NOVEMBER 2024
TOM WESTGARTH
MELANIE GARSON
KAYLA CROWLEY-CARBERY
ALEX OTWAY
JOSEPH BRADLEY
JAKOB MÖKANDER

State of Compute Access 2024: How to Navigate the New Power Paradox

TONY BLAIR INSTITUTE FOR GLOBAL CHANGE

Contents

01

Executive Summary

Compute resources are the driver of opportunity in the intelligence era. This infrastructure – facilitated by an ecosystem of institutions, talent, governance initiatives and technical partnerships – will shape economic developments, as well as the future of sovereign power and international influence.

National and regional compute ecosystems are evolving daily. Newly launched Danish supercomputer Gefion, for example, aims to be a never-seen-before "factory of intelligence" that will reshape Denmark's sovereign research capacity and role within Europe.[1] And in September 2024, the White House convened a landmark meeting of its compute-ecosystem leaders with US power and utility companies to align on strategy, ensuring that there will be sufficient grid capacity to power the country's leadership of this new industrial revolution.[2]

As many nations forge ahead, compute infrastructure – and the difference in its availability from country to country – risks becoming the basis of a new digital divide. Keeping up with the pace of acceleration requires leaders to have a measurement structure by which they can evaluate their compute capacity, as well as formulate strategy and policy. It is crucial for making sure that they will be able to meet demand for this now fundamental strategic resource, yet few leaders possess the right information on what compute they currently have – and what they will need in the future.

In 2023, in response to this data gap, the Tony Blair Institute for Global Change published the first ever global benchmarking of national compute capacity; our report, *State of Compute Access: How to Bridge the New Digital Divide*, provided leaders with a vital tool to navigate the ecosystem. In 2024 we have expanded on these findings with revised indicators to reflect the rapidly shifting landscape; we have also covered an additional 12 countries for more geographic representation.

In this follow-up report we show that the divide is still stark and growing.

- Globally, 60 per cent of installed servers[3] are located in Europe and North America. These regions account for only 17 per cent of the world's population.
- The processing power of the 500 fastest supercomputers has increased by a factor of 2.8, leaving those in emerging compute economies without supercomputers unable to compete.
- Microsoft has invested $10 billion in 462,000 H100/H800 GPUs, as it plays catchup with Meta in the compute race. And the new Colossus supercomputer in the US, built by xAI and NVIDIA in only 122 days, will have 2.5 times as many GPUs compared to Frontier, the country's number-one ranked supercomputer.[4]
- Sub-Saharan African countries suffer an average of 87 blackouts a year, whereas North America averages one per year.

There are also some key new trends emerging.

- The global compute race is accelerating. The United States maintains data-centre supremacy, having built more data-centre capacity in the past 12 months than the rest of the world combined, excluding China. Meanwhile, China's development has become increasingly opaque, which has created a new compute-security dilemma.
- Space and power constraints mean that many countries are focusing on quality of servers rather than quantity. Switzerland, Ireland, the Netherlands and Canada are leading this charge in upgrading data centres, replacing large numbers of inefficient servers with smaller numbers of AI-optimised servers.
- Physical geography is becoming a key factor in growth. Finland is making the most of a cooler climate (in which data centres are more energy efficient), having installed 39,000 new servers; Malaysia has built 210MW of new data-centre capacity, benefitting from the compute demands of neighbouring Singapore.
- Investment begets more investment: commitment to one part of the compute ecosystem yields massive impact. For example, Israel's investment in public cloud and high-performance computing (HPC) has dominoed investment in a 20,000 per cent increase in data-centre floor space; new investments in Côte d'Ivoire are likely to lead to an 84.3 per cent increase in server revenue growth over the next five years, which is a

projected growth rate higher than any other country in our study. And while not yet part of the TBI data set, new data-centre investments in the Democratic Republic of Congo have quickly received buy-in from its leading bank Sofibanque to help scale its infrastructure solutions.[5]
- Planning bureaucracy is throttling growth, with countries such as the United Kingdom facing obstacles to rapid expansion of both energy and data-centre infrastructure.

Meanwhile, obstacles that were identified in 2023 persist (and, in some cases, have intensified) in 2024.

- Access to sustainable energy resources remains the key driver when it comes to accelerating compute ecosystems.
- Emerging digital economies continue to demonstrate huge growth in innovation potential, as well as increased software-developer contributions. However, without opportunities for graduates to access compute resources, these countries risk losing the benefit of their homegrown talent.
- Geopolitical tensions and security fears across supply chains, in addition to attitudes to data portability, are driving new alliances that risk compute interoperability.

The acceleration and sophistication of compute ecosystems since our 2023 report is staggering. In sub-Saharan Africa for example, which has a large youth population, the growth of the software-engineering community – particularly in Ethiopia, Rwanda and Kenya – is larger than in any other countries we have studied. Elsewhere, new models of partnerships and investments have been created in the Middle East, India and Thailand, providing new opportunities to resolve geopolitical tensions between compute ecosystems.

There is also evidence that access to clean and sustainable energy resources in Europe, particularly some Nordic countries and Romania, is enabling the infrastructure for new innovation opportunities. And close cooperation between government and the data-centre industry in countries such as Singapore continues to deliver value in achieving compute access.

However, not all countries have evolved on their anticipated trajectory. Significant players such as the UK, Canada and Brazil have slowed down; they will need to overcome bureaucratic and investment challenges to ensure the implementation of current plans and supercharge their compute ecosystems.

Enabling governments to maximise on current compute trends and overcome the bottlenecks in their physical, institutional and regulatory infrastructure requires a dynamic whole-of-ecosystem approach. An approach that requires making key choices of where to maximise or cede elements of sovereign power to gain access to more compute power. This report and data explorer provide insights for all stakeholders to visualise the needs of their ecosystem as they look to 2025 and beyond. Leaders should use this report to strategise on how they can best invest their resources to deliver maximum impact and meet their compute aspirations. This, in turn, will mean they can meet the broader responsibility of leaving no nation behind in this AI-powered age.

02

Access to Compute: The New Foundation of Economic Growth and Geopolitical Influence

As countries digitise and pivot into the intelligence era, it is becoming clear that compute is the foundation of next-generation economic growth and influence. This capacity to process data at scale is being leveraged across every sector imaginable. From the creation of new video-simulation data to train driverless cars,[6] to new chatbots that can provide advice and support to victims of gender-based violence,[7] every field of humanity can be enhanced by large-scale compute-enabled processes.

The potential applications are only set to rise. The performance of AI hardware is doubling every 2.3 years,[8] and the performance-per-dollar of NVIDIA chips has been improving by 30 per cent per year. The combination of increasing performance and falling unit costs means that those that have access to the power to "manufacture intelligence" at scale will be at the forefront of harnessing the benefits of finding solutions to key challenges, from green transition to digital biology.[9]

The large tech companies, including Alphabet, Amazon, Meta, Microsoft, xAI and Oracle, are driving competition for the latest chips, such as NVIDIA's Blackwell GPU, launched in March 2024. This latest architecture is designed to offer four-times faster training and 30-times faster inference.[10] As today's largest AI models can be four times more computationally intensive than in 2023,[11] these chips – due to start shipping at the end of 2024 – will revolutionise the next generation of compute resources.

Compute is not just a source of scientific and economic progress, but the new benchmark of global power economically and geopolitically. Economic benefits are accruing; cloud computing looks set to make up about 8 per cent of India's GDP by 2026.[12] And infrastructure investments in quantum computing will reap a share of a market estimated to reach between $450 billion and $850 billion by 2040.[13]

With $210 billion having been spent globally to upgrade and maintain the existing server base in the past year, a new compute "security dilemma" is emerging as the two largest state players, the US and China, seek to check the other's power with export controls, investments and industrial espionage. With China's new "HPC iron curtain",[14] and its refusal to declare new supercomputers and other investments to global indices such as the TOP500 Supercomputers list, the balance of compute power will be increasingly difficult to verify. This strategy will lead to increasing uncertainty in the world order and intensify the race for compute supremacy.

As the commoditisation of computer resources intersects with the clash of great powers, questions of digital sovereignty and digital colonialism have proliferated. Caught between allegiances to infrastructure providers on the one hand and access to critical components on the other, emerging digital economies find themselves in strategic dilemmas as they trade off between a series of sub-optimal binary choices. This leaves opportunities for rising compute powers including Australia, India, Malaysia, Morocco, Saudi Arabia and UAE to build a new "compute diplomacy", leveraging their compute ecosystems in regional partnerships as non-aligned countries seek alternatives to boost their access to compute.

03

The State of Play in 2024

In 2022, based on the assumption that compute would be the basis of the next digital divide, and in anticipation of many of the dynamics outlined above, the Tony Blair Institute for Global Change felt there was a need to assess countries' national compute capacity locally and internationally. This would provide leaders with tools to consider their strategic and policy pathways in accessing the resources to power the new generation of progress in the intelligence era.

In November 2023, TBI in collaboration with Omdia, published the *State of Compute Access: How to Bridge the New Digital Divide*, the first global benchmarking of compute capacity. As new studies confirm,[15] the trends that we identified in last year's study are taking root, with an increasing compute divide emerging.

Yet the compute ecosystem is also shifting, and discussions of compute are now more firmly on national and global agendas. Since our last report on compute was released, the processing power of the 500 fastest supercomputers has increased by a factor of 2.8. This allows for entirely novel approaches to problem solving, such as Switzerland's new Alps supercomputer, now the sixth fastest in the world, providing high-resolution weather forecasts over the Swiss mountains.[16]

Measurements of compute capacity itself are being used in novel ways. It is now emerging as a potential tool for regulation, as policymakers in the US[17] and EU[18] develop legislation that uses compute processing power as a criterion for regulating AI models.

And the sectoral and national gaps that we identified are being acknowledged. The US,[19] UK,[20] and EU[21] have begun to open up national compute research resources to academics and startups. Similarly, in September 2024, the UN's High-Level Advisory Body on AI proposed a global AI-development fund[22] to support access to compute for developing nations.

However, despite these steps forward, the divide persists and is increasing or being reshaped.

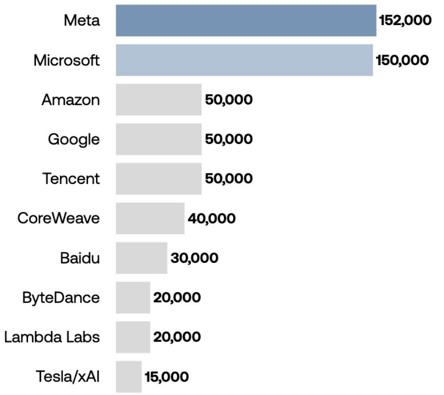
- There is a lack of competition: Omdia analysis for TBI shows that of the $138.9 billion in AI-server capitalisation in 2024, more than 50 per cent of this came from the top ten cloud-service providers.
- Gaps exist on multiple levels: between countries, between government investments and private-sector investment, and between the private-sector actors themselves. Microsoft has caught up with Meta to be a leading investor in AI compute, having ordered 462,000 H100/H800 GPUs this year, three times more than last year's investments.

FIGURE 1

Microsoft has led the way on H100/ H800 shipments in 2024

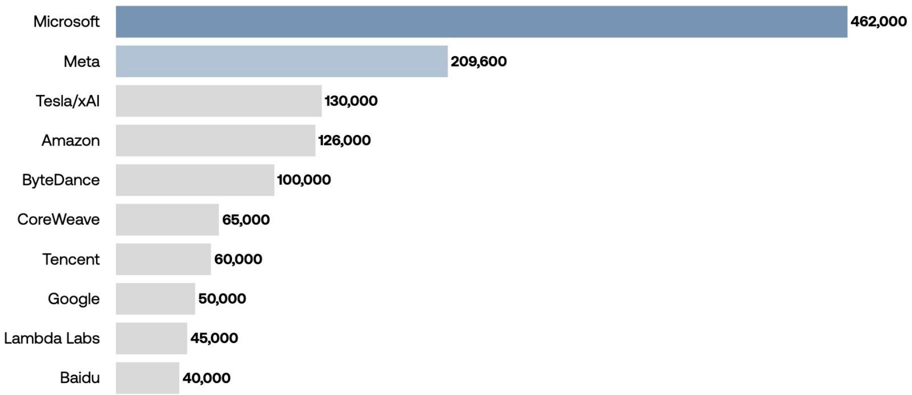
2023

NVIDIA Hopper GPU (H100 GPUs and derivatives) shipments to end users



2024

NVIDIA Hopper GPU (H100 GPUs and derivatives) shipments to end users



- Resource races have not only impacted the key supply chains underpinning compute ecosystems, but have awakened fears and

insecurities with the potential to increase the compute divide. These volatile supply chains and fears of digital colonialism are driving narratives of sovereign compute; such an approach eschews the cooperative models of data-sharing and resource-sharing that are the basis of the interoperable internet and shared compute infrastructure.
- Many leaders are still struggling to identify the right sequencing of their investments, including access to energy, skills and a clear strategic vision to scale their compute capacity.

Methodological Changes

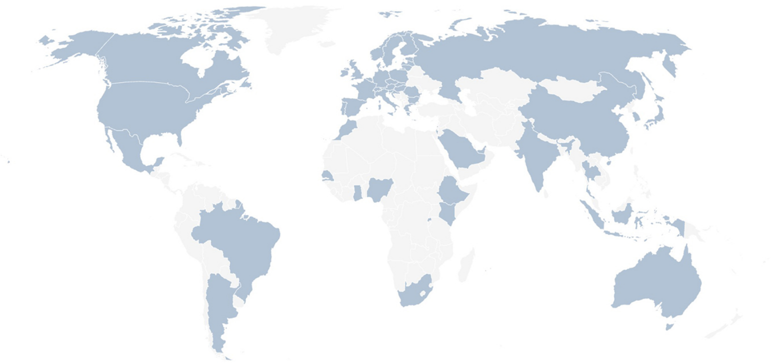
To accommodate these changes, this year's research has undergone two main changes to provide, to the extent possible, an accurate reflection of the changed state of access to the compute landscape.

First, an additional 12 countries were added to our research, bringing the total number of nations benchmarked to 67. The inclusion of Angola, Chile, Colombia, Côte d'Ivoire, Egypt, Malaysia, Mozambique, Peru, Philippines, Tunisia, Turkey and Vietnam enables a more global analysis of the key trends in the ecosystem.

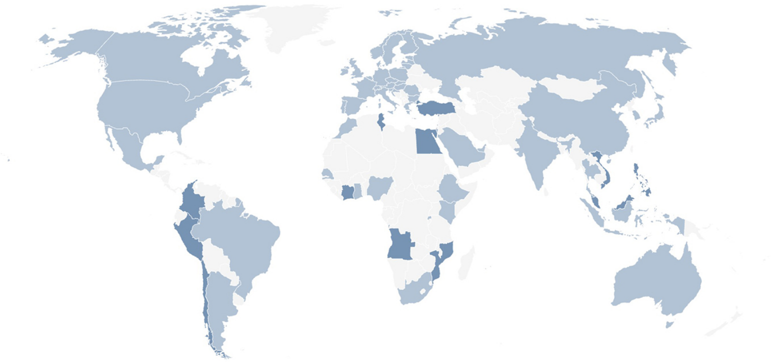
FIGURE 2

The new countries included in the 2024 report

2023



2024



We have also removed seven indicators that were either strongly correlated with other indicators or were not strong drivers of compute. These have been replaced with four new indicators that are more reflective of the opportunities and challenges of compute ecosystems, including server unit growth (five-year compound annual growth rate), AI skills demand, availability of clean energy and energy independence. These have been integrated into a revised compute-capacity measurement framework, which aims to represent the key groups of drivers at the heart of a functioning compute ecosystem. (An annex of the full methodology is available as a downloadable PDF.)

FIGURE 3

Measurement framework for State of Access to Compute 2024

Existing compute capacity

1. Server installed base
2. Server installed base per capita
3. Number of academic institutions with access to quantum computing
4. Access to parallel computing
5. Local supercomputer capacity

Connectivity & infrastructure

1. Power-grid maturity
2. Number of subsea cables and IX points
3. **Availability of clean energy** NEW
4. **Energy independence** NEW
5. ~~Availability of renewable energy~~ REMOVED
6. ~~Ecosystem differentiation~~ REMOVED

Policy & governance

1. AI strategy
2. Data-governance strategy
3. VAT on computer hardware
4. ~~Sustainability incentives~~ REMOVED

Local talent pool

1. Availability of training programmes
2. Participation in open-source community
3. **AI skills demand** NEW
4. ~~Share of STEM graduated~~ REMOVED

Compute-capacity buildout

1. Data-centre investment
2. Quantum investment
3. Server investment growth
4. **Server revenue growth** NEW
5. ~~Investment in high technology~~ REMOVED

Software & programming

1. Number of software engineers
2. Participation in open-source community
3. Cloud-service availability
4. ~~Relative supply to demand for software engineers~~
5. ~~R&D incentives~~ REMOVED
6. ~~Compute patents~~ REMOVED

This report is structured in two key sections. The first sets out the key findings from our data, including the core regional trends and the overarching themes that are exemplified by countries that have had an exceptional year in compute development.

The second section digs deeper into the choice and trade-offs that political leaders must navigate when trying to scale compute capacity. The impact of each decision varies according to an individual country's ecosystem, with the data explorer allowing leaders insights into a more nuanced approach for visualising their compute-access journey.

We also outline the current models of public compute and shared compute ecosystems that provide alternatives to building compute capacity alone, and which were a key part of our recommendations in our 2023 report. And we look at the new models of public-private partnerships and compute via international-development funds, which provide another set of options to address the compute divide.

Finally, we look towards 2025 and the levers that leadership can engage to accelerate access to the prosperity and growth of the intelligence era. The choice of model will be dependent on the core conditions in any given country or group of countries; however strategic thinking for national, regional or global cooperation will be imperative for any of them to succeed.

FIGURE 4

State of Access to Compute 2024: at a glance

2.8x

increase in the processing power of supercomputers in the TOP500

1.9x

increase in the annual megawatt buildout of data centres compared to last year

128

new institutions adopted quantum computers

$210bn

spent to upgrade and maintain the existing server base

04

The Regional Picture of Compute Development

Since the *State of Compute Access 2023*, there have been significant shifts in compute capacity at a regional level. As the number of countries included in the data has increased, we have been able to better understand which regions are cementing their positions in the short term as global centres of compute, as well as the variance within regions that may contribute to their success or decline.

The regional snapshots, alongside the country analyses, also point to where long-term investment into access to compute may lead to a rebalancing of global powers and new models of compute diplomacy. This will enable leaders to rethink their strategies to secure their positions in the compute era.

Energy is Central to European Growth

Overall, there has been an improvement in Europe's supercomputer capacity. While European High Performance Computing Joint Undertaking (EuroHPC JU)[23] is at the heart of the region's overall compute plans, some countries have had particularly substantial increases in domestic supercomputing capacity. Germany had the greatest gain in supercomputers following the US, adding nine additional supercomputers to the TOP500, and is set to host Europe's first exascale supercomputer, JUPITER.[24]

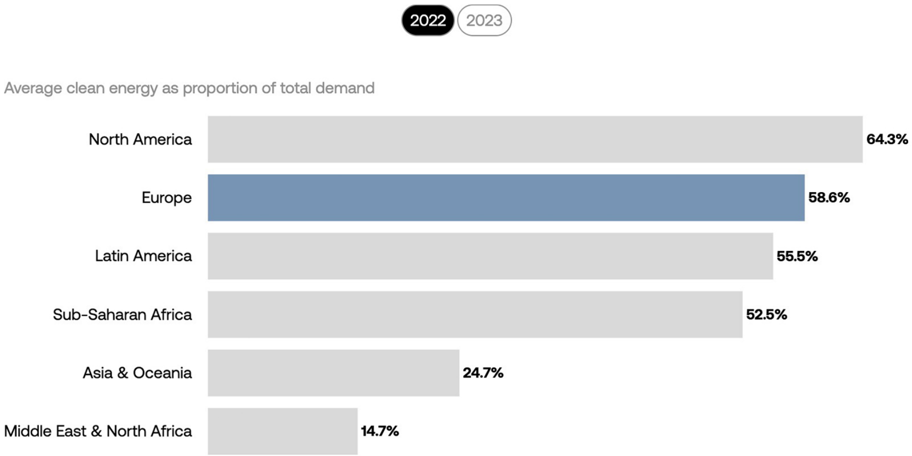
Yet Europe's talent gap persists. Reflecting recent EU reports on rising labour and skills shortages in all member states,[25] six of the eight countries that had declining numbers of software engineers were in Europe, with only Croatia demonstrating any growth.

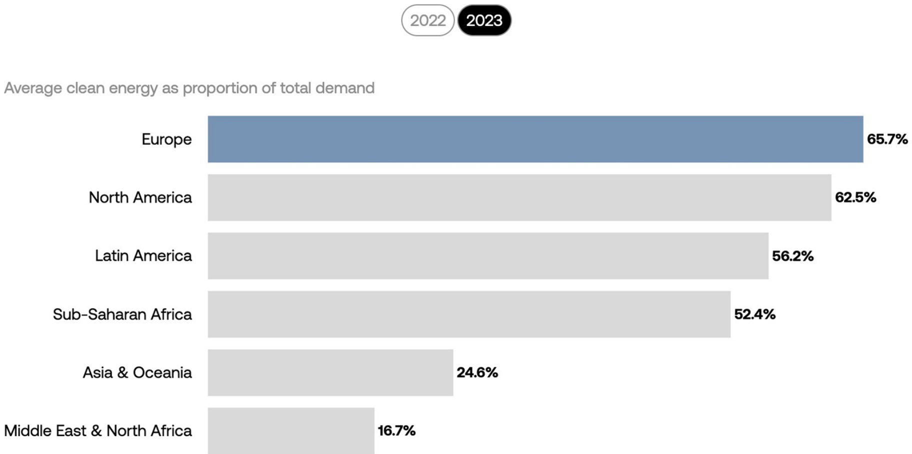
The most noticeable shift in Europe, however, has been driven by the increase in the availability of clean energy. Investment in renewables has been a priority across Europe to secure domestic energy supply and reduce heavy reliance on Russian energy.[26] Europe was the only region to have demonstrated significant growth in the availability of clean energy.

Companies building the largest data centres will prioritise putting them in locations that can build out energy infrastructure fastest. If Europe can translate its greening economy into one that can rapidly scale infrastructure, it will strengthen its competitive advantage.

FIGURE 5

Between 2022 and 2023, Europe demonstrated the biggest leap in clean-energy investment



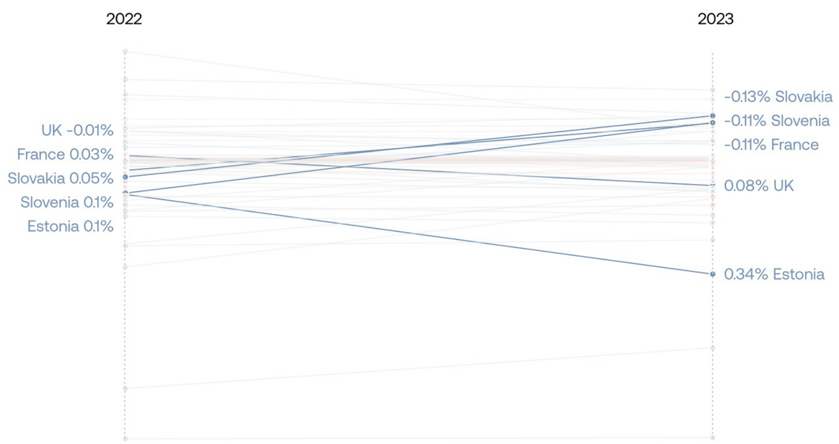


At the same time, there has been significant variance in energy independence across Europe. The UK and Estonia have grown substantially more dependent on trade for energy, whereas France, Slovakia and Slovenia have consolidated themselves further as net exporters. Romania is also showing increasing strength in this area, with a 7 per cent increase in energy produced from clean sources helping set the conditions for a potential $2 billion in data-centre investment from Google,[27] which was inked in a memorandum in July. This low reliance on fossil fuels and a positive energy-trade balance leaves these countries well positioned for energy-intensive compute infrastructure.

FIGURE 6

Energy supply in Europe has far greater variance than the rest of world

Notes: Imports as % share of total demand (negative value denotes net export of energy)



These resource variations are also driving the nature of server investments in Europe. Energy-availability constraints in Ireland, Germany and the Netherlands have led these countries to limit new data-centre connections to the grid.[28]

However, Nordic countries are making the most of their colder climates[29] and the availability of low-cost[30] renewable energy, with some of the highest projected rates of server unit growth. This expansion of compute infrastructure is likely to continue to grow as specialised companies such as atNorth, which offer build-to-order electricity generation, are emerging that will help meet the challenge of building scalable and sustainable data centres.

Middle East and North Africa: a Talent Boom Helps the Region Catch Up

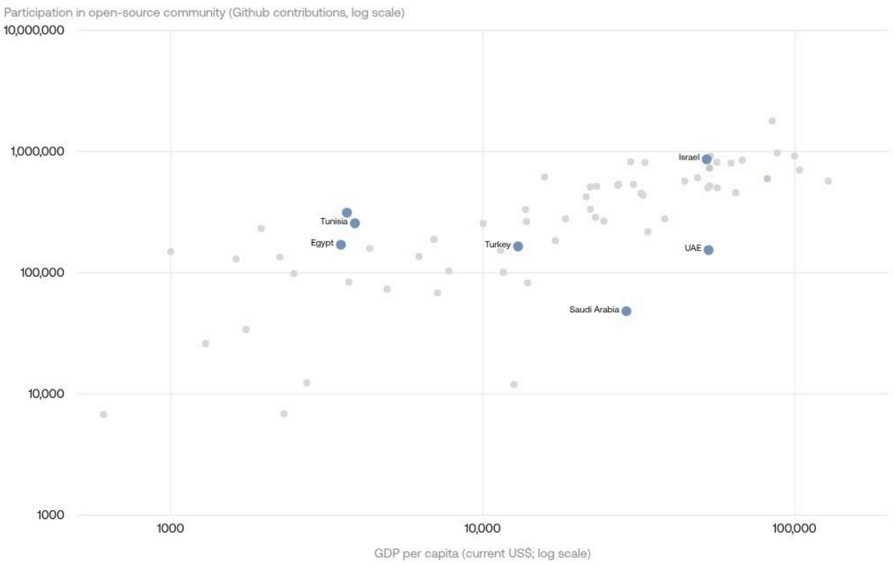
TBI's 2023 report highlighted that the challenges faced by the Middle East and North Africa (MENA) in accessing green energy have been a factor in stunting growth. But the UAE has made significant progress in this area. Over the past year, it achieved 70 per cent growth in installed renewable-energy capacity,[31] and aims to triple capacity by 2030.

Leading data-centre operators such as Equinix have been increasing their investments into compute infrastructure across the region, while Egypt has established itself as a hub for subsea cables.[32] The region has also seen a massive growth in talent, as training programmes have become more readily available to all countries in the region, with the UAE now in fourth place globally.

The talent boom in MENA is particularly noticeable as the open-source ecosystem expands at pace. While Israel remains a global and regional leader in GitHub contributions,[33] there was a 259 per cent increase in GitHub pushes per labour-force participant in Morocco, and Egypt's increased by 95 per cent. The UAE, Saudi Arabia, Tunisia and Turkey also saw strong growth in GitHub contributions. This open-source strength is also reflected in the development of new tools, with popular state-of-the-art open-source models such as Falcon and Jais being developed in the Middle East.

FIGURE 7

MENA's investment in open-source participation is on the rise



Despite this growth, both Saudi Arabia's and the UAE's open-source contributions are currently lagging countries at a similar stage of development. As a result, what has been a rapid talent boom will need to sustain over time for the region to have a globally competitive technology talent market. As we discuss later in this report, the strong human-capital base that is essential for driving compute growth also requires investment in infrastructure to keep skilled workers engaged locally.

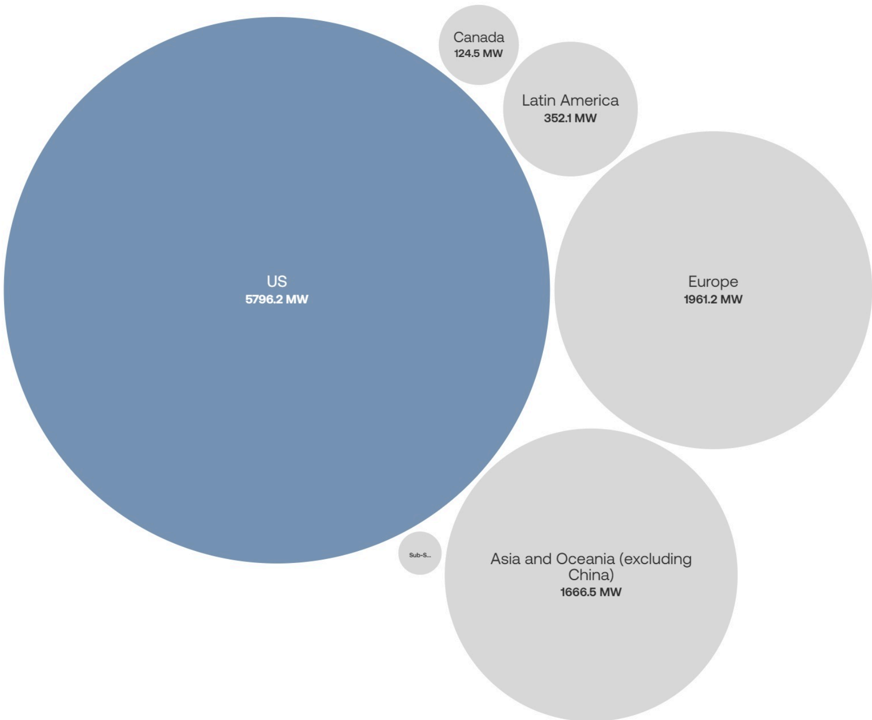
There has also been some slowdown in the growth of compute hardware – the forecast five-year compound annual growth in server revenue has dropped for the UAE from 29 per cent to 17 per cent. However, the region remains in the spotlight for new investments. OpenAI CEO Sam Altman met with Middle East investors this year[34] to discuss developing new chips for OpenAI infrastructure and Groq has agreed a new partnership with Aramco[35] to build the world's largest AI inference centre. If the region can couple significant investment potential with an abundance of clean energy, then MENA will emerge as an integral site for the development of the global digital landscape.

North America: United States Maintains Regional and Global Supremacy

The United States continues to build and invest at a scale that currently outstrips any other country. Our data-centre investment indicator shows that the US announced that 5,796.2 megawatts' (MW) worth of data centres are to be built over the next three years; with the exclusion of China, this is more than the known combined investment of all the other countries in the set.[36] This is a 164.4 per cent rise from 2023 and shows no sign of slowing, with Microsoft and OpenAI recently announcing plans to build liquid cooling data-centre campuses at close to gigawatt scale.

FIGURE 8

The US is forecast to outperform all other economies combined (excluding China) on data-centre investment



The processing power of US supercomputers in the TOP500 increased by 4.9 million teraflops. This was more than the increase in teraflops in TOP500 supercomputers from the rest of the world combined.

This continued development has been enabled by key strategic decisions to push towards HPC dominance. The National Science Foundation granted long-term funding to institutions such as the Texas Advanced Computing Center,[37] as well as the resources to build specialised data centres to house future supercomputers.

The US has also made massive investments in quantum – more than double that of Canada, its nearest rival. These have been facilitated since 2018 by a comprehensive set of quantum strategies.[38] The investment difference is so dramatic that lack of comparative resources is causing talent flights from Canada. Professor Yoshua Bengio of the University of Montreal stated that graduates in his research groups are now going to the US for jobs, and that “the brain drain is probably coming back”.[39] As the US builds new strategies such as its 2024 report Advancing International Cooperation in Quantum Science and Technology,[40] the country’s structural commitment to fight for compute dominance will continue.

Economic Challenges Mean That Central and Latin America Cannot Sustain Server and Data-Centre Growth

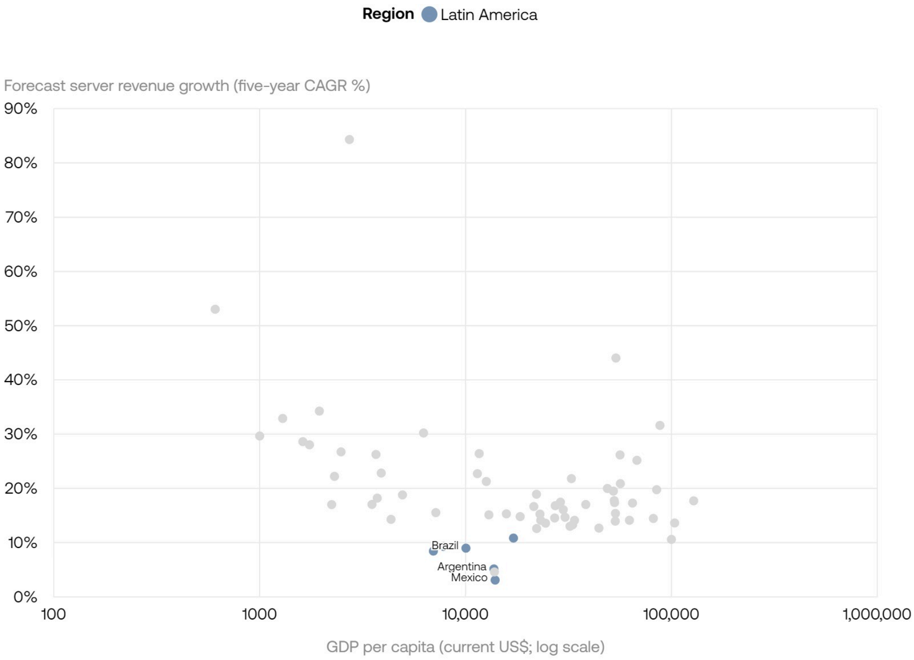
There has been significant growth in compute infrastructure across Central and Latin America over the past year. Mexico saw a 39.6 per cent increase in servers installed, while Chile and Argentina saw increases of 29.5 per cent and 16.5 per cent respectively, compared with a global average of 14.6 per cent.

Argentina was one of five countries to enter the supercomputer TOP500, while Brazil added two more TOP500 supercomputers, to reach a total of eight. The Brazilian government has signalled that compute is a priority for the nation, announcing an investment package in its AI Action Plan,[41] including upgrading the Santos Dumont supercomputer to make it one of the five fastest globally.

Despite these gains, the region’s growth looks as if it is slowing and is significantly lagging in forecast server revenue growth.

FIGURE 9

Low-revenue growth is a symptom of poor economic growth in Latin America



Average growth in the region was 8 per cent in 2023, compared with 21 per cent in other similar countries. This shift is most apparent in Mexico and Brazil, which have seen their five-year CAGR fall from 25 per cent to 3 per cent, and 19 per cent to 9 per cent respectively. This is partly driven by the shifting priorities of the largest builders and investors. Amazon, Microsoft and Google had built large campuses in countries like Mexico, but in 2024 have moved to building AI-optimised high-value server clusters at home, rather than general-purpose server and storage hubs abroad.

While this reflects the overall challenges to economic growth in the region (the World Bank[42] projects that the region is set to have the lowest rates of GDP growth), excessive red tape and grid issues[43] have also posed challenges to data-centre investment. These challenges are not unique to the region, but without a strong compute base, they are especially problematic.

Despite strong inflation management, investment and consumption are low, with high interest rates and fiscal deficits persisting. Whether or not the region will be able to develop shared compute capacity to minimise the financial burden remains to be seen.

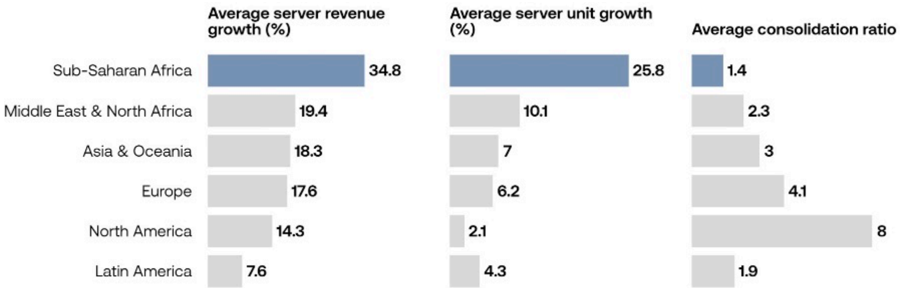
Rapid Growth in Sub-Saharan Africa, but Development Constrained by Grid Reliability

Compute ecosystems in sub-Saharan Africa are continuing to develop at significant pace; the addition of Angola, Côte d'Ivoire and Mozambique to this year's report reflects the increase of countries with rapidly maturing infrastructure that are delivering noticeable impact. In Côte d'Ivoire, server revenue growth is anticipated to reach 84.3 per cent over the next five years – a projected growth rate higher than any other country in our study. Recent investments such as the newly opened CIV1 data centre in Abidjan, which aims to bolster capacity across the West African Economic and Monetary Union, will be central to this growth.[44]

This investment is reflective of growth throughout sub-Saharan Africa, with an overall 25.8 per cent increase in the number of server units. The region is seeking to scale and catch up through massive investment in large numbers of less cutting-edge, lower-value servers.

FIGURE 10

Sub-Saharan Africa leads the way globally in its efforts to scale server infrastructure



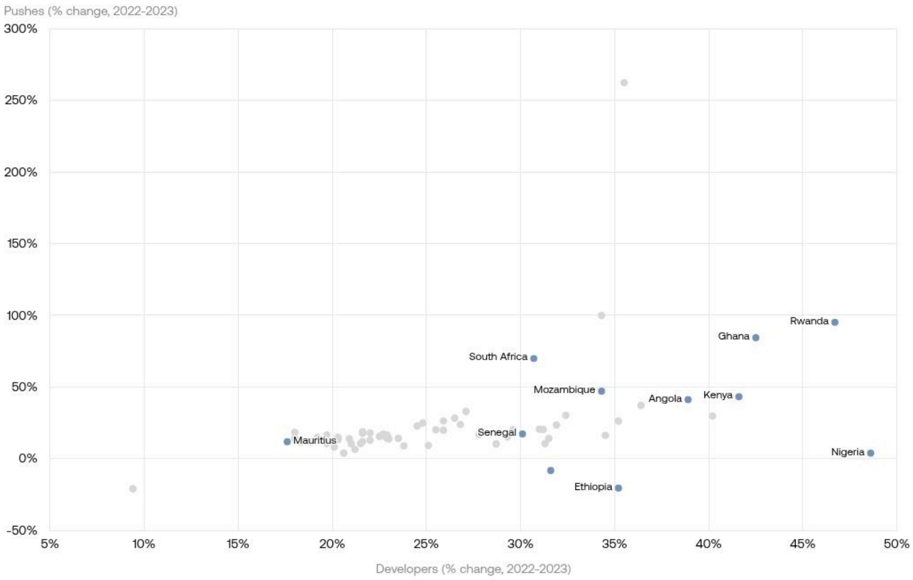
Notes: Server consolidation ratio divides forecast five-year CAGR revenue growth (%) by forecast five-year CAGR unit growth (%). Countries with a high ratio are expected to focus on consolidating existing stock; those with a low ratio on expanding stock.

As more mature ecosystems, such as South Africa, expand their compute ecosystems with new initiatives and collaborations,[45] these server investments have led to countries such as Rwanda, Senegal and Mauritius becoming leading contributors to the region's compute capacity. This looks set to continue, with countries such as Senegal now a top-ten country for forecast server revenue growth over the next five years.

Talent pipelines also continue to flourish in the region. Ethiopia, Rwanda and Kenya have seen the biggest year-on-year increase in software engineers across our data, at 40 per cent, 39 per cent and 29 per cent respectively. And Rwanda, Nigeria, Kenya and Ghana, in particular, have shown high growth in both the size of their developer community and its activity, while the region more broadly has the highest growth in GitHub developers year-on-year.

FIGURE 11

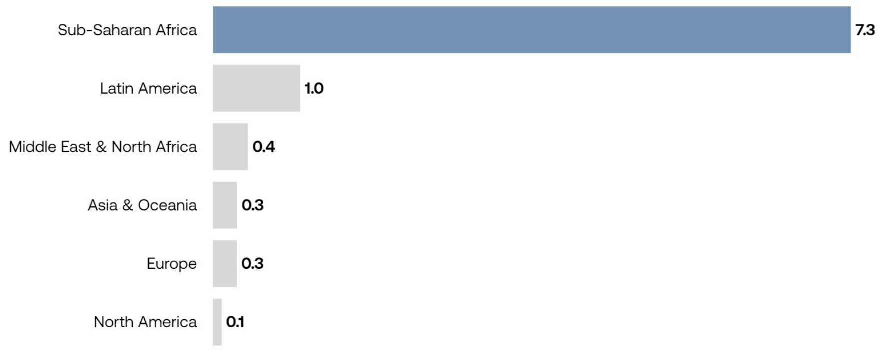
Africa's open-source community is growing in both size and activity



However, there are still some clear challenges hindering the growth of compute ecosystems in the region; sub-Saharan Africa's defining challenge remains creating the sustainable conditions for scaling server infrastructure. As noted in last year's report, problems with power-grid reliability hinder countries' efforts, with an average of 7.3 grid outages per month.

FIGURE 12

The number of power outages that sub-Saharan Africa suffers per month is unparalleled



Countries such as Nigeria are seeking innovative solutions, such as solar-plus-storage systems[46] or co-locating data centres near solar installations, to alleviate strain on the grid and enable their digital potential.[47] Similarly, in Kenya, Microsoft is investing $1 billion in a geothermal-powered data centre to increase the energy resilience of its compute ecosystem.[48] Such initiatives will be key to realising the value of the region's server investments.

Alongside the troubling infrastructure picture, there has not been an increase in the availability of compute training programmes via major commercial leaders, and some talent may be turning to informal channels to train themselves. This means that despite higher levels of STEM graduates, sub-Saharan Africa is not developing strong human-capital pipelines to compute roles. Indeed, despite a new wave of investment, the region has a similar number of servers to Spain; per capita, the UK has 20 times as many. Two changes are required: an increase in edge computing (where data processing takes place on a device or local server, rather than a data centre) to help leapfrog infrastructure limitations, and more opportunities for the region's skilled workforce within local compute ecosystems. Without these changes, the local talent pool is likely to seek opportunities elsewhere – and this could stymie the growth of the region's access to compute.

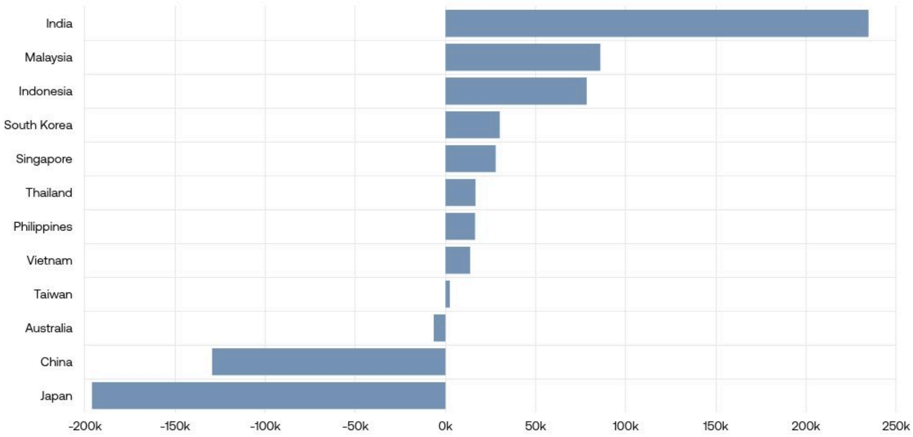
Asia-Pacific and Oceania See Slow but Varied Growth

There is huge variation in the growth of compute ecosystems across Asia and Oceania. It remains a significant region in terms of compute deployment, but imbalances are emerging in countries' hardware and talent bases.

In particular, there is a worrying diversity in the rates of server growth: for example, Japan's loss of 200,000 servers is just offset by India's growth, and China's loss is offset by the rest of the smaller nations. Some of these drops likely represent server consolidation;[49] for example, Japan has a consolidation ratio of 3.75, putting it just behind nations with physical space constraints such as Singapore and South Korea.

FIGURE 13

The annual change in Asia and Oceania's installed server base has been volatile

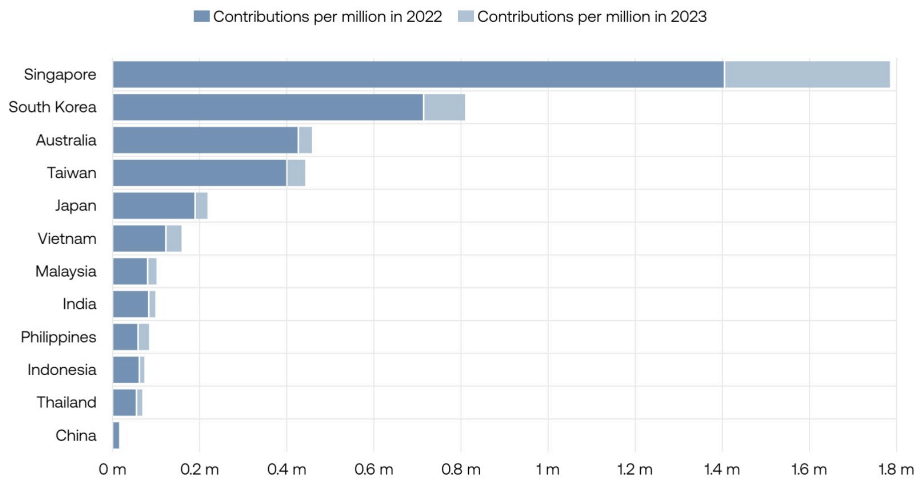


Other local factors have driven this variation. For example, in Japan, the new Labour Standards Act limits overtime for construction workers,[50] which is likely to slow down the rollout of new compute; the country also has a lack of power availability.[51] However, India has benefitted from a highly favourable investment environment,[52] which has prioritised data-centre construction at national and state level. Thailand has announced a $1 billion investment from Google[53] for data centres and cloud, while considering new developments in renewable and nuclear energy.[54] Cross-cutting action of this kind is essential for developing a balanced compute ecosystem.

On the human-capital front, Singapore continues to dominate with 400,000 more GitHub contributions per million people than in 2022. One example of its contributions on GitHub is the AI Verify[55] toolkit, which performs standardised tests on AI models to ensure appropriate governance.

FIGURE 14

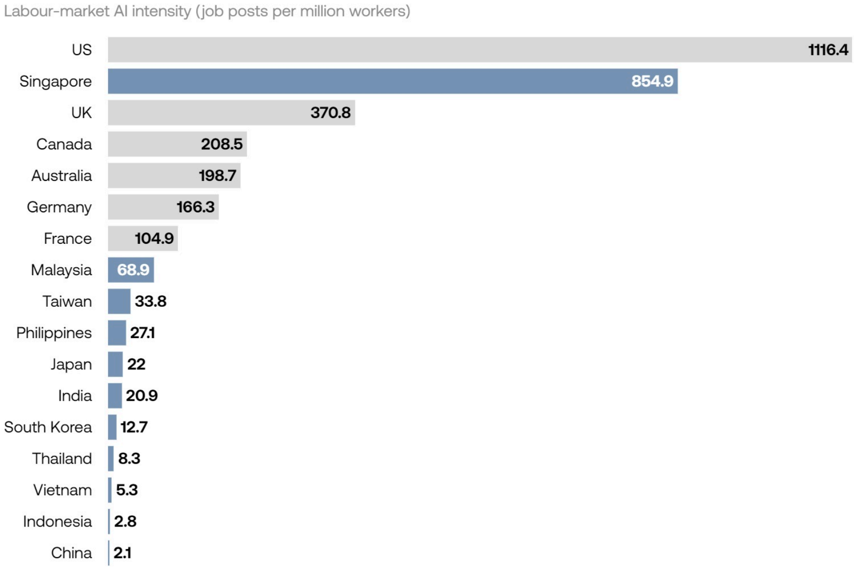
Annual change in GitHub contributions shows Singapore's dominance in Asia and Oceania



Similarly, the Singaporean AI labour market is significantly larger than that of its peers, with demand for an AI-literate workforce one of the highest in the world, and companies are hiring accordingly. This is part of Singapore reaping the benefits of close government involvement in the data-centre industry that has created opportunities for cooperation and critical thinking as to how to improve efficiency throughout their compute ecosystem.

FIGURE 15

Singapore leads the way on AI job postings in Asia and Oceania



Some island and peninsula nations, such as the Philippines and Vietnam, are struggling to develop server capacity as their geography makes the cost of compute prohibitive. Compounded by the uneven growth of the labour force, this issue means that the compute divide in Asia and Oceania seems set to widen. More research is needed on how to work with compute in such geographies: for example, Singapore's Tropical Data Centre Testbed[56] aims to understand how servers can be made energy efficient in humid climates.

05

The Evolution of Compute Ecosystems: New Trends and Dynamics

The analysis of this year's data revealed six major trends and dynamics that are cutting across regions and countries as compute ecosystems evolve. These are reflected in the rise and fall of several countries which have shifted significantly over the last year.

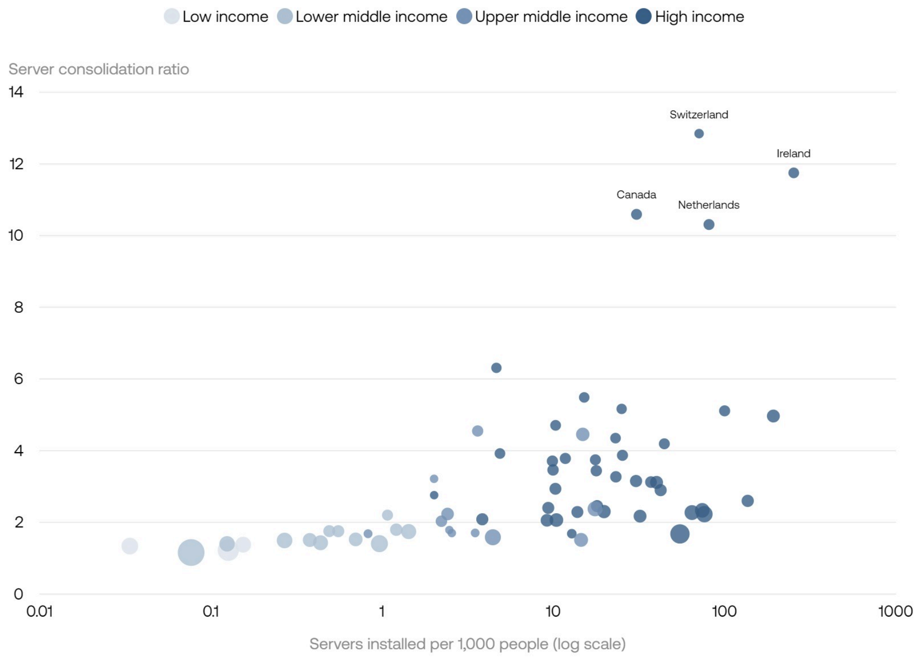
Mature Data-Centre Ecosystems Spend More Per Server Amid Space and Power Constraints

An initial survey of the data suggests that compute growth in digitally mature countries is slowing, and the compute gap might be closing. In the top 20 countries with the highest number of installed servers last year, no high-income country achieved a server growth rate higher than 8 per cent. India, Brazil, Indonesia and Malaysia, the only middle-income countries in the top 20, excluding China, all saw growth between 12 per cent and 17 per cent.

However, this is not a sign of any compute slowdown – rather, a refocusing of investment towards AI-optimised servers. Countries with a significant number of servers and data centres are beginning to replace general-purpose servers with smaller numbers of specialised AI servers. This can be seen most clearly in some of the worlds' data-centre hubs, including Switzerland, the Netherlands and Ireland. Switzerland, for example, is completing the upgrading of data centres for the AI age at a factor three times faster than its economic peer Sweden.

FIGURE 16

Switzerland is part of a breakaway group of countries blazing a trail on the adoption of AI-optimised servers



This trend reflects industry initiatives to overcome the problems of old servers holding up workloads in data centres[57] due to lower levels of energy and efficiency. AMD's Robert Hormuth says the company has customers who have been able to consolidate 15 servers to three, leaving enough power to add additional AI servers into data-centre racks.

Based on our current data, we predict[58] that more digitally mature countries are more likely to invest in a smaller number of highly valuable servers than continuing to invest in large numbers of less valuable servers.

As such, the number of servers will become a less useful indicator of sophisticated compute power, and paradoxically, investment in fewer servers may suggest greater strength and agility. This leaves leaders at a crossroads: should they continue to invest in servers at scale or adopt the new trend of fewer but more powerful servers?

Geography is Crucial to Building Economic Data Centres

FINLAND

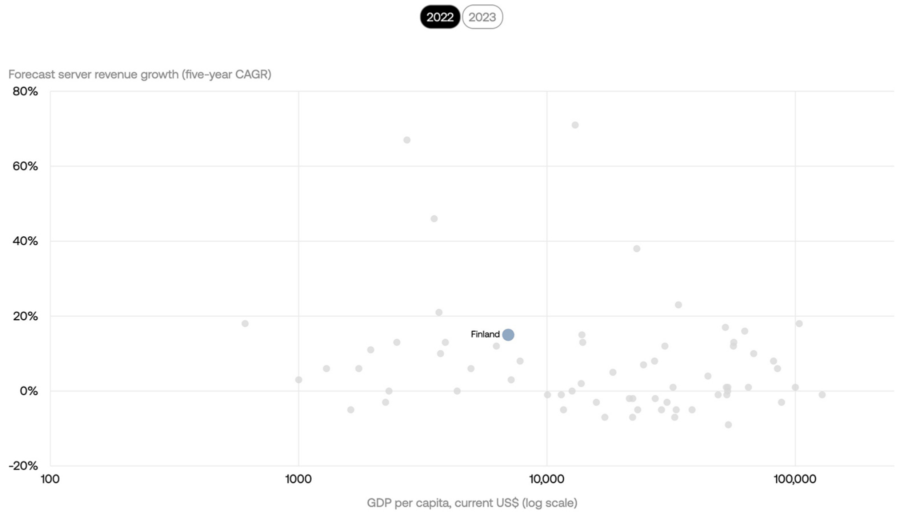
While the US dominates, many countries are leveraging their geographic comparative advantage to catch up. Traditionally not a compute powerhouse, Finland has had a notable rise this year.

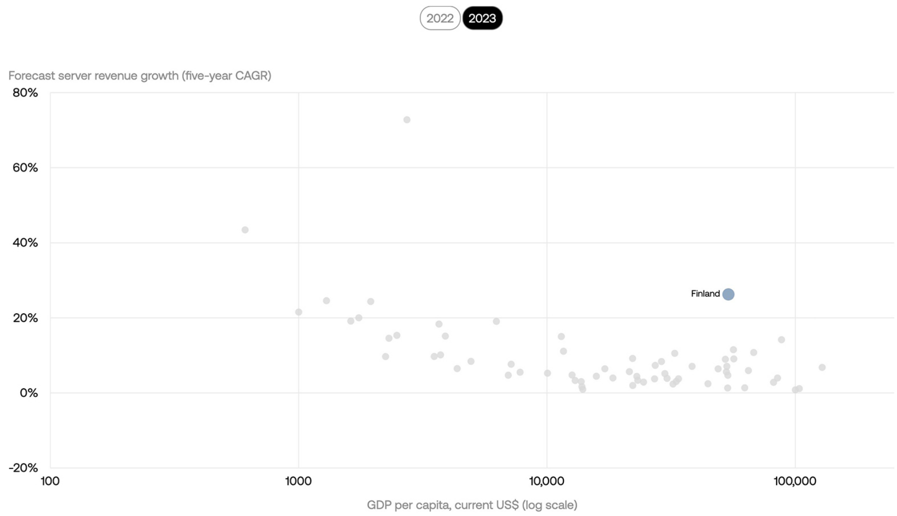
This is partly due to the increase of its server base by 15 per cent, lifting the country from 25th to 23rd position. Finland now leads Europe and high-income countries in server revenue growth, with forecast server revenue growth in 2023 improving from 35 per cent to 44 per cent, moving the country from seventh to third position.

Furthermore, Finland's forecast rate of server unit growth is unusually high for a country of its economic development, and has forecast growth rates more comparable to emerging-market economies.

FIGURE 17

Finland's projected server unit growth is an outlier among peer countries





This growth is projected to continue with major investments from tech giants. In May, Google announced a €1 billion expansion[59] of its Hamina data centre in Finland, and in February,[60] Microsoft announced that Finland would be among five European countries planned as its next data-centre regions.

This accelerated private-sector investment partly rests on Finland's comparative advantage: its climate and energy capacity. Cold weather makes for good computing, as lower air temperatures help cool data centres, making them more energy efficient.[61] It has been further boosted by Finland's impressive clean-energy buildout, using biofuels and an expansive network of nuclear and hydro. For this year's access-to-clean-energy indicator, Finland has the fifth-highest clean energy generated out of total demand at 92 per cent, up from 75 per cent in 2022. It has also increased its energy independence, with imports reduced from 15 per cent to 2 per cent of total energy demand.

Finland also plays an integral role within the European digital landscape. LUMI, Europe's most powerful supercomputer, is housed 600km north of Helsinki by the Finnish CSC (IT Center for Science). LUMI is used by seven other countries in its consortium across northern and northeastern Europe. Finnish startups are said to have exceptional competitive advantage[62] due to their opportunity to access high-performing computing resources – 20 per cent of LUMI's capacity is reserved for Finnish companies. By carefully positioning itself on the independence/capacity trade-off, Finland has been able to leverage continental economies of scale for its own compute agenda.

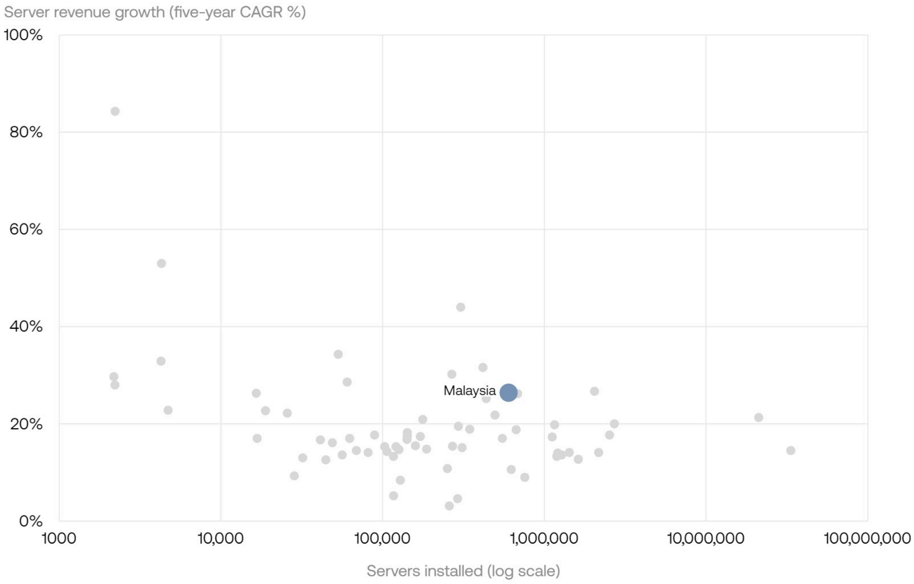
Overall, Finland's digital-infrastructure development offers lessons on harnessing natural resources and comparative advantage. As a nation that is less constrained by access to land and energy, Finland has not only been able to build domestic compute capacity, but also play an important role in supporting Northern Europe research and development (R&D).

MALAYSIA

In the past year, the Malaysian data-centre market has grown enormously. It has invested in building a further 210MW of data-centre capacity, more than the UK, Finland, Netherlands and Singapore. It also had the 13th-highest forecast server revenue growth, and the 18th-highest growth in server installed base.

FIGURE 18

Malaysia's compute buildout is strong, given its current stock of servers



Malaysia's recent development as a compute hub arises from its unique geographical position. Neighbouring Singapore placed a three-year moratorium on the construction of data centres in 2019[63] due to limited energy resources. Seizing the opportunity, Malaysia brought in investment from tech giants such as ByteDance and NVIDIA[64] by cutting approval times to build data centres from more than three months to as little as seven days. As a result, the area of Johor,[65] only a few kilometres from digital giant Singapore, is now providing compute at scale.

Malaysia has been able to maximise this opportunity by investing in its connectivity infrastructure at the same time. It is a strategic landing base for subsea cables, with only three cable landings fewer than Singapore. Malaysia has also recently completed a rollout of 5G,[66] improving connectivity for digital infrastructure, and is second to Singapore on the

ASEAN Digital Integration Index.[67] This growth is expected to continue and Malaysia is forecast to make up two-thirds of the South-East Asian data-centre market by 2035.[68]

Malaysia is a unique example of a country that has been able to capitalise on a neighbour that has high compute demands but limited resources to service those demands. This could also serve as a model for emerging shared-compute ecosystems in other regions as demand for compute outpaces the practicalities of sovereign investment.

Investment Begets More Investment

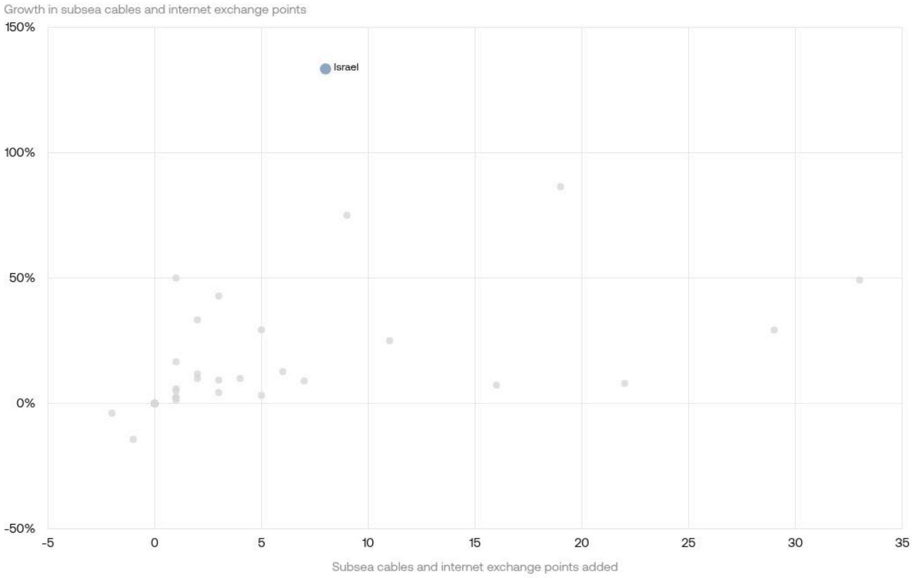
Israel's investment in public compute and infrastructure has cascaded throughout its ecosystem. In November 2023, NVIDIA completed the first phase of its Israel-1 AI supercomputer two months ahead of schedule – a supercomputer that will eventually offer eight exaflops of peak AI performance and 130 petaflops of HPC. While founder and CEO of NVIDIA Jensen Huang identified the talent in Israel – "home to world-leading AI researchers and developers creating applications for the next wave of AI"[69] – as part of the driver for this investment, it is Israel's recent commitment to infrastructure development has been game-changing.

In this year's data, Israel has risen from 38th to 24th on our cloud-service availability indicator, as it expanded from a single cloud or colocation data centre in 2022 to five additional data centres – an expansion of 20,000 per cent (from 4,000 to 804,000 square feet). Furthermore, Israel has shown the largest percentage growth in subsea cables and internet-exchange points, adding a further eight. This rapid growth has been driven in part by Israel's Project Nimbus, first conceived in 2019, to be the basis of the public cloud infrastructure in Israel to facilitate the move of a significant amount of government services to the cloud.[70]

With AWS and Google having won the tender to implement this $1.2 billion project, the subsequent investments play a key part in Israel's rise. This includes investments to increase the electricity grid, including building a new substation to facilitate the supercomputer and additional data centres, as well as a commitment to invest in smart grids and renewable energy[71] to power this growth. The accompanying investment into regional connectivity, such as the 254km subsea cable between the Mediterranean and the Red Sea,[72] is also creating new opportunities for Israel's compute ecosystem.

FIGURE 19

Israel has shown the highest percentage growth in subsea cables and internet exchange points



The infrastructure investment has also catalysed growth in training programmes as companies take advantage of the innovation potential. With an additional three Microsoft partners, two HPE partners and a Linux Foundation partner, Israel has risen from 35th place to 16th place for training programmes. The government has shored up its commitment to being at the forefront of AI research with its policy on AI Regulation and Ethics,[73] published at the end of 2023, which includes an inter-agency AI Policy

Coordination Centre and adopting a common set of principles based on the OECD AI principles. This has all been reflected in the 2024 AI Strategy,[74] which provides insights into how Israel is approaching its AI ecosystem and the role of infrastructure in its journey.

The ongoing conflict has slowed Israel's growth over the past year, but it has recently committed $250 million for a national AI programme,[75] 60 per cent of which will be implemented in 2024. Driven by a pressing need to ensure resources for innovation that can meet its national security requirements supported by new underground data centres that can withstand missile attacks coming online in 2025[76] and its recent accession to the Euro HPC JU cluster,[77] further investments into Israel's compute ecosystem should be expected.

Planning Regulation Can Throttle Growth

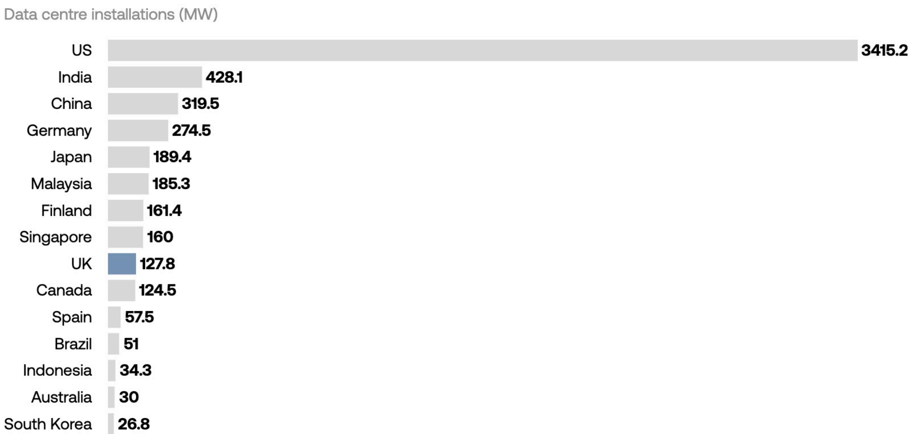
Some countries seem to have hit a plateau or begun to decline due to problems with their planning systems. For example, the UK has a strong compute ecosystem on which to build, with the third-highest number of cloud and colocation data centres; it is also third highest for server installed base, and has the second-highest score for our new indicator covering AI skills demand.

However, the UK's year-on-year growth has been slow to build on this strong starting point. For example, server installed base dropped significantly from 2.95 million servers to 2.7 million, a drop of 8.2 per cent. Some may speculate that this is due to server consolidation, but our data suggest that this is less likely for the UK. Countries with the highest server consolidation ratio include Switzerland, Ireland and the Netherlands, all with consolidation ratios above ten. The UK falls far below this, with a consolidation ratio of only 3.11.

The UK is simply building data centres at a slower pace than its peers. In 2024, the country installed 127MWs' worth of data-centre power. This trails Germany (274MW), Finland (161MW), India (428MW), Japan (189MW), Singapore (160MW) and Malaysia (185MW).

FIGURE 20

The UK is falling behind its peers on the construction of data centres



Many data-centre operators cite the planning system as a common explanation for why the UK has failed to build data centres quickly enough. One council recently refused plans[78] to build a data centre because it would spoil the views from bridges over a busy motorway. The new government described this decision as "pure economic vandalism", and has promised to ease planning restrictions on data centres.[79] Convincing voters that compute is a key driver of prosperity will be a challenge for the government.

While the UK picture appears to be shifting, AWS has announced[80] an £8 billion investment in building, maintaining and operating data centres in the UK; this type of investment needs to be sustained if the UK wants to maintain an edge over competitors.

Human Capital Drives Broader Compute Growth

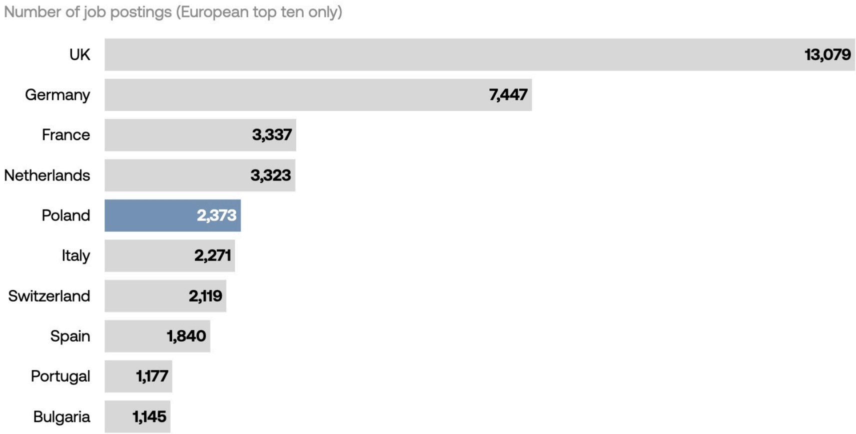
Poland has been experiencing a meteoric economic revolution, which is reflected in the significant growth of its compute infrastructure. The country has risen three places since last year to 14th, overtaking Brazil and Australia, and has added another five supercomputers to the TOP500 list including Helios, the 55th-fastest supercomputer. It has also dramatically increased its server capacity. Key local data-centre operator Atman has raised $345 million this year to expand and upgrade its campus[81] to accommodate 50,000 servers, so this trajectory looks set to continue.

Driven by forward-looking political decision-making and its role in the EuroHPC initiative, Poland established a National HPC Competence Centre.[82] This centre provides access to computing resources and training in high-performance computing, big data and AI. As well as the Helios supercomputer, the government has recently funded the development and upgrading of supercomputers at universities, such as Prometheus at the AGH University of Science and Technology in Kraków.[83]

Poland's infrastructure upgrades have been complemented by a huge rise in its talent pipeline. Benefitting from a large influx of Ukrainian skilled IT workers – with 59 per cent of Ukrainian startups choosing to relocate here[84] – Poland has overtaken Switzerland, Spain and Italy in opportunities for AI skilled workers.

FIGURE 21

Poland's AI recruitment drive is outstripping some of its European counterparts



This demand for AI skills is also being met with an improved supply of technical skills. Poland has upped its availability of training programmes, and the combination of a technically proficient English-speaking workforce with lower wages has led some experts to claim that the country has become an "outsourcing hub"[85] for the tech ecosystem. This ability to leverage the spillover effects of its geographic location together with opportunities from its position in the EU is cementing Poland's growth as an EU compute power.

06

Pathways to Building and Scaling Access to Compute

Our compute measurement framework sets out the key drivers of building compute capacity. However, unlocking the core enablers of compute capacity – be it energy, investment or skills – is by no means an easy task. Political leaders will face core bottlenecks when it comes to delivering new intelligence-era infrastructure. What will be crucial, particularly for developing and resource-constrained nations, is for countries to have a strong understanding of the levers of investment that will deliver the maximum amount of impact for their own compute ecosystems and those that will contribute to an overall increase in ability to access compute.

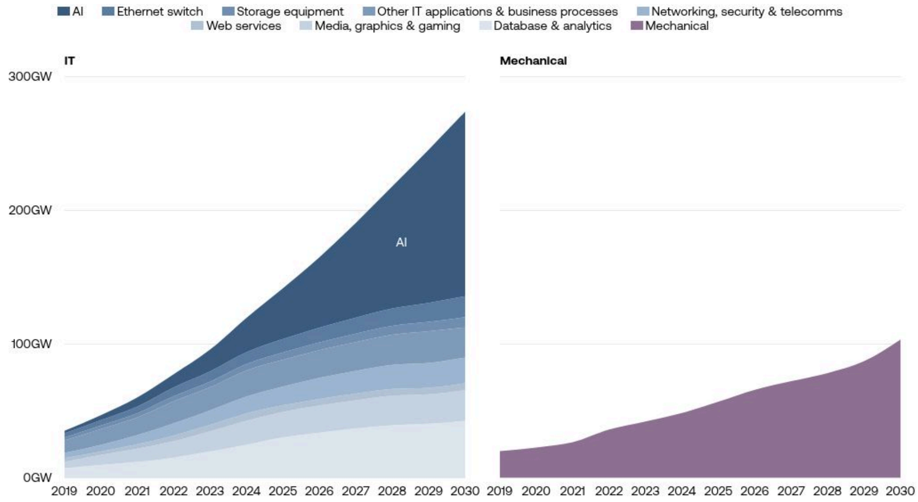
A Clear Compute-Centric Energy Policy

Access to sustainable energy sources remains a defining bottleneck to developing compute capacity, especially as AI workloads demand increasingly more energy. According to the International Energy Authority (IEA), data centres already account for 1 per cent of global electricity use and this is expected to increase.[86] Variation between countries complicates this picture further: data centres account for 7 per cent of energy consumption in Singapore and 14 per cent in Ireland.

Simply keeping pace with the current demand is difficult for many countries. In the UK, for example, the queue for new grid connections extends to the late 2030s.[87] As AI forms a larger share of compute workload, Omdia forecasts a doubling of energy demands from data centres by 2030, with AI workloads accounting for half of this increase.

FIGURE 22

Half of the world's data-centre energy demands are due to come from AI by 2030



This extra demand comes from the computational complexity of AI; giving a simple query to a chatbot powered by a large language model may use nearly ten times more energy[88] than the equivalent on Google Search. While chips are exponentially more energy efficient,[89] providing enough energy for training and inference of future frontier models, this is still likely to present significant energy challenges.

The huge increase in the US's compute capacity correlates with its increase in energy production.[90] The White House, recognising the scale of this challenge, recently established a Taskforce on AI Data-Center Infrastructure[91] to focus on generating more power. Some countries that are already feeling the strain of energy requirements have imposed a one-year moratorium on data-centre builds.[92] Data centres also increase the demand for water required for cooling. In some contexts, such as in Uruguay,[93] which is suffering from drought, this is leading to water shortages and protests.

Some technology companies are seeking innovative ways to tackle these problems: Microsoft is investing in nuclear power[94] for its latest data centres and Amazon[95] is investing heavily in cooling technologies that require significantly less water. However, this technology is not widespread, and water concerns may remain as an obstacle to building best-in-class infrastructure.

As countries consider their energy policies, those that seek to cement their position in the compute ecosystem and underpin their access to AI resources will need to put their compute-power needs at the centre of this policy. Both Australia and Japan are banking on green approaches to help consolidate their compute power. Pawsey's Setonix supercomputer in Western Australia not only employs a direct liquid cooling system, but also a unique groundwater cooling system that circulates heat without a net loss of groundwater, saving approximately seven million litres of water a year.[96] Similarly, Japan has some of the world's greenest supercomputers, with RIKEN's Fugaku being 70 times more powerful than its predecessors with 50-times greater power efficiency. Japan is also encouraging innovative approaches in cultivating energy-saving practices in its research community, such as credits that prioritise their priority in the job-scheduling queue when researchers maintain their energy consumption below a specified limit.[97]

Working with partners, governments should prioritise policies that can green data centres as well as drive innovation in energy-efficiency R&D to enable a successful long-term compute vision.

Manageable Planning and Construction Policies

Building compute capacity involves large, complex infrastructure projects that are likely to have huge delays.[98] Power stations, subsea cables and data centres take years to plan and bring online.[99] Epoch AI has noted that the timeline for infrastructure development is daunting:[100] it can take two years to establish electrical-transformer delivery pathways, five years to secure a place at the front of the grid interconnection queue, and typically a decade to construct transmission lines.

This presents a serious execution problem for emerging compute nations and contributes to the widening compute divide. However, it is equally problematic in the developed world, where projects are often cancelled[101] due to local protests. For example, a group of residents in Tokyo petitioned to block a 3.63-million MW data centre[102] on environmental and resource-depletion grounds. Projects often repeatedly fail to deliver on time and on budget,[103] and a shortage of skilled project managers prevents effective development in newer markets where it is considerably harder to build.[104] Some data-centre operators would like to self-generate power but if this capacity exceeds certain limits they would be treated as a utility. This is preventing investments in self-generation including in hydrogen fuel cells, solar, wind and gas turbines.

There are alternatives to localised, veto-prone planning processes that leaders can deploy, such as giving call-in powers to relevant government departments and reimbursing local residents who are disturbed by construction work.[105] There are also synergies to exploit, such as using heat from data centres to warm homes and offices.[106] Leaders should seek to ensure that new investments have general support to protect them from the disruption of short electoral cycles that could derail access to compute in the long term.

Access to Finance

Building compute is expensive and creates short-term fiscal pressures. The cost of data centres has risen steadily across the world.[107] In Singapore, prices rose from $6.40 per watt to $11.40 per watt in four years – even as chips become more efficient.[108] As leaders absorb these costs, countries have been cancelling compute projects, such as the £1.3 billion exascale supercomputer in Edinburgh.[109]

The increasing cost of data centres is compounded by the fragile fiscal position of countries, especially in the developing world[110] where access to credit is tighter. While investment can bring economic growth, leaders need to consider the risk that generative AI in particular may not be profitable enough to support a rich ecosystem in every region.[111]

Blending public and private finance provides opportunities to overcome the financial hurdles around compute. The new Danish supercomputer emerged from a ground-breaking public-private partnership between the Novo Nordisk Foundation and the Export and Investment Fund (EIFO) of Denmark, in which the EIFO has a 15 per cent stake in the business. Meanwhile, the Labour government in the UK has confirmed Blackstone's £10 billion investment[112] in a new AI data centre.

Governments seeking to lead in compute and access compute infrastructure will need to be flexible and innovative about how they fund compute development, while at the same time demonstrating to taxpayers that investment into national compute ecosystems is money well spent. AI has the potential to transform the quality and efficiency of the state and the lives of those who pay for it. Leaders who can connect the potential of compute to the lives of their electorate will find it much easier to raise the finance they need.

Investment in Compute-Centred Development of Human Capital

Even if political leaders can overcome the material constraints, there is still the issue of shortages in human capital to enable compute ecosystems. This includes shortages in expert capabilities to operate the data centres themselves, as well as wider digital skills.

It takes around 20 people to run a medium-sized 12MW site, but 42 per cent of operators struggle to retain staff.[113] This includes workers from across disciplines – from IT skills, facilities management and electrical-engineering expertise. This is driven by a lack of young people in training ready to replace retiring staff.[114]

There is also a wider digital-skills shortage, especially in the data-analysis and cyber-security domains[115] in all parts of the world. Shortages are so severe that companies are looking offshore to meet their needs[116] in combatting cyber-threats that are developing at increasing levels of sophistication.[117] Exacerbated by a growing digital-skills gap between men and women,[118] most governments are facing an urgent need to work with their education sectors and industry to produce the skilled workers needed to build compute capacity. These skills shortages also push leaders towards hyperscalers rather than domestic compute.

The EU recently announced €55 million in digital-skills training,[119] and the US has invested in simple ways to advertise immigration options available to AI-skilled workers.[120] Given the powerful network effects that have dominated the tech world for the past 20 years, it will be those first movers who can attract, train and retain global talent who will be at the forefront of compute for decades to come.

Ensuring a Geopolitically Resilient Compute Ecosystem

As critical parts of the tech ecosystem are becoming domains of geopolitical competition and are being securitised, leaders face the challenge of balancing development of their ecosystems with their need for security and sovereignty.

Caught in the middle of US export restrictions on AI chips, countries may find themselves unable to source the parts they need to build their compute infrastructure. This challenge is intensified in emerging digital economies that are built with Chinese infrastructure. Huawei has built 70 per cent of the 4G networks in Africa,[121] and China's biggest fibreoptic-cable manufacturer is building a massive cable to connect the East African seaboard to Pakistan.[122] China is now rushing and struggling to replace Western hardware with local alternatives.[123] And as the rivalry potentially expands into China's dominance of the rare-earth-minerals market, lack of raw materials for building digital products[124] will lead many leaders to pick the ally that can best meet the needs of their compute-development goals.

Geopolitical tensions continue to pervade every layer of the tech stack, as choices of ally could result in the complex networks of digitised assets and critical infrastructure being more susceptible to malicious activity. Storing digitised sensitive assets in the cloud may provide access to cutting-edge security but countries fear association with US cloud providers (Amazon, Microsoft, Google) which are governed by the CLOUD Act,[125] even if the server is situated outside of the US.

Narratives of digital colonialism[126] are making emerging digital economies reluctant to partner with larger tech companies, particularly those from the US. Continued concerns over privacy, espionage, political influence and economic capacity[127] are seen through the lens of a new form of imperialism, and prevent many states from engaging with tech partners to build compute efficiently and securely. This problem is especially acute for the smallest and least economically developed countries, as local providers are even less likely to be able to fill the digital vacuum and the challenge of accessing compute resources is likely to increase.

Developing local tech ecosystems to help counter these fears will require governments to nurture a pool of small but dynamic local champions,[128] as well as seek innovative models of private cloud capacity to provide greater assurances on security and privacy. Leaders who can create the conditions for new partnerships of trusted networks and mutually beneficial public-private partnerships will be those that will accelerate their compute ecosystems.

A Holistic Approach to Compute Development

A country's compute infrastructure does not exist in a vacuum. It needs to be developed as part of a holistic policy that integrates industrial policy, science-and-tech policy, defence policy and foreign policy underpinned by strong financial support. None of the obstacles to scaling a compute ecosystem can be overcome by an individual policy lever, and understanding the linkages within the network of policies is paramount to building a strong and functional ecosystem.

The State of Compute Access data explorer helps countries navigate these trade-offs and understand which of the levers can deliver maximum impact for their investment. The decision to focus on the energy grid or training programmes will have different net effects depending on the both the national and regional ecosystem in which they operate. And understanding the cost benefit of ceding strategic autonomy for long-term economic growth will be key for emerging digital economies' relationships with cloud providers and hyperscalers.

This strategic-autonomy challenge will be the defining decision for countries trying to unlock finance to build compute infrastructure. The lack of competition from top cloud providers means that countries wishing to build at scale using local or alternative options may find themselves lagging behind.

FIGURE 23

The Access to Compute 2024 data explorer is available as an interactive tool

Ultimately, not every country will be able to build and scale a compute ecosystem that meets all its needs, or it may find it illogical to invest in systems that may always be secondary to those best-in-class ecosystems. Instead, countries may need to focus on the areas where they can benefit from their comparative advantage – be that in clean energy, software development or basic research. Smaller nations with limited markets and energy-grid dependence on fossil fuels often struggle to attract investment from hyperscalers, who have less incentive to engage in these regions and have to consider the risks of possible oversupply, especially given the rate at which computer hardware becomes outdated.

Through taking a holistic approach countries can identify their competitive and comparative strengths and decide how these should be nurtured and leveraged to contribute uniquely to the global state of access to compute.

07

Towards New Models of Compute Ecosystems

In the 2023 *State of Compute Access* report, we explored three key options that countries can pursue to increase their access to compute: building public compute, building shared regional infrastructure and engaging with the international-development community. Over the past year organisations and governments have accelerated efforts to democratise access to compute both nationally and regionally. These new models offer avenues and incentives for countries to develop their competitive advantages as participants in compute ecosystems.

Public Compute

One option to build strategic autonomy while leveraging compute for the public benefit is to invest in "public compute". Definitions of public compute vary, but this generally relates to state-built components of the compute stack, including data centres, cloud infrastructure[129] and supercomputers. This compute can then be provided to critical sectors and also be used as a lever to encourage responsible use of AI. As the Commonwealth and the Ada Lovelace Institute have argued,[130] compute should be provided on the basis that companies meet certain safety requirements, or that end outputs can help to contribute towards the digital commons.

This requires investing in the key enablers of a model compute ecosystem. Having an abundance of energy and a mature data-centre ecosystem has made it easier for countries like Malaysia to build their own public-compute capacity. Countries with abundant energy resources and stable business environments, such as Dubai, may focus on attracting and hosting large-scale data centres.

Alternatively, nations with advanced manufacturing skills and startups might invest in developing domestic chip production, as seen in China's efforts to counter US restrictions,[131] while countries with more limited capital, resources and skills but a strong research ecosystem may wish to pursue a strategy of model development. For example, state-of-the-art models can be fine-tuned to be better aligned with national language, values and interests. The Netherlands has begun building GPT-NL[132] with this purpose in mind. Furthermore, this can allow for a more prioritised approach to public benefit, building models with specialised sector-specific objectives in mind.

By strategically leveraging compute for public benefit, nations can not only address immediate sectoral needs but also reinforce both their long-term technological sovereignty and their competitive position in the global digital landscape.

Public-Private Partnerships

Since the 2023 report, new models of mobilising private capital for the development of AI infrastructure have emerged. These provide opportunities to unlock multi-trillion-dollar long-term investment and the potential for technological advancement on a global scale.

Groq's partnership with Aramco[133] to create the world's largest AI inference centre, strategically located in Saudi Arabia, aims not only to support the Saudi government's Vision 2030 to drive digital transformation and economic diversification, but also has the potential to open access to cutting-edge AI capabilities for those nations that lack the resources to build and maintain their own AI platforms. Through creating a global hub for AI infrastructure, Saudi Arabia will reinforce its leadership in the global digital economy and provide a blueprint for how public and private capital can come together to build the advanced AI infrastructure needed to drive the next wave of technological innovation, economic growth and inclusivity on the global stage.

Regional Compute Sharing

Nations with limited energy and domestic compute infrastructure could also seek to build and participate in regional compute models. Estonia, despite being one of the world's most digitally advanced nations, has limited compute resources. By leveraging EuroHPC JU (and, in particular, access to its fastest supercomputer LUMI)[134] it overcomes potential issues in accessing supercomputing resources and AI servers via the cloud.

Similarly, the ASEAN-HPC taskforce seems to provide opportunities across the region so that ASEAN countries can access shared HPC resources such as Japan's Fugaku supercomputer.[135] The hosting of the ASEAN HPC school in Indonesia at the end of 2023 highlights how different countries in the region can take on roles that play to their strengths within a compute-ecosystem approach.[136]

A regional approach provides a larger market for compute providers to sell into, so may provide a stronger rationale for investment. Furthermore, transnational investment will help to reduce the public-spending requirements on any one individual nation.

Some nations with unique geographical relationships and abundant resources (such as Malaysia's proximity to Singapore) are able to participate in such models easily. However, in regions with weaker harmonisation and more fraught political alliances, this prospect becomes considerably more challenging. Many African countries, for instance, heavily restrict data portability across borders.[137] However, the ability for Ukraine to shift its data to the cloud at the onset of the Russian invasion highlights the possibilities opened when data portability challenges are quickly resolved.[138] Key for other countries without strong political integration will be to find practical[139] and regulatory mechanisms to secure corridors so that shared compute access can be realised.

International Development for Access to Compute

If countries lack the right political environment to build shared compute, an alternative could involve looking to the international development landscape for support.

For example, the US Department of State, in collaboration with companies such as OpenAI and NVIDIA, have launched the Partnership for Global Inclusivity on AI.[140] This partnership has allowed the Department of State and NVIDIA to offer compute credits to emerging market economies. Likewise, the Gates Foundation AI Grand Challenges[141] adopts a similar approach, focusing on health care. The Gates Foundation has also stated curating a “Gavi for compute”,[142] to redistribute compute resources in Africa.

Other multi-stakeholder agencies have begun their quest to close the international compute divide. The UN, for example, recently recommended the development of an AI development fund,[143] with objectives including to provide access to compute-poor nations. Such an approach at the level of the UN may create better standards for building and adopting compute; however it may come at the expense of speed. Large multi-stakeholder organisations can require sign-off at multiple international levels, meaning that delivering programmes can take longer than initially anticipated.

Many of these projects are in their infancy, so there are unique opportunities for countries to shape this agenda by understanding their own national demand and engaging with these new initiatives.

08

Conclusion: Access to Compute in 2025

When TBI first designed the Access to Compute report in June 2022, prior to the launch of ChatGPT, the signs of its future significance were already beginning to become evident. But the trajectory of its acceleration has far exceeded initial predictions. With the proliferation of generative-AI tools, and frontier models using ever-increasing amounts of compute capacity, compute is being touted as “the most precious commodity in the world” and potentially the “currency of the future".[144] Crucially, becoming a “manufacturer of intelligence” is fast becoming a fundamental infrastructure for many states.[145]

Looking towards 2025, leaders are going to have to articulate their compute vision and intelligence needs, and chart a dynamic whole-of-society pathway to achieve this within an unpredictable world order.

Building on current trends, emerging compute ecosystems will face the choice of whether to invest in multiple low-end servers, or leapfrog and make strategic investments into fewer more expensive servers that may serve them better over the long term. This can be done through enabling national champions such as in Japan, building strategic partnerships such as in Saudi Arabia or Denmark,[146] or participating in regional initiatives such as in Estonia.

Supercharging an industry needs more than supercharged servers and computers, and leaders will continue to face difficult trade-offs to balance dependencies and growth. The scarcity of relevant data-centre industry skills may require countries to revisit their immigration policies as they opt for attracting migrants as an alternative to training or retraining their domestic workforce. Some countries such as Canada have developed technical talent visas for people in relevant technology professions – for example, for tech entrepreneurs in their startup visa.[147]

Leaders may be faced with the choice to prioritise key strategic sectors and industries that have access to compute resources. If access to compute is a bottleneck to sectoral progress, which sectors face this bottleneck the most, and how do those bottlenecks constrain the long-term political objectives of the nation? For example, some of the US's National AI Research Resource, such as the Berkeley National Laboratory[148] and FutureHouse,[149] has been dedicated to scientific research. This will provide the necessary resources to strategically important sectors for that nation, but it will also limit innovation, creativity and growth in other sectors.

This prioritisation challenge extends throughout a government's spending choices. Supporting the ability to build data centres without increasing the availability of energy and water will have material impacts on other important construction projects. For example, housing projects in West London have stalled due to data centres creating a shortage of electricity[150] and water.[151]

And as the race for compute supremacy mirrors wider geopolitical tension, countries may find themselves in a position of having to bandwagon to one of the supercomputer superpowers for assured compute access, increasing tech bipolarity in an era of geopolitical upheaval. Alternatively, countries seeking to accelerate their own progress may be able to take advantage of the rivalry and develop alternative ecosystems to meet the needs of non-aligned countries, creating a new multi-polarity in the world order. Leaders wishing to avoid either dynamic may seek to develop sovereign capabilities, but risk investing in expensive infrastructure that is still behind the pace of the global leaders. Or they may find that key levers of their ecosystem, in particular talent, has left to work in more developed ecosystems.

This in turn will leave leaders with stark choices as to their institutional and regulatory infrastructure – in particular, approaches to cross-border data-sharing and securing the processing of sensitive national data. Resolving the dilemma between data sovereignty and trusted data-sharing for innovation and growth will be one of the greatest psychological hurdles for many leaders.

In his recent book, *On Leadership*: *Lessons for the 21st Century*, Tony Blair writes that "the priorities are the big things, and because they're big, there can't be 15 of them". The report and explorer provides a tool for leadership to understand compute as a priority and identify the core priorities that will help give their country better access to compute on their journey into the era of intelligence. These priorities may vary for each leader but ensuring that access to compute does not create a new digital divide and further global digital fragmentation is a priority for every leader.

09

Acknowledgements

We would like to extend our thanks to the experts who offered their advice and guidance in the development of this report and data explorer.

Steering committee

- Max Beverton-Palmer, NVIDIA
- Alex Creswell, Graphcore
- Laura Foster, techUK
- Lennart Heim, RAND
- Alison Kennedy, UKRI
- Chloé MacEwen, Microsoft
- Christopher Marsh-Bourdon, Oracle
- Keegan McBride, Oxford Internet Institute
- Pippa Robertson, Ofgem
- Rodolfo Rosini, Vaire Computing
- Andrew Scott, 7percent Ventures
- David Snelling, Fujitsu

External reviewers

- Jon Crowcroft, University of Cambridge
- Richard Stirling, Oxford Insights

Omdia

- Vladimir Galabov
- Elaine Hutton
- Dominika Koncewicz
- Pascal Korte
- Aaron Lewis

An annex of the full methodology behind this report and data explorer is available as a downloadable PDF.

Endnotes

1 https://blogs.nvidia.com/blog/denmark-sovereign-ai-supercomputer/

2 https://www.cnbc.com/2024/09/12/openai-anthropic-and-google-leaders-met-with-white-house-to-talk-ai-.html

3 Servers that have been deployed into a data centre. Servers are pieces of hardware or software that provide functionality to a data centre; they are connected to networks to make data accessible to computers.

4 https://www.theregister.com/2024/10/29/xai%5Fcolossus%5Fnetworking/

5 https://www.datacenterdynamics.com/en/news/drc-bank-sofibanque-moves-to-oadc-data-center-in-kinshasa/

6 https://wayve.ai/thinking/introducing-gaia1/

7 https://kwanelesouthafrica.org/chatgbv/

8 https://epochai.org/blog/introducing-epoch-ai-machine-learning-hardware-database

9 https://www.tweaktown.com/news/101335/nvidia-ceo-jensen-huang-joins-king-of-denmark-to-launch-sovereign-ai-supercomputer/index.html

10 Based on NVIDIA internal estimates including both hardware and software. https://www.home.saxo/en-gb/content/articles/equities/nvidias-blackwell-chips-raising-the-bar-in-generative-ai-30082024

11 https://epochai.org/blog/training-compute-of-frontier-ai-models-grows-by-4-5x-per-year

12 https://www.oliverwyman.com/our-expertise/insights/2022/sep/future-of-cloud-and-its-economic-impact.html

13 https://www.bcg.com/publications/2024/long-term-forecast-for-quantum-computing-still-looks-bright

14 https://www.hpcwire.com/2023/11/21/chinas-hpc-iron-curtain-creating-a-top500-problem/

15 https://osf.io/8yp7z

16 https://ethz.ch/en/news-and-events/eth-news/news/2024/09/press-release-new-research-infrastructure-alps-supercomputer-inaugurated.html

17 https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/

18 https://artificialintelligenceact.eu/high-level-summary/

19 https://new.nsf.gov/focus-areas/artificial-intelligence/nairr

20 https://www.computerweekly.com/news/366584173/Bristol-goes-live-with-UK-AI-

supercomputer

21 https://ec.europa.eu/commission/presscorner/detail/en/IP%5F23%5F5739

22 https://www.un.org/ai-advisory-body

23 https://eurohpc-ju.europa.eu/index%5Fen

24 https://www.fz-juelich.de/en/ias/jsc/jupiter

25 https://commission.europa.eu/news/tackling-labour-and-skills-shortages-eu-2024-03-20%5Fen.

26 https://www.iea.org/topics/russias-war-on-ukraine

27 https://www.datacenterdynamics.com/en/news/google-signs-digital-infrastructure-mou-with-romanian-government/

28 https://www.ft.com/content/53accefd-eca7-47f2-a51e-c32f3ab51ad5

29 https://datacentremagazine.com/articles/the-nordics-a-leading-sustainable-data-centre-destination

30 https://www.researchandmarkets.com/reports/5314774/data-centre-nordics-2022-to-2026-report

31 https://www.thenationalnews.com/business/energy/2024/09/15/uaes-clean-energy-share-reaches-2783-of-total-mix-energy-minister-says/

32 https://submarine-cable-map-2024.telegeography.com/

33 GitHub is a site for hosting and reviewing code. Code that is developed in this way is usually called "open source", and other open-source platforms exist. A "contribution", also called a commit or a push, is one instance of a developer adding or editing code on a platform like GitHub. An increase in contributions would represent an increase in the innovation ecosystem and the number of skilled professionals contributing towards it.

34 https://fortune.com/2024/09/24/sam-altman-blog-cheap-ai-water-middle-east/

35 https://groq.com/news%5Fpress/aramco-digital-and-groq-announce-progress-in-building-the-worlds-largest-inferencing-data-center-in-saudi-arabia-following-leap-mou-signing/

36 The combined worth of all the other countries in the set excluding China is estimated at 4,140.6MW.

37 https://news.utexas.edu/2024/08/21/the-state-of-the-supercomputers/

38 https://www.quantum.gov/strategy/

39 https://www.theglobeandmail.com/business/article-canadas-ai-infrastructure-does-not-compute/

40 https://www.quantum.gov/wp-content/uploads/2024/08/Advancing-International-Cooperation-in-QIST.pdf

41 https://www.g20.org/en/news/brasil-launches-a-usd-4-billion-plan-for-ai-and-prepares-global-

action

42 https://www.worldbank.org/en/news/press-release/2024/04/10/revitalizing-growth-an-urgent-agenda-for-latin-america-and-the-caribbean

43 https://www.cbre.com/insights/reports/global-data-center-trends-2023

44 https://www.datacenterdynamics.com/en/news/raxio-launches-data-center-in-the-ivory-coast/

45 https://physicsworld.com/a/africas-quantum-future-offers-a-beacon-of-hope/

46 Installations that include a solar array to capture solar energy, and an energy-storage solution to store solar energy for future use. https://www.velosolar.com/solar-plus-storage/

47 https://datacentremagazine.com/critical-environments/solar-plus-storage-powering-nigerias-data-centre-revolutio

48 https://datacentremagazine.com/critical-environments/solar-plus-storage-powering-nigerias-data-centre-revolutio

49 The decrease in server shipments to China is likely driven by China's investment in AI-optimised servers.

50 https://www.turnerandtownsend.com/insights/japans-digital-transformation-risks-and-opportunities-in-the-data-centre-market/

51 https://www.datacenterdynamics.com/en/analysis/navigating-the-insular-growth-of-japans-data-center-sector/

52 https://www.cbre.co.in/insights/articles/indias-booming-data-centre-market-a-powerhouse-for-the-digital-age

53 https://www.googlecloudpresscorner.com/2024-09-30-Google-Announces-Plans-to-Invest-US-1-Billion-to-Build-Data-Center-and-Cloud-Region-in-Thailand,-Support-Initiatives-to-Expand-AI-Opportunities-for-Thais

54 https://www.nationthailand.com/sustainability/renewable-green-energy/40038917

55 https://github.com/aiverify-foundation/aiverify

56 https://news.nus.edu.sg/nus-and-ntu-launch-first-of-its-kind-tropical-data-centre-testbed/

57 https://www.fierce-network.com/ai/there-are-millions-old-servers-use-amd-says-thats-problem-ai

58 "Server consolidation ratio" is a ratio of forecast growth in server revenue to server units over the next five years. Countries with a higher ratio are forecast to invest in a smaller number of highly valuable AI servers.

59 https://www.businessfinland.fi/en/whats-new/news/2024/google-announced-an-expansion-of-its-hamina-data-center

60 https://azure.microsoft.com/en-us/blog/microsoft-supports-cloud-infrastructure-demand-in-europe/

61 https://finland.fi/business-innovation/search-engine-giants-find-innovative-energy-for-data-centres-in-finland/

62 https://www.eurocc-access.eu/finnish-start-ups-have-an-exceptional-competitive-advantage-this-is-something-that-should-be-announced-more-loudly-says-harri-koskimaki-from-candour-oy/

63 https://www.datacenterdynamics.com/en/analysis/cracking-green-conundrum-singapore-amid-data-center-moratorium/

64 https://www.freemalaysiatoday.com/category/opinion/2024/05/18/malaysias-tech-scene-set-to-soar-with-surge-in-investments-by-tech-giants/

65 https://www.malaymail.com/news/malaysia/2024/10/17/why-do-foreign-investors-prefer-to-set-up-data-centres-in-johor-instead-indonesia-communications-ministers-gives-three-reasons/153987

66 https://www.ericsson.com/en/press-releases/2/2024/10/malaysia-carries-out-first-5g-live-broadcast-by-rtm-in-collaboration-with-dnb-and-ericsson.

67 https://www-file.huawei.com/-/media/corporate/local-site/my/pdf/huawei%5Fmdec%5Fwhite%5Fpaper%5F7%5Foctober%5F2022.pdf

68 https://www.ft.com/content/4d8ab5e8-a7a6-4850-a631-5e9e2a4c13bb

69 https://www.datacenterdynamics.com/en/news/nvidias-israel-1-supercomputer-starts-operations/; https://www.datacenterdynamics.com/en/news/nvidias-israel-1-supercomputer-starts-operations/.

70 https://www.gov.il/en/pages/press%5F24052021

71 https://www.saurenergy.com/solar-energy-news/israels-data-centre-industry-majorly-driven-by-renewable-power

72 https://www.techloy.com/israel-to-build-fibre-optic-link-between-europe-and-asia/

73 https://www.gov.il/en/pages/ai%5F2023

74 https://www.gov.il/BlobFolder/news/most-news20231218/he/Israel National AI Program 2024.pdf

75 https://www.datacenterdynamics.com/en/news/israel-looks-to-build-ai-supercomputer/

76 https://www.datacenterdynamics.com/en/news/techtonic-planning-16mw-underground-data-center-in-bet-shemesh-israel/

77 https://eurohpc-ju.europa.eu/israel-joins-eurohpc-joint-undertaking-2023-10-06%5Fen

78 https://www.telegraph.co.uk/money/net-zero/25bn-super-hub-data-centre-blocked-spoiling-m25-views/

79 https://www.datacenterdynamics.com/en/news/uk-labour-party-to-ease-planning-restrictions-on-data-centers-report/

80 https://www.aboutamazon.co.uk/news/job-creation-and-investment/aws-plans-to-invest-8-billion-in-the-uk

81 https://www.datacenterdynamics.com/en/news/atman-raises-345m-to-build-data-centers-in-poland/

82 https://www.eurocc-access.eu/about-us/meet-the-nccs/ncc-poland/

83 https://www.cyfronet.pl/en/18378,artykul,prometheus%5Fsupercomputer.html

84 https://www.careersinpoland.com/article/news/report-59-of-startups-relocating-from-ukraine-choose-poland

85 https://emerging-europe.com/opinion/forget-silicon-valley-poland-is-europes-emerging-tech-epicentre%EF%BF%BC/

86 https://www.iea.org/energy-system/buildings/data-centres-and-data-transmission-networks

87 https://www.ofgem.gov.uk/blog/preparing-faster-more-efficient-electricity-connections-process

88 https://www.goldmansachs.com/insights/articles/AI-poised-to-drive-160-increase-in-power-demand

89 https://epochai.org/blog/trends-in-machine-learning-hardware#energy-efficiency

90 https://www.eia.gov/energyexplained/us-energy-facts/

91 https://www.datacenterdynamics.com/en/news/white-house-to-launch-task-force-on-ai-datacenter-infrastructure-speed-up-data-center-permitting/

92 https://www.datacenterdynamics.com/en/analysis/the-ongoing-impact-of-amsterdams-data-center-moratorium/

93 https://www.theguardian.com/world/2023/jul/11/uruguay-drought-water-google-data-center

94 https://www.nytimes.com/2024/09/20/climate/three-mile-island-reopening.html

95 https://sustainability.aboutamazon.com/natural-resources/water

96 https://pawsey.org.au/australias-setonix-ranking/

97 https://www.asianscientist.com/2024/10/print/whats-next-in-supercomputing/

98 https://doi.org/10.1080/07421222.2022.2096544

99 https://www.spinergie.com/blog/new-challenges-lay-ahead-as-demand-for-offshore-cables-increases

100 https://epochai.org/blog/can-ai-scaling-continue-through-2030

101 https://www.bbc.co.uk/news/articles/c7285e02qy4o

102 https://www.reuters.com/world/asia-pacific/tokyo-residents-seek-block-building-massive-data-centre-2024-07-10/

103 https://committees.parliament.uk/committee/127/public-accounts-committee/news/201373/uk-lacks-skills-and-capacity-to-deliver-major-infrastructure-pac-warns/

104 https://images.connect.arcadis.com/Web/

Arcadis/%7B856573cb-2409-4448-96a6-852638aa24fd%7D%5FInternational-Construction-Costs-2024.pdf

105 https://www.samdumitriu.com/p/how-to-get-data-centres-built-quickly

106 https://www.gov.uk/government/news/thousands-of-homes-to-be-kept-warm-by-waste-heat-from-computer-data-centres-in-uk-first?IYA-mail=88a8a1cd-eaa0-40c6-83ed-0602552be92b

107 https://reports.turnerandtownsend.com/dcci-2023/data-centre-cost-trends

108 https://doi.org/10.1126/science.aba3758

109 https://bmmagazine.co.uk/in-business/uk-government-cancels-1-3-billion-ai-and-tech-funding-amid-economic-tightening/

110 https://www.un.org/sustainabledevelopment/blog/2024/01/overview-world-economic-situation-and-prospects-2024/

111 https://www.goldmansachs.com/insights/top-of-mind/gen-ai-too-much-spend-too-little-benefit

112 https://www.bbc.co.uk/news/articles/c3e957k9d1yo

113 https://www.jll.co.uk/en/trends-and-insights/workplace/how-the-data-center-industry-is-bridging-the-skills-gap

114 https://www.telehouse.net/blog/skill-shortages-in-the-data-centre-industry/?gad%5Fsource=1&gclid=Cj0KCQjwmOm3BhC8ARIsAOSbapVLY2um%5FkmkGi%5FYoI3BQsW3Yn4e76sA9a1ZvXk

115 https://impact.economist.com/perspectives/talent-education/bridging-skills-gap-fuelling-careers-and-economy-india

116 https://www.incommon.ai/blog/tackle-ai-talent-shortage-with-offshoring/

117 https://www.ncsc.gov.uk/collection/annual-review-2023

118 https://doi.org/10.54675/RAPC9356

119 https://digital-strategy.ec.europa.eu/en/news/digital-europe-programme-invests-over-eu108-million-european-digital-skills-capacity-and-tech

120 https://ai.gov/immigrate/

121 https://foreignpolicy.com/2019/03/19/for-africa-chinese-built-internet-is-better-than-no-internet-at-all/

122 https://www.ipcs.org/comm%5Fselect.php?articleNo=5776

123 https://www.reuters.com/technology/china-rushes-swap-western-tech-with-domestic-options-us-cracks-down-2023-10-26/

124 https://www.nature.com/articles/d41586-023-02330-0

125 Clarifying Lawful Overseas Use of Data Act.

126 https://journals.sagepub.com/doi/10.1177/0306396818823172

127 https://sur.conectas.org/en/digital-sovereignty-or-digital-colonialism/

128 https://www.researchgate.net/profile/Daniel-Isenberg/publication/309328841%5FFostering%5FScaleup%5FEcosystems%5Ffor%5FRegional%5FEconomic%5FGrowth%5FInnovations%5FC
Mas%5Fand%5FScale%5FUp%5FMilwaukee/links/624432cc57084c718b7644e5/Fostering-Scaleup-Ecosystems-for-Regio
Innovations-Case-Narrative-Manizales-Mas-and-Scale-Up-Milwaukee.pdf#page=70

129 https://ukdayone.org/briefings/gb-cloud-building-the-uk-s-compute-capacity

130 https://www.adalovelaceinstitute.org/blog/the-role-of-public-compute/

131 https://www.lowyinstitute.org/the-interpreter/are-us-export-controls-making-china-s-chip-industry-more-innovative

132 https://gpt-nl.com/

133 https://groq.com/news%5Fpress/aramco-digital-and-groq-announce-progress-in-building-the-worlds-largest-inferencing-data-center-in-saudi-arabia-following-leap-mou-signing/

134 https://www.silicon.co.uk/ai/lumi-the-most-powerful-supercomputer-in-europe-579392

135 https://astnet.asean.org/about-asean-hpc-taskforce/

136 https://www.eeas.europa.eu/delegations/association-southeast-asian-nations-asean/asean-continues-high-performance-computing-hpc-school-programme-after-successful-partnership-eu-and%5Fen?s=47.

137 https://gdprlocal.com/gdpr-and-data-protection-laws-in-africa-a-comparison/

138 https://d1.awsstatic.com/institute/AWS-Institute-Accelerate-public-service-transformation-4-Security-Ukraine.pdf

139 https://www.sciencedirect.com/science/article/pii/S0167404823005151

140 https://www.state.gov/united-states-and-eight-companies-launch-the-partnership-for-global-inclusivity-on-ai/

141 https://www.gatesfoundation.org/ideas/science-innovation-technology/artificial-intelligence

142 https://www.exemplars.health/stories/its-an-understatement-to-say-ai-will-be-a-game-changer-for-global-health

143 https://www.theguardian.com/business/2024/sep/19/global-ai-fund-needed-to-help-developing-nations-tap-tech-benefits-un-says

144 https://lexfridman.com/sam-altman-2-transcript

145 https://blogs.nvidia.com/blog/denmark-sovereign-ai-supercomputer/

146 https://www.tomshardware.com/tech-industry/artificial-intelligence/nvidia-ceo-jensen-huang-and-the-king-of-denmark-plug-in-the-countrys-first-ai-supercomputer-gefion-leverages-1-528-nvidia-h100-ai-gpus

147 https://www.canada.ca/en/immigration-refugees-citizenship/campaigns/tech-talent.html

148 https://nsf-gov-resources.nsf.gov/2023-08/

AI%5FResearch%5FInstitutes%5FMap%5F2023%5F0.pdf

149 https://www.futurehouse.org/research-announcements/wikicrow

150 https://www.newcivilengineer.com/latest/grid-shortages-thousands-of-stalled-london-homes-unblocked-14-03-2024/

151 https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://www.ft.com/content/8d8bf26f-5df2-4ff6-91d0-369500ed1a9c&ved=2ahUKEwiBg8WIouuIAxVyxAIHHY3pEh8QFnoECBYQAQ&usg=AOvVaw1Tad
https://www.ft.com/content/8d8bf26f-5df2-4ff6-91d0-369500ed1a9c

TONY BLAIR
INSTITUTE FOR
GLOBAL CHANGE

Follow us

facebook.com/instituteglobal
x.com/instituteGC
instagram.com/institutegc

General enquiries

info@institute.global