Nature Computes Better

Opportunity space

v1.1

Suraj Bramhavar, Programme Director

CONTEXT

This document describes an opportunity space - an area that we believe is likely to yield breakthroughs, from which one or more funding programmes will emerge.

You can find out more about opportunity seeds within this space and meet the Creators **here**. In tandem, our programme hypothesis related to this opportunity space has now been published. You can read this document **here.** [PDF]. We have also launched a programme, Scaling Compute, in this opportunity space. Find out more **here**.

This opportunity space is not currently soliciting feedback — you can stay up to date with this opportunity space, plus others across ARIA, **here**.

An ARIA opportunity space should be

+ important if true (i.e. could lead to a significant new capability for society),

+ under-explored relative to its potential impact, and

+ ripe for new talent, perspectives, or resources to change what's possible.

We can redefine the way computers process information by exploiting principles found ubiquitously in nature. We can better understand how the natural world around us performs computation and build dramatically more efficient computers.
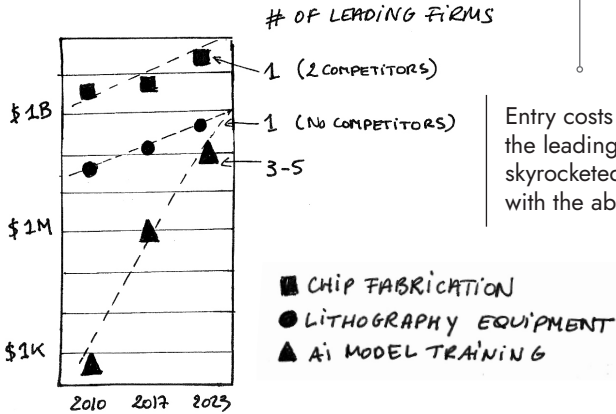
BELIEFS

The core beliefs that underpin/bound this area of opportunity.

1. The growth of AI exacerbates an already unsustainable demand for compute $\longrightarrow$ **we need alternative scaling pathways.**

2. Natural systems are orders of magnitude more efficient than silicon microprocessors at a wide range of computational tasks $\longrightarrow$ **a stronger understanding of how living systems compute is needed to advance both engineering biology and the creation of new hardware.**

3. Investigating the role of statistical physics and nonlinear dynamics in novel hardware represents a significantly underexplored opportunity $\longrightarrow$ **exploiting these approaches is likely to yield new modalities for AI processing.**

4. Modern AI has massive and broad applicability but is underpinned by a narrow set of mathematical kernels $\longrightarrow$ **this presents a unique opportunity for the development of next-generation computing paradigms.**

OBSERVATIONS

Some signposts as to why we see this area as important, underserved, and ripe.

For the first time in computing history, increasing performance requires exponentially increasing costs. The economics of Moore's Law are fundamentally broken.

Vibrant communities have developed around quantum computing and neuromorphic computing, but far less attention has been paid to energy-minimisation in physical systems as an extremely efficient computational mechanism.

Entry costs to innovate at the leading edge have skyrocketed, leaving few firms with the ability to participate.

 # OF LEADING FIRMS

- $1B
- $1M
- $1K

1 (2 COMPETITORS)
1 (NO COMPETITORS)
3-5

2010 2017 2023

■ CHIP FABRICATION
● LITHOGRAPHY EQUIPMENT
▲ AI MODEL TRAINING

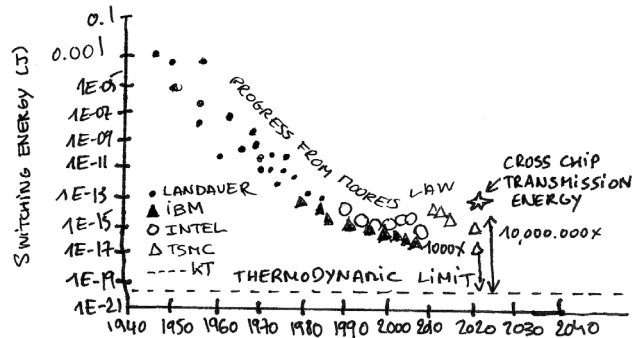*These features are wastefully recreated within a paradigm designed specifically to avoid them*

Algorithmic advances are plentiful (and being investigated exhaustively), but rarely change the fundamental dynamics of the industry. Alternative hardware paradigms present a riskier alternative but can ultimately lead to much larger performance improvements.

Recent popular demonstrations (e.g. Stable Diffusion, ChatGPT) are increasingly showing impressive new capabilities by incorporating phenomena found in nature (stochasticity/probability) as core features.
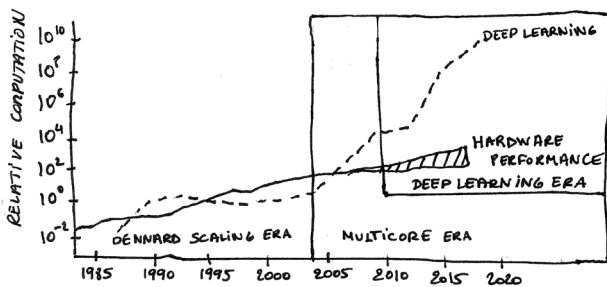
Access and governance of this critical technology has become a major geopolitical issue.

The energy required to flip a bit in the digital domain has plateaued at about 1000x the fundamental limit. Sending that bit across a chip requires yet another 10,000x more energy. A single user's session with chatGPT requires ~150x more power than a human brain consumes performing ALL functions. *There is ample room for improvement*

SWITCHING ENERGY (J)

0.1
0.001
1E-05
1E-07
1E-09
1E-11
1E-13 ● LANDAUER
1E-15 ▲ IBM
       ○ INTEL
1E-17 △ TSMC
1E-19 ---- KT THERMODYNAMIC LIMIT
1E-21
1940 1950 1960 1970 1980 1990 2000 2010 2020 2030 2040

PROGRESS FROM MOORE'S LAW

CROSS CHIP TRANSMISSION ENERGY

1000x 10,000,000x

COMPUTING POWER DEMANDED BY DEEP LEARNING

RELATIVE COMPUTATION

10^{10}
10^8
10^6
10^4
10^2
10^0
10^{-2}

1985 1990 1995 2000 2005 2010 2015 2020

DEEP LEARNING
HARDWARE PERFORMANCE
DEEP LEARNING ERA
DENNARD SCALING ERA
MULTICORE ERA

AI is now the primary economic driver of computational power, and its capabilities are primarily derived from well-known mathematical primitives. *It provides a unique combination of specificity and impact*

Scalability, manufacturability, and true economic advantage have historically limited the adoption of new computing technologies. These commercial constraints can serve as an opportunity to galvanise creative approaches and improve the chances for new technologies to break through.

SOURCES

A compiled, but not exhaustive list of works helping to shape our view and frame the opportunity space (for those who want to dig deeper).

The End of Moore's Law

Compute Demand in the AI Era (Figure 3)

The Geopolitics of AI and Semiconductors (1,2,3,4,5)

The Role of Probabilities in ChatGPT

Energy-Based AI Algorithms

Energy-Minimization in Hardware

Thermodynamic Linear Algebra

Computing with Biology

A Burgeoning Community

Computing with Light

AI & Novel Electronic Memories

The Evolution of Transistor Switching Energy (Figure 2)

ENGAGE

This opportunity space is not currently soliciting feedback — you can stay up to date with this opportunity space, plus others across ARIA, **here**.

If you require an accessible version of this document and/or form, please contact us at **info@aria.org.uk**.