Editorial

Some steps towards a safe and sustainable AI

🔴 Check for updates

The continuing advance of artificial intelligence requires initiatives to address the potential harms of the technology and efforts to develop new energy-efficient electronic hardware.

At the start of November, the AI Safety Summit took place at Bletchley Park in the United Kingdom. Hosted by Prime Minister Rishi Sunak, attendees included the US Vice-President, Kamala Harris, the European Commission President, Ursula von der Leyen, and the owner of X (formerly Twitter), Elon Musk. Released at the event was a document called the Bletchley Declaration[1], which recognizes the need to manage the safety risks posed by advanced artificial intelligence (AI). The declaration was signed by 28 countries including the United States and China.

Two days before the summit, US President Joe Biden issued an executive order[2] on safe, secure and trustworthy AI. The order requires companies developing powerful AI systems to notify the government and share safety test results. And it calls on the US National Institute of Standards and Technology to set the safety standards.

These are valuable initiatives in efforts to minimize the potential risks and harms of AI while maximizing its benefits. At the same time, there is, in much of this discussion, a potential expectation that the recent pace of development in AI technology can be sustained. And whether this will be the case is far from certain.

Improvements in the capabilities of machine learning techniques can be linked to an increase in the size of the artificial neural networks they rely on. Training the leading AI models is now prohibitively expensive, and their size places considerable demands on the underlying hardware and its energy consumption.

Back in July, Alexander Conklin and Suhas Kumar at Rain AI in San Francisco and Sandia National Laboratories argued that the growth in leading AI models has probably slowed over the past few years, and that we are entering an era of economics-limited computing[3]. This has been followed in recent days by an analysis of efforts to monitor the energy consumption — and carbon footprint — of AI models[4]. Written by Charlotte Debus and colleagues at the Karlsruhe Institute of Technology and Helmholtz AI, it highlights the vital need to quantify and report energy consumption in a standardized way.

To sustain the development of AI — and create a more sustainable future for the technology — new energy-efficient hardware is required. A key target in this search is systems that can process data at the edge: that is, close to where the data are being generated. In such scenarios, devices often need to operate under strict size and power constraints.

In a Perspective in this issue of *Nature Electronics*, Dhireesha Kudithipudi and colleagues discuss the design of AI accelerator hardware that is intended for deployment on edge platforms. They focus on hardware that can support lifelong learning models, which learn on a continual basis, acquiring new skills without compromising old ones. The researchers — who are based at various institutes in the United States, Germany and Italy — identify desirable capabilities for such edge AI accelerators and provide guidance on the metrics to evaluate them.

Conventional computing hardware is based on a von Neumann architecture, where processing and memory are separated and data need to be shuttled back and forth between them. This is inefficient when handling machine learning tasks — limiting computing speed and wasting power. Neuromorphic computing, which comes in a variety of different forms, is a possible solution, and the capabilities of such methods can be seen in the latest chip design from researchers at IBM Research in San Jose, California[5]. Led by Dharmendra Modha, the team have developed a brain-inspired chip that co-locates memory and processing elements, eliminating the need to access off-chip memory. The chip — which is termed NorthPole — can perform image recognition tasks faster, and at a lower energy cost, than existing architectures. (See also our Research Highlight on the work.)

An alternative approach, which can also fall under the label of neuromorphic computing, is computing in memory. Here the computational tasks are moved to the memory unit in which they are stored. A complex range of computing-in-memory technologies have been developed so far. This though makes it difficult to establish a comprehensive understanding of the approach.

In a Review Article elsewhere in this issue, Zhong Sun and colleagues attempt to address this problem by providing a full-spectrum classification of computing-in-memory technologies. The researchers — who are based at Peking University, the Technion-Israel Institute of Technology, Southeast University in Nanjing, and University College London — develop the classification by identifying the degree of memory cells participating in the computation as inputs and/or output. The resulting taxonomy allows the advantages and disadvantages of each of the different technologies to be compared.

References

1. *The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023* (UK Government, 2023); https://go.nature.com/49BFvBG
2. *President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence* (The White House, 2023); https://go.nature.com/46sbeno
3. Conklin, A. A. & Kumar, S. *Nat. Electron.* **6**, 464–466 (2023).
4. Debus, C., Piraud, M., Streit, A., Theis, F. & Götz, M. *Nat. Mach. Intell.* https://doi.org/10.1038/s42256-023-00750-1 (2023).
5. Modha, D. S. et al. *Science* **382**, 329–335 (2023).