

AI comes with risks, and more can be done to minimise them

The technology is changing the world, and the Republic should update its regulations to keep Singaporeans safe.

Carol Soon and Beverly Tan

On July 24, Singapore announced its latest initiative related to artificial intelligence – AI Trailblazers. Its goal is to generate 100 use cases for generative AI within 100 days, so that the Government and enterprises can test and develop AI solutions for deployment.

It is important to stay on top of this rapidly evolving technology. The applications of AI have been wowing the world over the past year. While AI has been researched and developed since the 1960s, the recent launch of free generative AI tools such as OpenAI's ChatGPT and Stability AI's Stable Diffusion has ushered in remarkable applications such as virtual news anchors and live-streaming hosts on e-commerce sites.

AI promises immense benefits, but it also poses challenges and risks unless it is governed properly.

BENEFITS BUT ALSO RISKS

Predictive AI harnesses its training data to make complex calculations and predictions,

whereas generative AI learns from training data to create new content such as text, images, audio and videos. Generative AI can be particularly useful in increasing productivity by providing hyper-personalised support in various contexts such as writing social media copy or generating logo designs based on the prompts it is fed.

But Generative AI also has the potential to pose powerful threats, particularly when it comes to cyber security. Malicious actors can use AI to generate code for malware.

There are also worries about AI entrenching biases in hiring processes – candidate ranking tools may favour an “ideal” candidate profile based on the dominant race or gender in a particular field. Generative AI also allows for the mass creation of disinformation, such as fake news articles and deepfakes in the form of misleading images, video or audio.

WHAT OTHER COUNTRIES ARE DOING

Internationally, governments are racing to regulate the use of AI to keep their people safe while reaping the benefits of the technology.

The European Union was the

first group of countries to legislate on AI, passing its Artificial Intelligence Act (AIA) in June 2023. A key feature of the AIA is its risk-based approach that sets out different requirements for AI applications deemed to be of different risks – unacceptable risk (will be completely banned), high risk (will be subject to legal requirements), and low risk (mostly unregulated).

Other countries have published voluntary guidelines that organisations can follow while developing or using AI. Britain, Australia and the United States are empowering users to contest decisions made by AI and seek remedies if they are wrongfully impacted. Countries such as Japan and China are making advancements in AI education and literacy.

WHAT MORE CAN SINGAPORE DO

Singapore is part of the Global Partnership on Artificial Intelligence, and alongside 41 other countries, has agreed to adopt the Organisation for Economic Cooperation and Development's AI ethics principles to promote the responsible use of AI. But it can do more to address the harms AI can bring.

Existing initiatives include the Personal Data Protection Commission's (PDPC) Model AI Governance Framework published in 2019, as well as the AI Verify Framework and Toolkit that was piloted in 2022 and launched in

2023, both of which are general guidelines for all sectors. On the other hand, the Monetary Authority of Singapore and the Smart Nation and Digital Government Group have also launched their voluntary sector-specific guidelines for the finance sector and public officers, respectively.

The problem is that these existing initiatives are purely voluntary, and there are no disincentives for organisations that do not adopt the guidelines. The most comprehensive framework in Singapore, AI Verify, serves as a toolkit for organisations to assess their own AI systems. But companies are not required to adhere to AI ethics principles. While existing frameworks are important and a

The problem is that these existing initiatives are purely voluntary, and there are no disincentives for organisations that do not adopt the guidelines. The most comprehensive framework in Singapore, AI Verify, serves as a toolkit for organisations to assess their own AI systems. But companies are not required to adhere to AI ethics principles.

necessary first step towards promoting responsible use of AI, their efficacy depends mainly on voluntary compliance.

While Singapore has always positioned itself as a price taker, in due time, as with other technological threats, legislation might be necessary to ensure that the technology is not misused. Moving forward, the Government might consider imposing legal requirements to ensure that AI developers perform due diligence when developing their AI systems, and that deployers of AI properly assess the risks of the AI systems they use. For example, the developers of AI could be legally required to ensure the security and privacy of the data they collect and use to train their models, similar to the EU's AIA.

Singapore could also consider requiring the labelling of AI applications that impersonate humans and AI-generated content, the way AIA has done so. A survey conducted by ToolTester early in 2023 found that at least half of the participants could not differentiate between content written by AI and that by a human journalist, with the newer GPT-4.0 model fooling almost two-thirds of participants.

Efforts to improve the governance of AI also need to include a wider range of actors and sectors. According to the announcements made by the PDPC and the Infocomm Media Development Authority, a handful of small and medium-sized enterprises (SMEs) have been

consulted in the process of setting up existing frameworks and toolkits. However, these SMEs are mainly from the finance and healthcare sectors.

While there are close to 600 AI-related start-ups registered in Singapore's Startup SG ecosystem, and possibly many more SMEs that deploy AI in their work processes, SMEs often work with limited resources. This means they may be unwilling or unable to comply with stringent regulations on AI. More SMEs from a wider range of sectors should be consulted to better understand and address their needs.

Apart from focusing on organisations, Singapore's existing AI regulations could also include provisions to protect individuals. The EU, Britain, Australia and the US emphasise the importance of allowing individuals who have been wrongfully impacted to challenge AI-driven decisions and seek remedies for them. Singapore could also consider mandating companies using AI to provide a similar feedback mechanism.

Finally, AI education in Singapore can be expanded to dive deeper into AI ethics. While AI Singapore has made a commendable effort in making AI-related courses available and accessible to the public, existing courses focus on explaining what AI is and building AI models. AI literacy programmes should move beyond the technical aspects of using AI to cover different AI ethics principles and the ethical implications of using AI in daily life.

No doubt, more benefits will come our way as AI technology develops. That should not stop us from addressing the risks, before matters spiral out of control.

• Dr Carol Soon is principal research fellow at the Institute of Policy Studies at the National University of Singapore, and vice-chair of the Media Literacy Council. Beverly Tan is a research assistant at the same institute. Both are authors of the IPS Working Paper 52 on the regulation of AI.