

*RSIS Commentary is a platform to provide timely and, where appropriate, policy-relevant commentary and analysis of topical and contemporary issues. The authors' views are their own and do not represent the official position of the S. Rajaratnam School of International Studies, NTU. These commentaries may be reproduced with prior permission from RSIS and due recognition to the author(s) and RSIS. Please email to Mr Yang Razali Kassim, Editor RSIS Commentary at [RSISPublications@ntu.edu.sg](mailto:RSISPublications@ntu.edu.sg).*

---

## **The Paradox of Scaling AI: New Age or Future Winter?**

*By Manoj Harjani & Teo Yi-Ling*

### **SYNOPSIS**

*As AI matures, constraints to its future progress are also emerging. The question is whether this paradox will lead to another “AI winter” or provide the necessary brakes for a potential runaway train. What could this mean for Singapore’s ambition to be a “living laboratory” for global AI solutions?*

### **COMMENTARY**

DEPENDING ON who you speak to, we are either entering an age where artificial intelligence (AI) is propelling humanity forward or are inexorably developing superintelligence that will cause an existential catastrophe. The reality of AI’s progress is far more prosaic. While there has been considerable advancement in research and in commercial applications, AI has generally been consistent in disappointing both techno-optimists and pessimists.

This does not mean that AI lacks transformative potential. Indeed, the past five years alone have witnessed significant developments, particularly in applications of machine learning. Nevertheless, as AI matures, its limitations have also come to light, ranging from biased output to ever-increasing computing resources required to train and deploy models. These limitations are not trivial as AI becomes more embedded in daily life.

### **Paradox: Avoiding Another “AI Winter”**

It is this paradox — that as AI scales, we are discovering more obstacles to its future growth — which governments, companies and researchers must reckon with. It is unclear whether these obstacles will lead to another “AI winter” where investment and

research will decline or focus attention on current shortcomings in existing AI-based systems and their knock-on societal implications.

Despite what some techno-optimists might [suggest](#), we are a [considerable distance](#) from achieving AI that can scale itself. Humans are still very much “in the loop” when it comes to AI’s prospects for achieving scale. However, it remains to be seen whether this human factor, rather than data or hardware, will be instrumental in avoiding another “AI winter”.

One challenge is that researchers appear to be [prioritising](#) the development of novel techniques rather than making existing applications work better for society. In contrast, when we look at companies, there might be a “winter by stealth”: on the surface, AI innovation continues apace, but brakes are being applied selectively where applications are generating obviously negative consequences.

Recent examples of this include Twitter’s algorithmic bias [bounty challenge](#) for its image cropping tool, and Meta [shutting down the use of facial recognition](#) on Facebook.

However, many governments have yet to tangibly address the larger issue of how to make AI technology accountable to society. High-minded lists of ethical principles and abstract national strategies do little to ensure that societal harms are mitigated and appropriately penalised, let alone incentivise the creation of safe and trustworthy AI-based systems.

The European Union’s approach is a clear exception in this regard. While far from perfect, its [draft AI legislation](#) attempts to introduce a [risk-based framework](#) for regulating AI and protect consumers from potential harms through stiff [penalties](#).

### **What is “Success” for AI?**

These developments beg the question of what “success” will look like for AI. Currently, success seems to mean that the output or outcomes of AI deployment function as expected. Whether or not this expectation accounts for the successful implementation of ethical AI principles — “[ethics by design](#)” — is less clear. There have been significant [examples](#) of correctly-functioning AI-based systems producing discriminatory and unfair outcomes.

Globally, the conversation about ethical AI has [moved](#) from identifying and defining principles to describing what [trustworthy AI](#) is. While this is a welcome and important change, a concern is that this may result in box-ticking exercises that, when completed, bestow upon an AI-based system a false gloss of trustworthiness.

To avoid such “trust-washing”, it is important to interrogate the ethics of actions undertaken throughout the development and deployment process. A [continuous and progressive assessment](#) contrasts with current suggestions for ethical AI [audits](#), which have to contend with sunk costs as they typically occur after the fact.

If a claim of observing ethics by design is to mean anything at all, ethical practice must

be active, real-time, and integrated into development workflows, not simply a consequential debriefing or reckoning.

The question then becomes whether a chain of accountability for trustworthiness can be established through such exercises, and whether integrity is carried all the way along its links. It will also be important to address the prevailing sentiment in some quarters that taking ethics into account “chills” or stifles AI development.

Governments can play an important role here: setting clear and transparent standards for investment in and procurement of AI-based systems. This will incentivise research and applications to prioritise trust and safety, and can be complemented by safety regulations similar to the EU’s draft AI legislation, thereby ensuring that consumers are protected from harm.

### **Implications for Singapore: Is ‘Living Lab’ Goal Still Viable?**

If it is intended for AI to become a key driving force of Singapore’s Smart Nation initiative, this is not yet evident in how resources are currently being allocated. Only [around 13%](#) of the government’s overall ICT procurement budget for the 2021 financial year (~S\$500 million out of an estimated S\$3.8 billion) was earmarked for AI-related projects.

In addition, it is currently unclear how much additional funding from the [Research, Innovation and Enterprise 2025 Plan](#) launched in 2020 has been allocated to AI Singapore, the national research programme for AI, on top of the existing [S\\$150 million committed in 2017 over five years](#).

Two years have passed since the [National AI Strategy](#) was first launched. [Questions around the viability of a “hub strategy”](#) remain, and are joined by new concerns around [ensuring trust and safety](#). Is Singapore’s goal of being a “living laboratory” for global AI solutions still viable, and if so, what should the characteristics of it be in the light of these?

This is an opportunity to re-evaluate Singapore’s notion of success for AI and re-align resource allocation more closely with the relative importance attached rhetorically to AI within the larger Smart Nation initiative. Singapore is still a leader in the region when it comes to AI, but needs to take concerted action in order to sustain its larger ambitions on a global scale.

---

*Manoj Harjani is a Research Fellow with the Future Issues and Technology (FIT) Cluster, and Teo Yi-Ling is a Senior Fellow with the Centre of Excellence for National Security (CENS) at the S. Rajaratnam School of International Studies (RSIS), Nanyang Technological University (NTU), Singapore.*

---