ambivalence or an unwillingness to trust.

"In this context, the consideration of universal standards for fairness becomes increasingly pertinent as AI and machine learning continue to advance. The overarching objective should be to foster transparency and, above all, clarity for citizens and consumers who are navigating the complexities of data and information that they provide in digital and analogue experiences," he adds.

ISIS Malaysia fellow Farlina Said points out that increasingly, the conversation on governance of AI is focused on overcoming its so-called black box nature and building accountability in the production and roll-out of AI, which is why international organisations like the Organisation for Economic Co-operation and Development recommend the explainability of the technology as a principle.

Associate Prof Dr Mohd Naz'ri Mahrin, dean of Universiti Teknologi Malaysia's (UTM) Razak Faculty of Technology and Informatics, believes that designers and developers of AI algorithms need to ensure that these algorithms are explainable and verifiable.

"They should not come up with algorithms that are difficult to understand, especially by end users," he says, acknowledging the black box characteristic of algorithms during implementation. "When we use AI to obtain the results of an analysis, most of the time we experience difficulties in understanding the decision-making process. We look at the data and apply the algorithm to get the results. But what happens in the AI system and how does it arrive at the specific outcome or recommendation?"

Decision-makers, he says, should be cognisant of the workings of AI models and, to that end, AI engineers and computer scientists have a role to play in helping users better understand the workings of AI.

"Decision-makers must know how the AI produced the results and be able to explain them, and this leads to the next principle, verifiability, which is checking whether the algorithm did the analysis right," he adds.

Gan concurs, adding that decisions derived from AI models should also be alterable if required. "That's why proactive businesses and governments are needed to evaluate and ensure that the decisions as well as the results of an AI system are not inadvertently skewed or biased, and how a data set that is used to train a model is representative of a desired scenario.

"Additionally, there is a need to determine whether the assumptions and business logic upon which a system was built may contain inherent societal bias and when bias should be included in a model as a fair indicator of the outcome. The use of statistical confidence in inferences becomes essential when making business decisions. For example, some bias that appears skewed towards one group may be appropriate if it can deliver an accurate indicator of the outcome such as when assessing the likelihood of disease contraction based on gender or ethnicity," he says.

COLLECTIVE RESPONSIBILITY

AI governance is widely acknowledged as a collective responsibility. Gan believes that although regulators are already taking steps to establish relevant AI frameworks, laws alone do not inspire public trust.

"There must be collective commitment among all players in the AI ecosystem for trust to exist sustainably within the AI space," he adds.

Governing AI, says Farlina, is challenging.

"The pace of AI development and its incorporation in technology is fast, which means that impact can be unpredictable. So, even while developing with principles such as testing and explainability, there can still be a slight chance of risk, though the rule of thumb is perhaps to withhold from rolling out to the public if the technology and impact cannot be explained," she says.

Outcomes may also not be as expected; for instance, a developer looking for one bias may completely overlook others.

Farlina continues: "The second challenge is how to regulate a technology that can act differently or have variations in outcome for different sectors. We may have to look at the available governance structures, the mandates ministries and agencies regulate as well as the risks the government would hope to mitigate.

"Cybersecurity, for instance, has a short history of governance in Malaysia, beginning with the Communications and Multimedia Act 1998 and the Computer Crimes Act 1997. Yet, as technology becomes ubiquitous, there is a need for data protection acts and an agency at the federal level (National Cybersecurity Agency) to oversee standards and implementation."

ASIA SCHOOL OF BUSINESS

> ***"Too often, regulators get dragged into technical details or overreach. Imagine just banning AI in a country. It's like banning the rain! And it would kill innovation."***
> *Sanjay, Asia School of Business*

> ***"Trust is our licence to operate, and we do that by prioritising responsible data practices and ethics to foster a culture of responsible innovation."***
> *Joachim, CelcomDigi*

Equally, she feels that it's worth looking at legal philosophy to understand what is expected of AI governance and why.

"For instance, one may see principles as norms shaping the environment while another may see them as the underlying standards for rules. So, to understand the types of principles the field of AI needs for governance is to clarify the expectations for the principles, the scope the principles should cover and how a governance structure should work," she says.

Sanjay also raises the need for balance in regulation — putting in place principles that stand the test of time without stifling innovation.

"This is not easy. Too often, regulators get dragged into technical details or overreach. Imagine just banning AI in a country. It's like banning the rain! And it would kill innovation," he adds.

The key, he says, is to focus on principles and not on technological solutions.

"These principles have been articulated. For example, the EU's General Data Protection Regulation provides a nice summary of principles for data protection and privacy. Once principles are articulated, they will need to be reviewed and updated," he adds.

Gan concurs, adding that given mankind's experiment with AI is still at an exploratory stage, institutional safeguards should naturally evolve at pace.

"To stay abreast of technological evolution, as well as the challenges and risks that come with it, it is essential that forums or working groups — represented by both the public and private sectors — are established to enable discussions as well as ideate on governance constructs, public policy and ethics," he says.

Lessons abound from the EU, China and Singapore on AI governance.

"Many advanced AI markets emphasise guidelines and regulations that safeguard and boost domestic research and development (R&D). As they aim for a competitive edge on the regional and global scale, they establish innovation incubators to foster the progress of AI capabilities. In Singapore, the Monetary Authority of Singapore in collaboration with financial industry partners established Veritas, an R&D framework that supports responsible AI adoption and data analytics," says Gan.

CelcomDigi chief corporate affairs officer Joachim Rajaram points to China's governance model, which is said to focus strongly on R&D and digitalisation initiatives.

"Local-level leaders are also encouraged and empowered to drive innovation within their local jurisdictions through the provision of incentives such as start-ups, hackathons or research initiatives," he adds.

Another key takeaway, says Gan, is the building of an AI ecosystem encompassing skills development, technology capabilities, partnerships and cross-functional collaboration. He points to China's New Generation Artificial Intelligence Development Plan, unveiled in 2017, as an example. The strategy, which outlines China's aims to be a global AI leader by 2030, highlights initiatives for R&D, talent development, industrialisation, regulations, ethical norms and security.

WHAT CAN ORGANISATIONS DO NOW?

Although the risks and dangers of AI have grabbed the headlines in recent years, AI is not going away. Generative AI technology, for example, attracted more than US$1.37 billion in venture capital in 2022, thanks in part to the release of tools like DALL-E 2 and ChatGPT.

Gan says KPMG's 2023 Generative AI Survey found that 71% of respondents plan to implement their first generative AI solution within two years.

"However, many companies have not gotten far with risk mitigation strategies for generative AI, with approximately only 31% having evaluated risk and risk mitigation strategies and have/are implementing them," he adds.

In the absence of institutional safeguards, there are several viable strategies for businesses and the private sector to adopt to ensure the responsible and ethical use of AI.

Gan says that businesses can start by assessing their current environment followed by establishing a strategy on how they can leverage AI, such as pilot testing across different areas and functions.

"This will be beneficial for organisations to rethink their approach to mitigating AI risks and place proper controls before expanding the technology. This process will allow the organisation to have a structured journey — from ideation to proof of concept — with a constant focus on the business objectives, organisational alignment as well as its people and customers.

"To safeguard their organisations further, it is important to ensure that trust and governance are established. With an AI risk management framework and tools concurrently set up, this will encourage secure and legal use of AI within an organisation. Combined with effective oversight and continuous monitoring, organisations can build trust in AI," he says, adding that to fully unlock the potential of AI, it's important that data be streamlined.

"Organisations may consider modernising their data systems, models and algorithms. This will help in overcoming challenges such as disparate data, integration and inefficient workflows.

Culture-building within organisations through the instilling of governance and accountability for all employees in the areas of AI deployment and development is also key, says Joachim. "This not only enshrines trust and transparency but also ensures that technology is used to augment human decision-making, not replace it."

At CelcomDigi, the emphasis is on "the centrality of trust in everything that we do, given our mandate as the data guardians of over 20 million subscribers nationwide. Trust is our licence to operate, and we do that by prioritising responsible data practices and ethics to foster a culture of responsible innovation", he says.

Key initiatives focus on AI ethics literacy and trust building.

"We have a taskforce that works on building a framework to govern the responsible use and AI deployment. We are also currently studying appropriate models that will best suit our ambitions. We believe organisations have an obligation to first deepen their compliance maturity internally. They can leverage technologies (for example, privacy-enhancing technology) that have adjustable controls; allowing them to remain flexible and adaptable as the law modifies, grows and develops, to future-proof our processes and governance.

"In parallel to having the right controls, there needs to be constant capacity-building among employees through training and upskilling modules to keep up with the developing technologies and to use them responsibly," he says.

At the community level, the organisation continues to increase digital literacy for target segments of the population, especially the youth and schoolchildren, to prepare them for the digitally evolving workforce.

"It is crucial that we help nurture digital citizens who are mindful of ethical responsibilities when acquiring digital knowledge and navigating the digital space safely."

KPMG MALAYSIA

> ***"This will be beneficial for organisations to rethink their approach to mitigating AI risks and place proper controls before expanding the technology."***
> *Gan, KPMG Malaysia*

UTM

> ***"Algorithms cannot be developed by ignoring human concerns, so by putting people first, principles such as verifiability, accountability and explainability will be part of the students' mindset."***
> *Naz'ri, UTM*

ROLE OF UNIVERSITIES

Gan believes that institutions of higher learning and research institutes have a pivotal role to play in shaping the AI ethical landscape.

"By educating individuals about the potential risks and benefits of AI, as well as best practices for its safe and responsible use, these institutions can achieve a balance which contributes to fostering trust in AI adoption over time," he says.

UTM, which is in the process of establishing the country's first AI faculty, is cognisant of the role played by universities and institutions of higher learning.

Having recently concluded the study plan for the faculty's academic programme, Bachelor in Artificial Intelligence, Naz'ri says prospective students will not only be exposed to AI techniques but also the risks of bias and issues relating to data quality.

"Students will learn to assess AI risks, conduct AI risk assessments and understand the capability of AI, [learning] that it is not a silver bullet for solving problems. What's also important is understanding that whatever solutions they develop, human values should not be ignored," he says.

The setting up of the AI faculty at UTM was announced during the tabling of the 2024 Budget, along with an initial allocation of RM20 million. According to Naz'ri, who is a member of the UTM task force for the establishment of the faculty, the study plan for the industry-infused programme has been approved by the UTM Senate and Department of Higher Education. The faculty, located at UTM's KL campus, is expected to commence the programme in September 2024, with a target enrolment of 120 students.

"Students will learn about AI ethics early in the programme. Also included in the curriculum are maths for machine learning, algorithm and data structures, machine learning, deep learning, natural language processing and generative AI to name a few," he says, adding that the study plan has received encouraging feedback from Zack Kass, who was formerly head of go-to-market at OpenAI and Prof Dr Hamido Fujita, an expert on AI technologies. The former, Naz'ri says, has reiterated the importance of putting people first when it comes to designing and developing AI systems.

"We don't want to produce students who look at AI as if there is no human element involved. Algorithms cannot be developed by ignoring human concerns, so by putting people first, principles such as verifiability, accountability and explainability will be part of the students' mindset," he says. E

TECH FUN

Cooking on the go

THERMOMIX

BY **KIRAN JACOB**

ANYONE with a demanding job knows the struggle of food preparation, be it for the family or oneself. It involves handling various ingredients, cooking them in different pots and pans and the inevitable task of cleaning up, consuming precious hours.

Thermo-cookers simplify food preparation by incorporating functions like chopping, grating, kneading and steaming. This streamlines the cooking process into an all-in-one experience, reducing the number of items to clean afterwards.

Top-of-the-line thermo-cookers boast advanced features such as fermenting, sous vide, slow cooking and even rice cooking. Numerous online videos showcase the versatility of thermo-cookers, demonstrating the creation of dishes from various cuisines like Indian, Indonesian and Italian. In addition, they can assist in baking.

Thermo-cooker brands such as Thermomix offer access to a vast library of over 60,000 recipes. Enthusiastic home chefs can enhance their cooking experience using these devices. However, access to the Cookidoo platform, which is essential for this feature, requires a subscription.

As a high-end product, the Thermomix comes with a substantial price tag. The Thermomix TM6, equipped with a cutter, is priced at RM7,686 according to its website.

But fret not, as there are multiple alternatives on the market. These include the Xiaomi Smart Cooking Robot (RM2,888) and the Kitchen Idea Kody 21 smart cooking machine (RM2,099).

Being cheaper does not necessarily mean it is of lower quality. For example, the Thermomix comes with a cooking temperature capped at 120°C. Meanwhile, the Xiaomi Smart Cooking Robot allows cooking at temperatures up to 180°C.

The advent of air fryers has made cooking much more convenient. As with thermo-cookers, they elevate the cooking experience by allowing people to experiment with different recipes and cuisines. E

XIAOMI