CERT
Centre for Responsible Technology

The imperative for responsible technology

By Datuk Prof Dr Mohd Faiz Abdullah

9 January 2026 | Safety, Security, Society, Sustainability

Much digital ink has been spilled wrangling over the promise and pitfalls of new technologies, with public debate swinging between two poles. On one hand are the techno-optimists who present each new technology as the panacea for humanity's deepest social and economic ills. On the other are avid consumers of dystopian narratives who sound the alarm for humanity's imminent apocalyptic futures.

Certainly, this is not intended as a slight against the genre-defining works of Orwell, Huxley, or Atwood, whose enduring relevance lies in their ability to provoke reflection on power, agency, and human vulnerability. Rather, the point is that societies cannot govern technology based merely on anecdotal inkling, let alone hubris.

Instead, what is needed is robust discourse on what these technologies actually are, what they mean for us, and how their adoption or regulation can serve broader societal goals rather than be driven by novelty, profit, or inevitability.

Crucially, this is needed to avoid falling into Amara's law – the tendency to overestimate the impact of technology in the short run while underestimating its effects in the long run. The former fuels hype, panic, and knee-jerk reactions, while the latter leaves institutions unprepared when structural and systemic shifts eventually take root.

These are not abstract concerns. They are foundational for shaping societal decisions on whether, where, and how new technologies should be adopted, and for informing the necessary reviews of laws, regulations, and policies required to align governance with new realities. The stakes are particularly high when technologies begin to mediate not just efficiency or productivity, but human cognition, judgment, and social interaction.

Few technologies illustrate this clearer than artificial intelligence, which has emerged as the defining technology of this epoch. AI is no longer confined to ivory towers or specialised research laboratories, and its applied functions are among the fastest diffusing technologies in modern history.

As a general purpose technology – the likes of which include the steam engine, electricity, and the computer – the list of its use cases grows each day, yet public discourse often treats it as a singular, homogeneous technology. This, in turn, flattens the distinctions that matter for individual adoption decisions, national policy directions, and all those in between.

A more productive starting point is to disaggregate AI by function and use. There is a meaningful difference between general-purpose and domain-specific models, between tools that augment human judgment and those that replace it, and between systems designed to inform decisions and those empowered to make them autonomously.

Consider the contrast between an AI model system that helps a doctor detect anomalies in medical imaging, and one that independently determines medical insurance premiums. Similarly, a model that helps summarises policy papers carries a different risk-reward profile than one that allocates public funding on its own authority.

Viewed this way, it is clear that the differences go beyond mere semantics. These distinctions determine where risks accrue, what responsible use looks like, and how accountability should be determined when automated systems fail or cause harm. Any credible governance regime must be built on this distinction if it is to maximise benefits while mitigating risks.

Appreciating the operational differences across AI technologies would also help sharpen our understanding of the trade-offs associated with its use. While today's frontier large language models require staggering amounts of energy and water to train and run, they are not the only way in which AI can be developed or deployed. More efficient models, such as those specifically built for computational chemistry or healthcare, can often be deployed with far lower energy and resource drains on the planet.

Beyond policy conversations, appreciating these distinctions also enables society to be cognisant of what these technologies fully entail. Generative AI, for example, can help its users explore ideas and engage with complex problems. If prompted, it can also generate the "answers": paragraphs of text that are grammatically correct, logically coherent, and rhetorically convincing.

There is little doubt that this reduces effort and saves time. While some may scoff, the reality is that it is an especially attractive proposition to others. And for many users, this is precisely its appeal. Thinking, after all, is demanding work. It requires one to first grasp the foundations of a given topic, and to formulate a position on it that is aligned with personal values and societal norms. Naturally, the option to bypass this process is a tempting shortcut in the name of efficiency.

But this raises the question: is efficiency and productivity the right metric for society to be optimising for in domains that rely on creativity, reasoning, and reflection?

What is gained in speed may be lost in depth, independence, and epistemic responsibility. Offloading thinking to the machine risks our thoughts being "enframed" – to borrow the concept Heidegger espoused in The Question Concerning Technology – to the incentive and bias structures of the model and its creators. Then there are the further philosophical questions of how this changes Descartes' cogito, ergo sum and whether AI thinks at all or is rather just an impressive stochastic parrot.

Be that as it may, early research is already suggesting that overreliance on generative AI could result in "metacognitive laziness", or a reduction in cognitive abilities such as critical thinking and problem solving.

Our experience with social media platforms is instructive here. It too was adopted rapidly when first introduced, and for quite a while was widely heralded as tools that would enable the next wave of democratisation. Only much later did their less visible negative effects come into focus: shortened attention spans, mental health impacts, and the erosion of shared public discourse. Generative AI may yet lead us down a similar rabbit hole, but this time, with cognitive ability itself at stake.

Like with social media before it, generative AI invites the concern of what widespread reliance and the attendant weakening of cognitive capacities worldwide might mean for democratic participation. Democracy rests on its citizens being able to adequately evaluate information, weigh competing claims, and exercise judgment – but what happens when AI mediates both ends of the process?

None of these considerations should be read as arguments against the adoption of AI, or of new technologies more broadly. Rather, they underscore the need for rigorous, evidence-based research. Such work must be multidisciplinary, examining technical capabilities alongside societal effects and their attendant environmental costs in order to properly assess both opportunities and risks that come along with it.

This is the foundational rationale for the Centre for Responsible Technology: to help shape technological development in line with the principle of responsibility, encompassing public interest, human rights, accountability, and effective regulation. It reflects a sober recognition that technological outcomes are not predetermined, and that to assume otherwise would be to mistake momentum for destiny.

Disclaimer: The views and opinions expressed in this op-ed are those of the author(s) and do not necessarily reflect the views of the Centre for Responsible Technology (CERT), the Institute of Strategic & International Studies (ISIS) Malaysia or the Malaysian Communications and Multimedia Commission (MCMC).