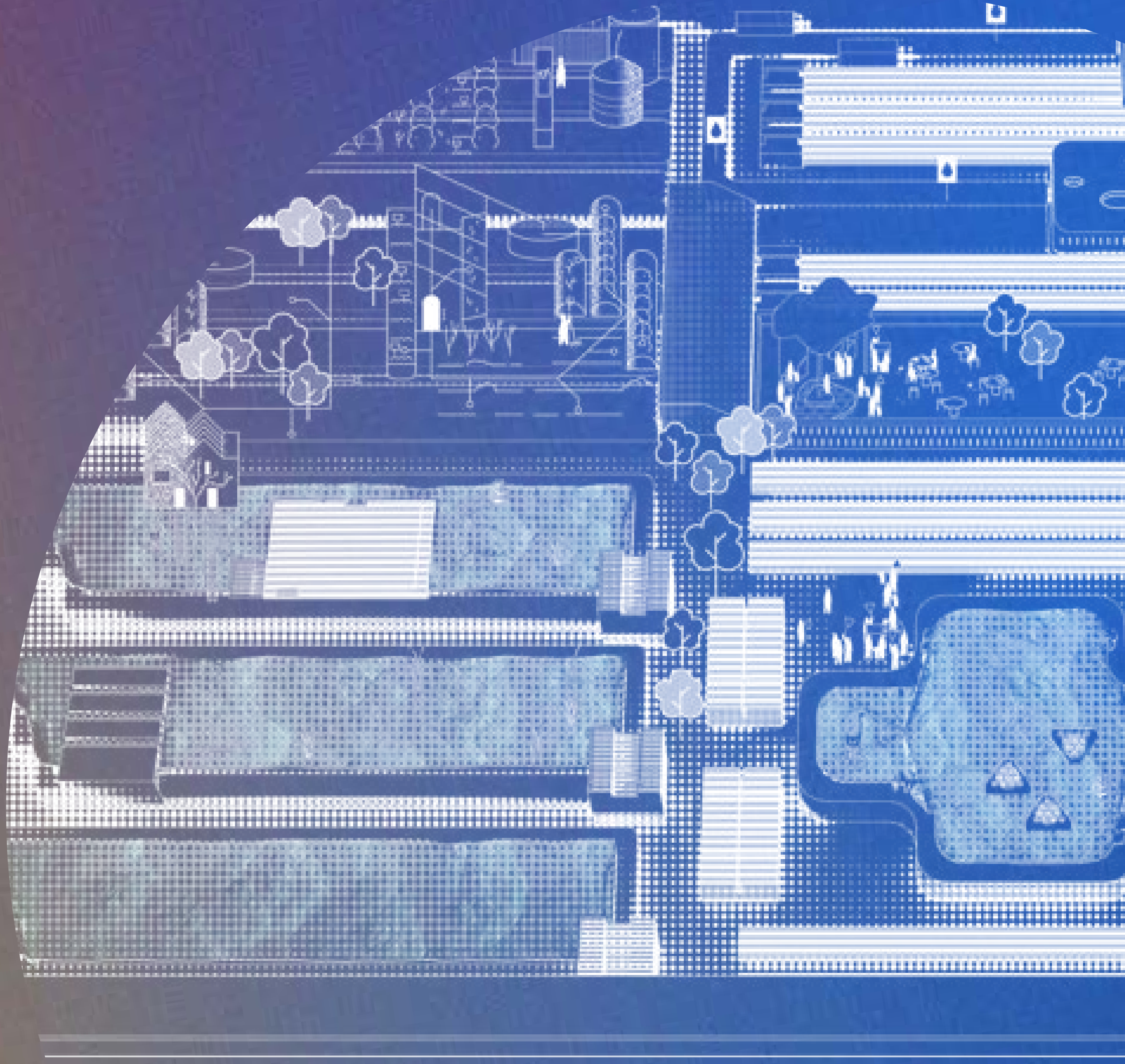
Issue 04

Data Centre Futures:

Data and Compute

How the evolving nature of data and compute might redefine the design and locations of data centres



Foresight
June 2026

Data Centre Futures – looking beyond the asset

Twenty-first century life — from the smallest everyday interactions to large-scale international and even extraterrestrial systems — is increasingly reliant on flows of data. While we would all feel the effects if these flows were disrupted, few of us pause to reflect on the vital digital infrastructure that transforms this data into meaningful insights, the communications, services, and operations that enhance our everyday lives.

If left unchecked, data centres could destabilise the urban metabolism, crowding out resource flows that are vital for short and long term urban development.

Data centres embody a paradox. Their rapid growth is placing increasing strain on key resources, even as our dependence on them continues to intensify. Managing this tension is critical.

Within an ‘urban metabolism’, data centres are hubs through which flows of water, energy, materials, and people — as well as information — move. And as critical infrastructure, which needs to be protected, they can play a central role in shaping safe, resilient and regenerative places.

Data centres are critical nodes in one of the most prolific infrastructural periods in history. Thinking of them merely as ‘assets’ narrows our field of vision and limits the potential for wider, positive impact through their design.

So, what defines a ‘good’ data centre and how can the facility be both a ‘good neighbour’ and a ‘good ancestor’?

Data Centre Futures – looking beyond the asset

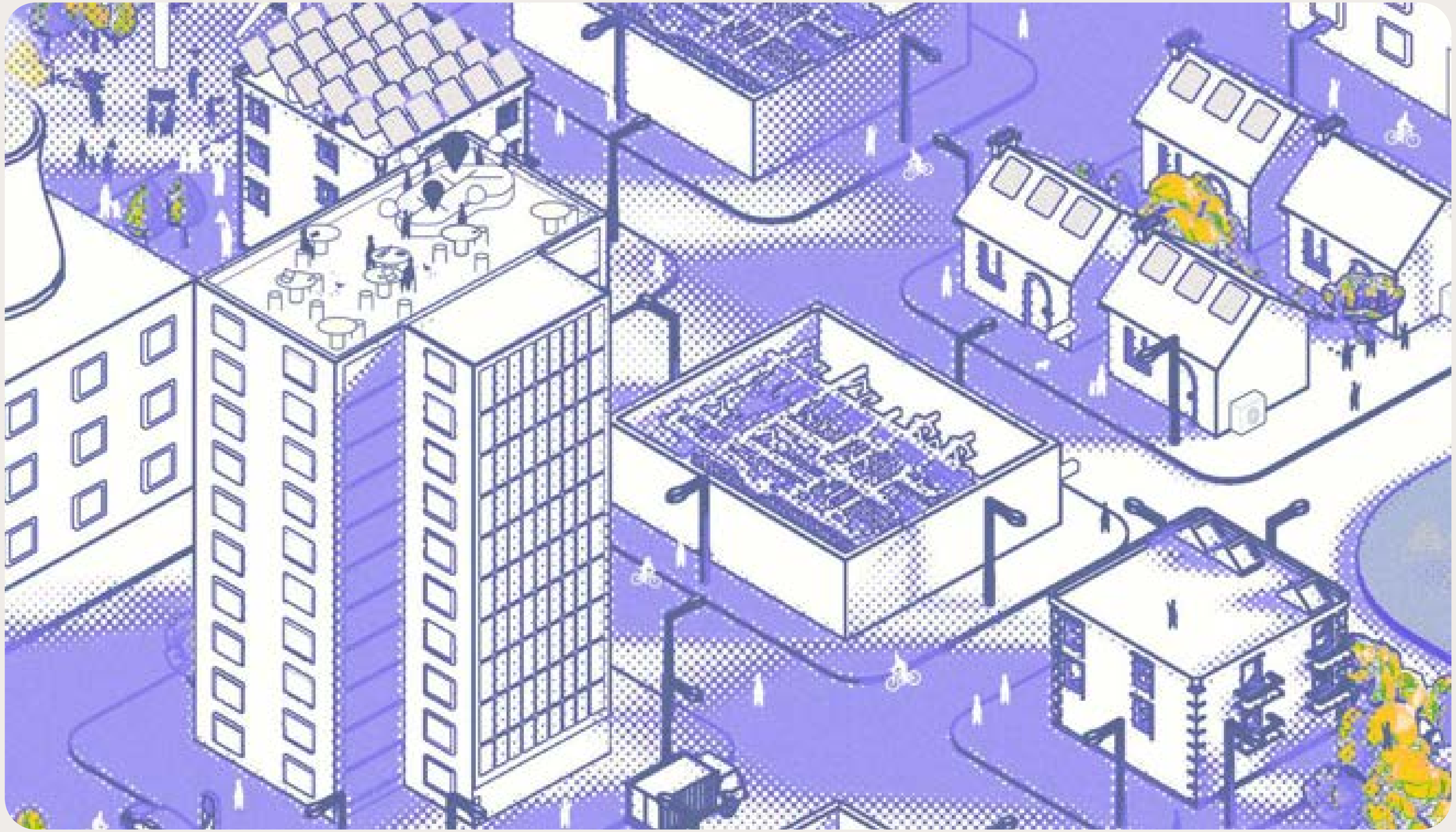
Being a good *neighbour* requires a rethink of 'performance' metrics; a reframing of 'efficiency'; a refocusing on their impact in the areas in which they are built. How can data centres improve their local place and their local ecosystem? Being a good *ancestor* requires an awareness that the infrastructure of today becomes the relics of tomorrow. It also means accounting for the impact of the decisions we make today on future generations.

Creating meaningful impact starts with long-term thinking and designing systems that can adapt to change. What happens when technological and socio-cultural evolutions make data centres, in their current form, redundant?

With this and subsequent issues, we explore the Future of Data Centres, building upon our already broad range of insights on Arup.com.

We aim not to defend a position, but rather to elevate the conversation around data centres, encouraging more of us to ask better questions about dangerous assumptions *and* possible futures in this fast-evolving landscape.

Hit pause, look up and look around.



Data Centre Futures series

In collaboration with technical experts across Arup’s global offices, the Foresight team presents the Data Centre Futures series.

Each issue explores a key theme, emerging issues, trends shaping future context, critical reflections and informed speculation on longer-term possibility.

This **fourth and final issue** of this series focuses on **Data and Compute.**

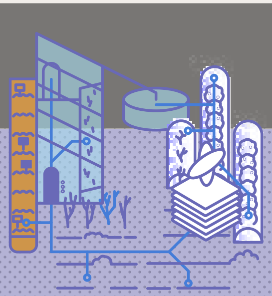
We are entering a period of accelerated technological change, driven by the emergence of several new computing approaches and multiple general-purpose technologies. These developments ripple across the data lifecycle. Data now arrives in vast volumes, from an expanding array of sources, and often in real time. It is processed across multiple architectures and pushed through distributed networks, all while being shaped and constrained by heightened awareness of security and sovereignty.

Yet, this rapid evolution is colliding with a very different reality. The long life, high cost and relative permanence of the physical infrastructure built to enable it. Data centres are being designed and financed as enduring assets at the very moment the technologies they support are becoming more fluid, uncertain and contested. This creates important near-term risks of redundancy and lock-ins alongside significant financial and environmental costs for operators, clients and wider economies.

01 Water

02 Energy

03 Land and space use



04 Data and Compute

Foresight would usually take a longer view to identify deeper patterns of change, in this case, however, there is also an urgent need to confront near-term uncertainty and design for optionality. This moment also creates real opportunities, giving today's data centre designers the chance to become more specialised, adaptable and creative than the sector previously required.

At its best, this sector is becoming one of the most important enabling industries of the next decade, quietly powering breakthroughs in science, public services, and everyday productivity.

We can shape the next era rather than inherit it. A good data centre is more than an asset: it earns its social licence by supporting local communities, ecosystems and resource systems. As a good ancestor, it is designed for long-term value, adaptable, resilient, and leaving a positive legacy rather than stranded infrastructure.

This report examines these shifts across the data lifecycle and asks:

How could emerging and future technological developments drive future design and location of data centres?

How do we ensure that today's dominant digital infrastructure archetype doesn't limit the potential for new types of computing and technologies in the future?

Our findings point to the need for adaptive, context-specific data centre design and location strategies, moving beyond uniform models to reflect distinct data ecosystems.

Mapping uncertainty

Technology drivers impacting data centres

Data centres are critical, capital-intensive infrastructure with long life cycles despite rapid industry change. Decisions made today will shape exposure to technological, regulatory and environmental risk for decades. That lifespan is also a strategic advantage: designing for flexibility now can keep sites useful across multiple technology generations.

Technological change - what we know

Artificial intelligence, advanced analytics and pervasive data collection are driving a surge in demand for computing power, making it one of this decade's most critical resources[1]. As data sources multiply, from IoT devices and industrial sensors to autonomous systems, the volume, velocity and diversity of data continue to accelerate. Demand for compute is unlikely to ease as capabilities expand, bringing new applications, dependencies and expectations.

At the same time, data governance is rising up the political agenda, particularly in a more geopolitically fraught environment. Data security, sovereignty and asset protection are shaping infrastructure decisions. Location is no longer only about cost, connectivity or convenience, but also strategic resilience and autonomy. These pressures extend beyond IT hardware to the wider systems data centres rely on, including energy, water, land and critical materials, which are becoming more constrained and contested, as covered in earlier issues of this series.

Yet scarcity and pressure often drive innovation, pushing new forms of efficiency, new design responses and new technological solutions.

What we don't know

The current data centre landscape will struggle to sustain the pace of projected growth without major changes to power supply, capacity expansion, and operating models. That pressure is also accelerating innovation, forcing better efficiency, clearer workload choices, and more adaptable operating models.

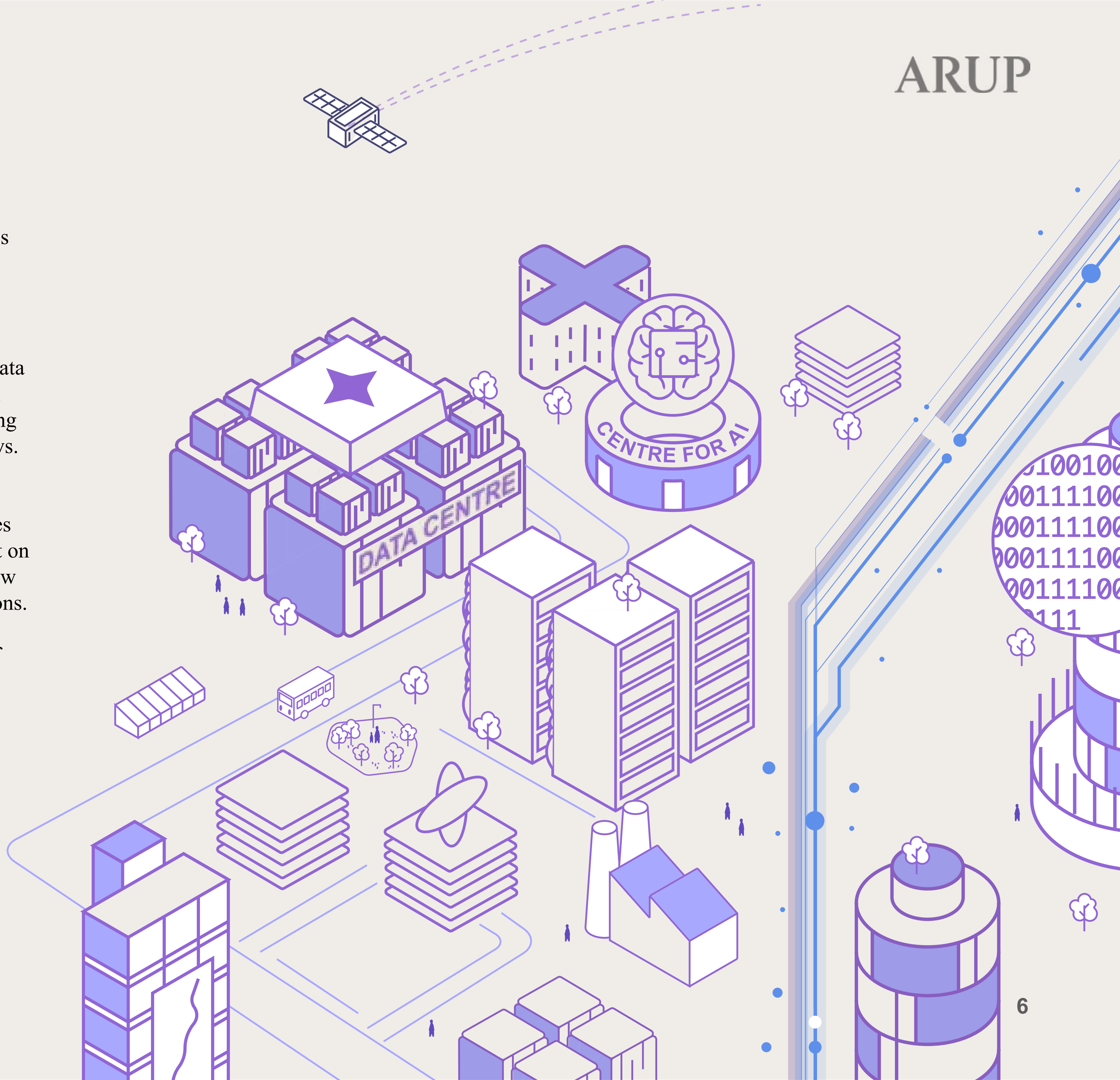
> Emerging technologies and evolving computing paradigms will impact where and how data centres are designed, built and operated. But we don't yet know the precise timing or nature of major breakthroughs nor the pace and depth of their adoptions.

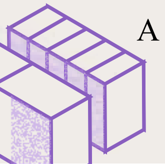
Convergence across technology families and sub-families may generate new spatial, operational and investment needs that are not yet fully visible. The uncertainty also extends beyond technology itself. The future shape of data governance, sovereignty and regulation will vary from place to place, influencing infrastructure decisions in different ways.

Past disruptions such as the move from centralised servers to cloud architectures or the rapid rise of AI workloads reliant on specialised chips, show how quickly new demand can outpace existing assumptions. This makes long-term planning more difficult but also reinforces the value of adaptive and resilient strategies.

To explore this mix of signals and unknowns, we map the landscape of technologies and sub-technologies on a radar. The radar does not predict outcomes, but helps identify in a structured way which emerging technologies might reshape the design, location, and role of data centres.

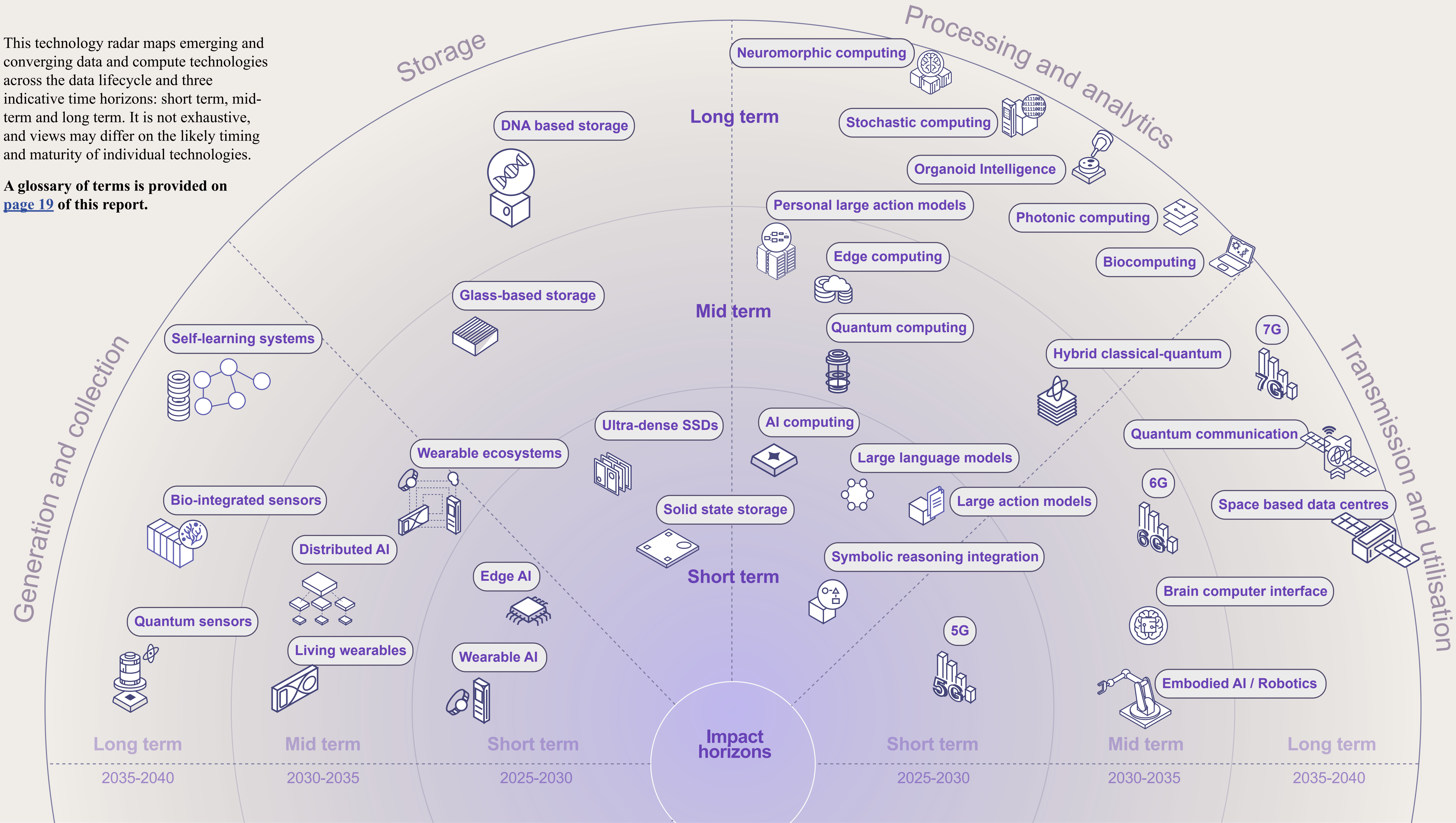
A glossary of terms is provided on page 19 of this report.





This technology radar maps emerging and converging data and compute technologies across the data lifecycle and three indicative time horizons: short term, mid-term and long term. It is not exhaustive, and views may differ on the likely timing and maturity of individual technologies.

A glossary of terms is provided on page 19 of this report.



Storage

Processing and analytics

Generation and collection

Transmission and utilisation

Long term

DNA based storage

Neuromorphic computing

Stochastic computing

Organoid Intelligence

Photonic computing

Biocomputing

Personal large action models

Edge computing

Glass-based storage

Self-learning systems

Mid term

Quantum computing

Hybrid classical-quantum

7G

Ultra-dense SSDs

AI computing

Wearable ecosystems

Large language models

Quantum communication

Bio-integrated sensors

Solid state storage

Large action models

6G

Space based data centres

Distributed AI

Symbolic reasoning integration

Short term

Edge AI

Brain computer interface

Quantum sensors

Living wearables

Wearable AI

5G

Embodied AI / Robotics

Impact horizons

Long term 2035-2040

Mid term 2030-2035

Short term 2025-2030

Short term 2025-2030

Mid term 2030-2035

Long term 2035-2040

Arup expert piece:

Heterogeneous compute: Designing for the multi-paradigm data centre

The coexistence of classical, quantum and other paradigms such as biocomputing will fundamentally fragment the traditional data centre model. We are likely to be moving from a unified facility to a campus of specialised zones, as the infrastructure requirements for each paradigm are radically different.

Location strategy will diverge. While classical computing continues to rely on power grids, quantum will drive facilities toward geologically stable zones to mitigate vibration, and may require proximity to helium supply chains. Biocomputing, conversely, will anchor itself to bio-tech hubs for access to biological reagents and containment protocols.

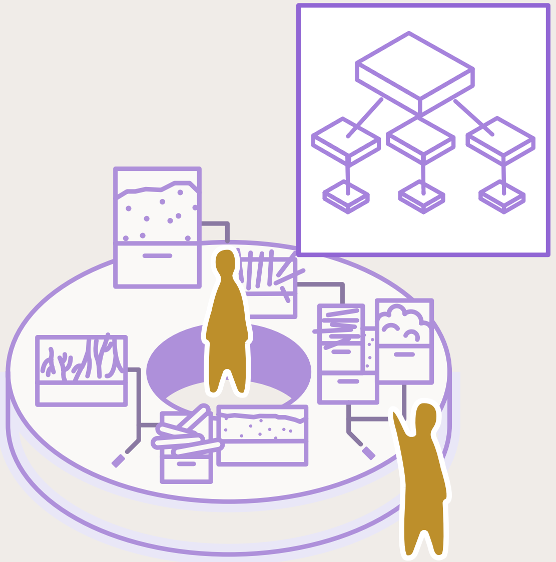
> Design and infrastructure will be defined by extreme segregation. We cannot place a BSL-2 biocomputing lab requiring negative pressure and HEPA exhaust next to a classical server aisle.

Similarly, a quantum processor's need for millikelvin temperatures and EMI shielding is incompatible with the 'noisy' electrical environment of high-density AI racks. This forces a laboratory model over a warehouse model, requiring different flooring, air handling, and waste systems within the same campus.

Organisations are already designing for coexistence, not replacement. Classical compute remains dominant, but existing work on quantum computing facilities and cryogenic environments shows how different future paradigms impose fundamentally different requirements, vibration, EMI, thermal stability, and power quality. That's driving data centres toward modular, heterogeneous campuses and more deliberate location strategies.

The decision to stop scaling classical and switch to another paradigm isn't driven by a single metric like flops or transistor size. It's driven by breakpoints where the physics or economics of classical computing simply collapse.

Jimmy Chan
Associate Director,
Advanced Digital
Engineering, Arup



First, the optimisation wall, where classical algorithms face combinatorial explosion. When a problem requires millions of years of classical compute, like complex logistics or materials simulation, you switch to quantum.

Second, the power wall. A classical AI rack needs 20,000 watts for pattern recognition. When cooling infrastructure can’t dissipate the heat or power budgets collapse, you switch to neuromorphic computing for energy-efficient edge AI.

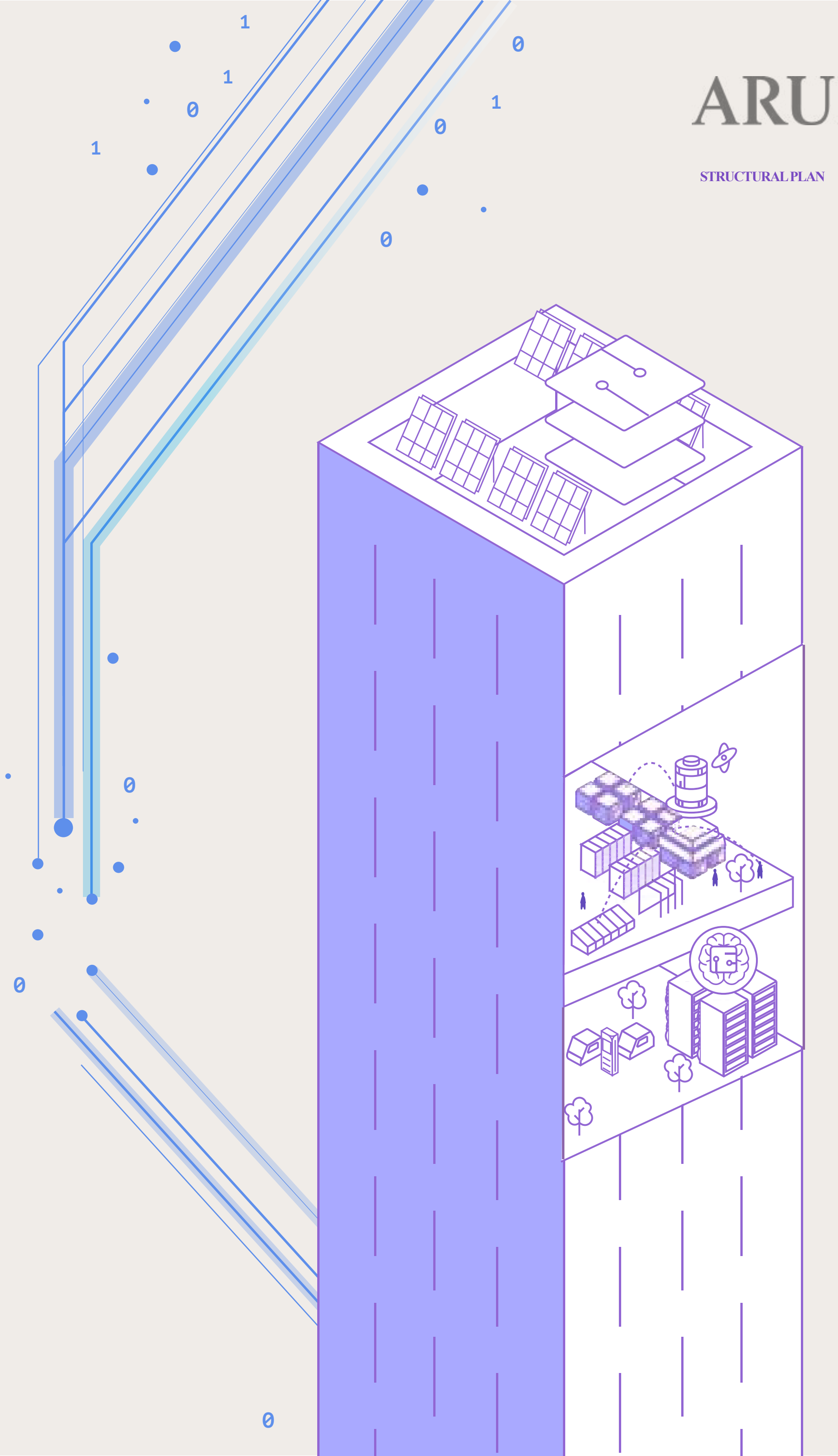
Third, the density wall. Silicon storage hits quantum tunnelling limits around one terabyte per cubic inch. DNA stores 215 petabytes per gram. When the data centre real estate for archival storage exceeds the cost of a wet-lab interface, we switch to biocomputing.

Finally, the latency wall. When the speed of light becomes the bottleneck, such as high-frequency trading or sensor-level processing, we switch to photonic computing at the edge.

> Classical computing is expected to still dominate by volume and economic footprint in 2045. Quantum and biocomputing breakthroughs are real, but they are narrow breakthroughs, not general-purpose ones.

Even optimistic roadmaps position them as co-processors accessed through classical orchestration layers, not standalone replacements. So in 2045, the data centre is still classical at its core, just with far more specialised compute islands embedded within it.

A glossary of terms is provided on page 19 of this report.



Drivers of change

Driver 1

Compute shifts outward, creating a ‘ladder-shaped’ data centre network

Models shift from ‘talking’ to ‘doing’, requiring ‘always on’ inference.

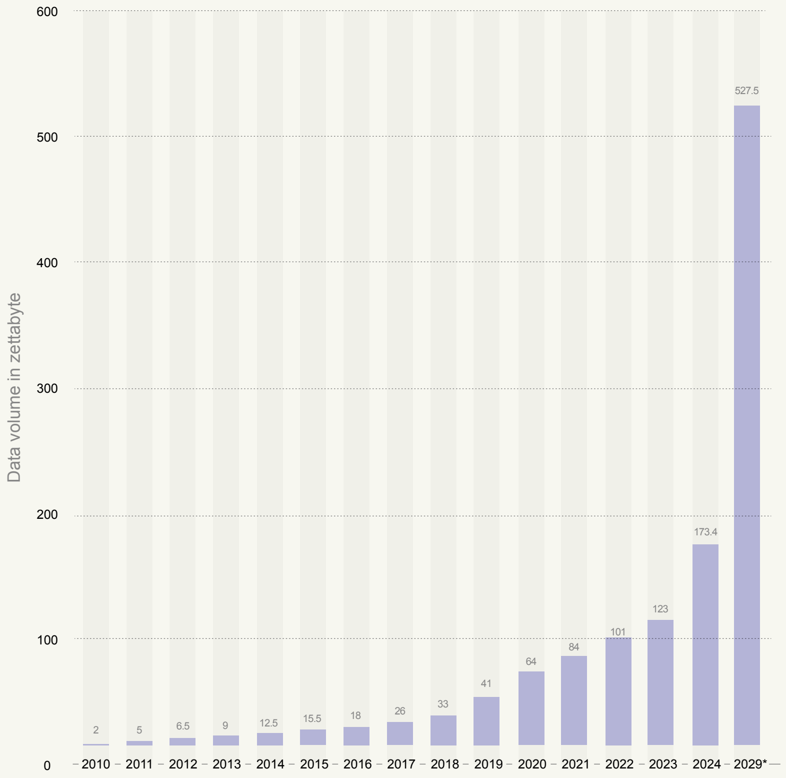
AI is moving from generating content to taking actions, shifting compute demand. Large Language Models (LLMs) are giving way to Large Action Models (LAMs) that observe, decide, and execute continuously, integrated into daily life, from personal devices and homes to workplaces and enterprise systems.

At the same time, data from sensors, wearables, IoT devices, autonomous systems, and industrial equipment is transforming the volume, speed, and immediacy of information flows. This drives demand for always-on inference, low latency and tighter links between AI and operational systems.

Because decisions increasingly need to happen in real time, compute is moving closer to where people and machines act, at the edge’, rather than only in distant hyperscalers. Training still depends on large, centralised clusters, but inference is becoming more distributed.

The shift towards decentralisation is reinforced by distributed usage of the models, which uses local compute, while training still relies on large, centralised clusters.

Enterprises adopting edge computing aren’t replacing centralised infrastructure; they’re creating ‘ladder-shaped’ networks that link local edge nodes for real-time tasks, regional hubs for aggregation, and core sites for heavy workloads like training and large-scale storage. This creates a more resilient system, with critical actions happen close to where they’re needed, while heavy lifting stays efficient at scale.



Volume of data or information created, captured, copied, and consumed worldwide from 2010 to 2029

The strategic focus shifts to the portfolio, not a single site.

Space as an experimental infrastructure tier

At the far end of the ladder, space is emerging as an experimental infrastructure tier for new compute architecture. Orbital data centres are not near-term substitute for terrestrial capacity but could offer the advantage of abundant solar energy, water-free heat rejection, and zero land use, though high launch costs[2] and regulatory uncertainty remain significant challenges.

Technologies from radar underpinning this driver: Self-learning systems, bio-integrated sensors, quantum sensors, distributed AI, living wearables, wearable ecosystems, wearable AI, edge AI, edge computing, 5G, 6G, 7G, space based data centres, large action models.

Drivers of change

Driver 2

Storage shifts from one-size-fits-all to optimised tires

Emerging archival media reshape storage economics

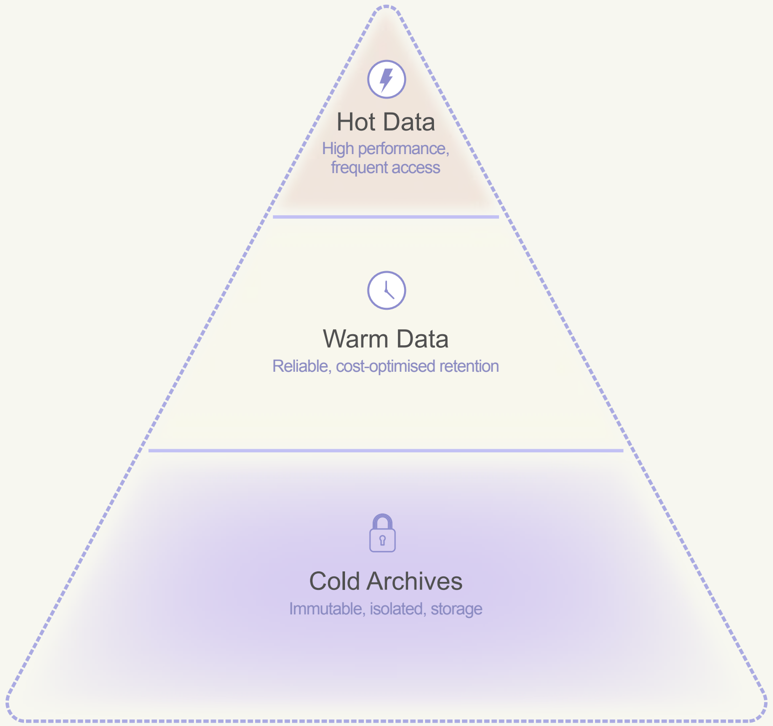
With data generation accelerating faster than storage supply,[3] the industry is pushing beyond silicon-based storage[4] toward next-generation archival technologies such as glass and DNA-based storage. These media could change storage economics by increasing density, reducing refresh requirements and lowering cooling demand. DNA, for example, could theoretically hold up to 1.8 exabytes per gram[5], greatly reducing the footprint of archival storage.

But these technologies are not yet ready for broad deployment DNA storage remains costly due to synthesis and sequencing expenses,[6] and major uncertainties persist around how biological storage can be integrated seamlessly into existing digital systems and compliance frameworks.

From one-size-fits-all storage to tiered, security- driven architectures

At the same time, storage architecture is becoming more tiered and purpose-built. Rather than relying on one system for all data, organisations are separating fast working sets, durable mid-term retention, and deep archival layers, each matched to different media and retrieval patterns.

This points to a future in which storage is optimised through deliberate use of multiple media types, each tailored to its operational purpose.



Think of it as organising your items in your home into three places: what you use every day stays within arm's reach, things you might need this month go in a closet, and things you almost never touch go in the attic. Data centres are doing the same with data, splitting it into 'hot' data that must be retrieved instantly, 'warm' data that should be reliable but not ultra-fast, and 'cold' archives that are cheaper to store but slower to access, using different storage technologies for each so the whole system becomes faster and cheaper. In practical terms, a multi-media storage stack can cut cost and energy for cold data while improving recoverability when it matters most.

Security pressures will reinforce this shift, as ransomware increasingly targets backup systems first, driving demand for immutable storage and isolated recovery environments with stronger physical separation.

Technologies from radar underpinning this driver: DNA based storage, glass-based storage, ultra-dense SSDs, solid state storage.

Drivers of change

Driver 3

Security and sovereignty drive design and investment decisions

Data sovereignty and cross-border legal exposure shape where data, backups, and keys can reside.

> Security is increasingly judged by recoverability, containment and sensitive data control, shaping data centre location, replication, and connectivity.

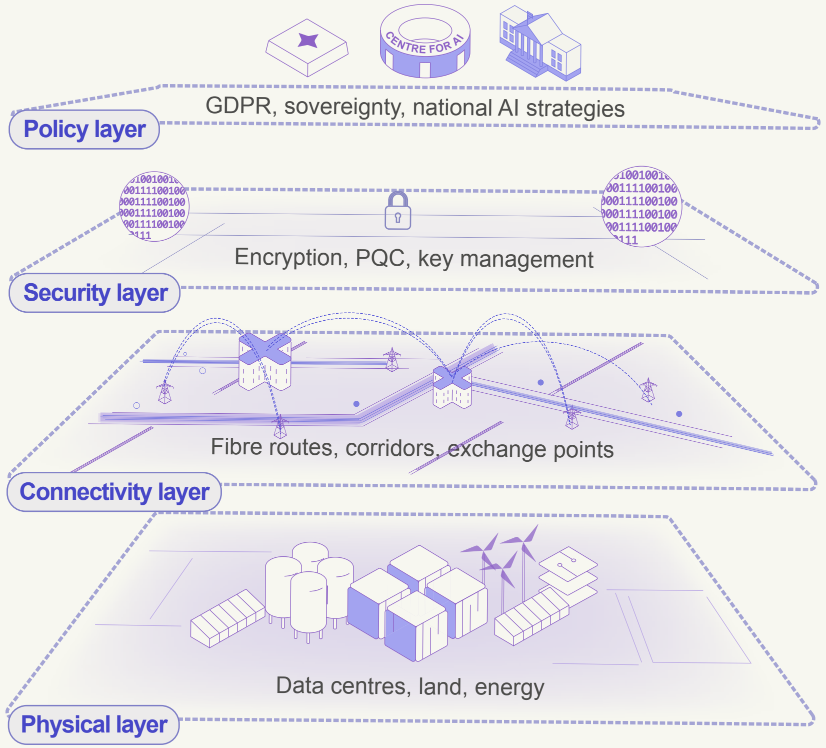
Data sovereignty, rules that require data to stay in a specific location or be controlled under its laws, is now a practical constraint, prioritised by 88% of IT decision-makers (Lenovo)[7]. Regulations such as the EU's GDPR and sector-specific rules restrict cross-border access, retention, and liability, pushing organisations to localise data, while adding complexity to storage strategies.

As data centres are increasingly treated as critical national infrastructure, high security requirements are being embedded in government AI strategies, raising the importance of early location and design decisions.

Quantum-safe security makes connectivity a facility and corridor issue

Quantum computing adds a further pressure. Encryption that protects data in transit today could become vulnerable within a decade. Creating a current "harvest now, decrypt later" risk. Governments and standards bodies are accelerating the shift to post-quantum cryptography for this reason.

This makes secure connectivity a core facility issue. Sites with diverse routes and strong carrier ecosystems gain value, while facilities need dedicated to key management zones, stronger traffic separation, and layouts that support crypto agility[8]. Security is becoming a more engineered and auditable design requirement, moving from policy intent to physical, testable design choices.



Technologies from radar underpinning this driver: Quantum computing, Quantum communication, hybrid classical-quantum, brain computer interface, space based data centres, 5G, 6G, 7G.

Drivers of change

Driver 4

Facilities and sites shift from homogenous to specialised

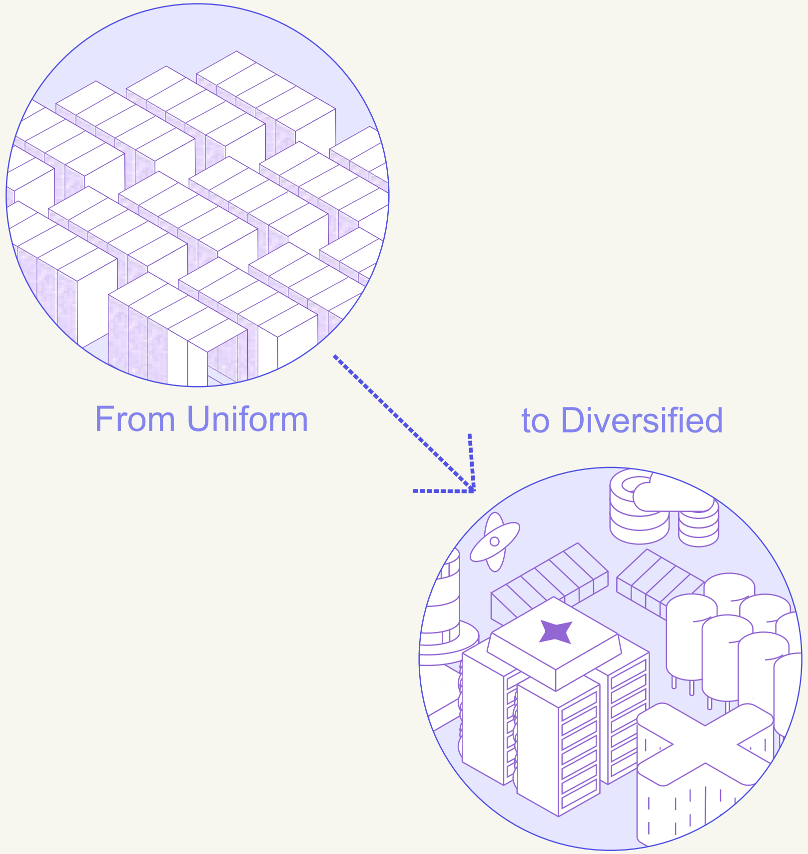
Until recently, most workloads improved by scaling similar hardware. Now workloads are diverging. AI training and inference, simulation or streaming analytics and methods that analyse data while limiting exposure of sensitive information often require different chips to achieve the best performance and energy efficiency.[9] [10]

As the limits of conventional silicon become more visible, new hardware approaches are maturing in parallel: photonic processors optimised for communication-heavy tasks, neuromorphic chips designed for low-power inference, hybrid classical-quantum systems for complex optimisation, and emerging biocomputing models exploring radically different physical behaviours. This is pushing data centres away from server halls into differentiated environments[11] designed to support radically different requirements for power, cooling, vibration, and environmental stability[12]. A photonic cluster may need precise temperature stability. Liquid-cooled GPU racks can exceed 100 kW per rack[13], while quantum require extreme thermal control.

> To stay adaptable, operators are moving toward modular, reconfigurable infrastructures so new generations can be integrated without redesigning the whole system.[14] [15]

Specialisation is also reshaping facility geography. Advanced compute types are concentrating into ‘capability hubs’ where the right ecosystem, supply chains, and expertise exist to deploy them reliably.[16]

If we get this right, heterogeneity becomes a strength: facilities that can host new workloads faster, with better performance per unit of power and space.



Technologies from radar underpinning this driver: Neuromorphic computing, stochastic computing, organoid intelligence, biocomputing, photonic computing, quantum computing, AI computing, large language models, large action models, personal large action models, symbolic reasoning integration.

External expert piece

From always-on to context-aware: rethinking digital infrastructure for data ecologies

Marina Otero
Architect, researcher, and advocate for new models of data storage Harvard GSD's 2022 Wheelwright Prize for research into the future of data storage

There is a significant shift in how major big tech companies and corporations are operating. As technology evolves rapidly and the lifespan of components like semiconductors and servers grows shorter, many companies are becoming less interested in building fixed data centres. Instead, they are favouring more flexible configurations – powered warehouses or power shells – that allow hardware to be placed, replaced, or relocated depending on changing conditions such as energy availability, tax incentives, or community responses. Extreme examples include building components on wheels that can be treated as temporary equipment and bypass regulation.

> This could create a highly fluid landscape in which territories become temporary hosts for digital capacity, often without long-term commitments.

When demand shifts or priorities change, facilities risk being left under-used or unsuitable for future needs. And as these infrastructures reshape power grids, water systems, and surrounding environments, once the machines move on, they could leave behind indebted territories that were organised around a future that never arrived.

Rethinking the 'always-on' infrastructure, for 'ecologies of data'

This evolution raises a more profound question about the logic governing digital infrastructure. Data centres have long been built on the premise of constant demand: 24/7 operation, continuous low latency, and immediate data accessibility.

However, our cities do not function uninterruptedly at maximum capacity; their rhythms fluctuate throughout the day and night. Only a limited range of digital activities genuinely always requires immediate response, such as medical devices or aviation control systems. Treating every workload as if it belonged to this category leads to over-provisioned, energy-intensive facilities that impose unnecessary environmental burdens. What if we challenged this approach and scrutinised the local impact of data centres?

I define as distinct ‘ecologies’ — hot, warm, cold, open, private, ephemeral, or eternal. Think of these ecologies as typologies reflecting the nature of the data. Each ecology implies different modes of storage and security levels, often more appropriate than current practices. Cold or long-term data, for example, does not require the immediacy of a Tier 4 facility (a very high resilience, high availability data centre standard) and can be housed in simpler environments optimised for slower access patterns. By aligning infrastructure design and placement with the nature and sensitivity of different data ecologies, organisations can move toward context-specific systems rather than defaulting to maximum intensity settings. In doing so, resources are used more consciously, favouring social justice, environmental protection, and effective public oversight.

The same logic applies to computational power and density, particularly with AI. Training does not need to run 24/7; instead, it can be scheduled with variable peaks, pausing at specific times and adjusting energy and computing use in alignment with eco-social values.

That leads to practical priorities: clarify which types of compute are truly critical and must be always on, design for workload flexibility and load shifting (moving computing to different times and places to match available power and reduce strain), integrate large scale storage with mediums that require less energy, and connect these choices to community governance. Load-shifting operations are being implemented in some cases, but not in ways that deliver meaningful benefits to society. We could design infrastructure that is responsive or adaptable to climate needs and territorial constraints, considering variables such as water stress or energy availability.

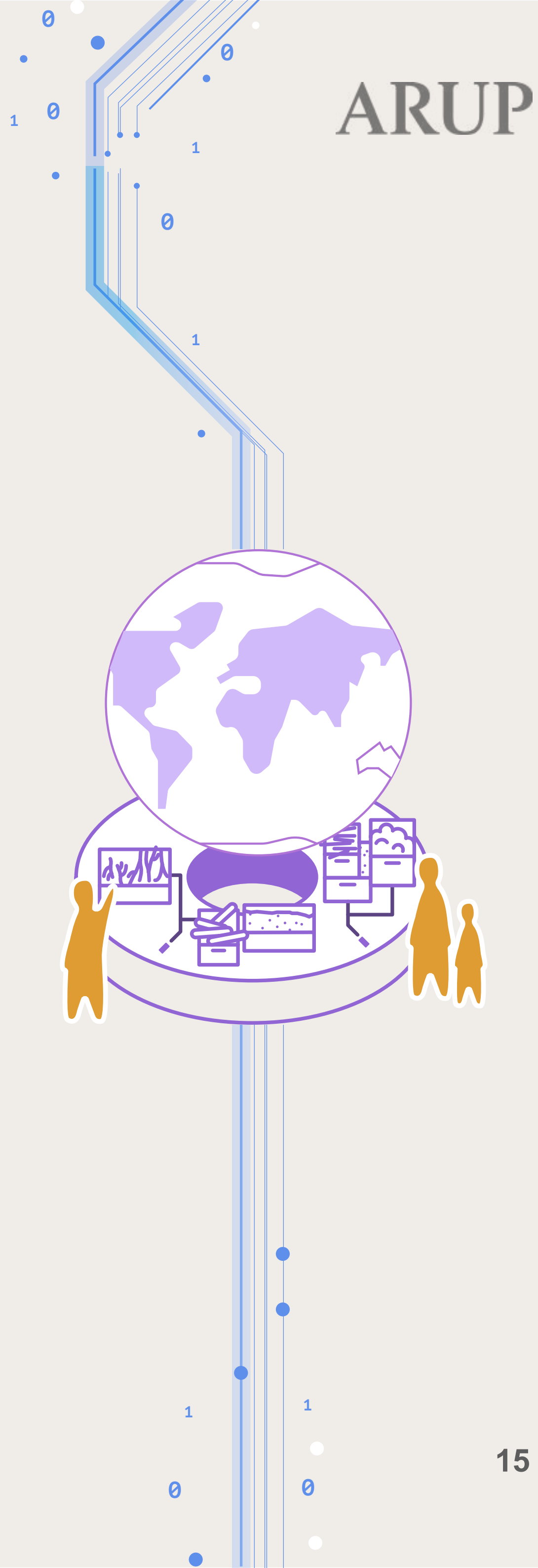
> The aim is to create a more flexible way of managing workload, so infrastructure can respond to real needs rather than imposing one uniform model onto all society and territories.

An uneven transition ahead

The transition to adaptive, responsible digital infrastructure will vary globally, with some prioritising extractive, high-intensity models and others favouring flexible, territorially attuned, long-term systems.

Tensions are evident in the United States, where rapid innovation can compromise security and increase social and environmental risks. Approaches differ globally: China follows a state-coordinated path, while Europe emphasises regulation and trustworthy AI to align development with ethical and social values. Regions focused on sovereignty, energy security and environmental governance may move faster in treating digital infrastructure as a public asset. Achieving this requires a shift towards systems designed for citizens and territories, resilient, distributed and responsive to local conditions.

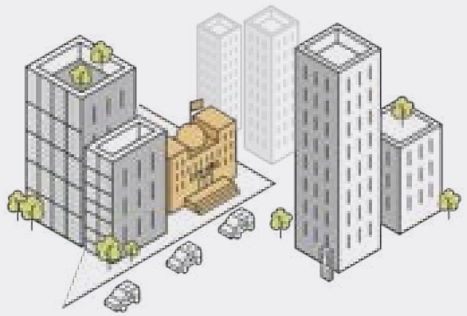
For data centre leaders, this means designing and operating portfolios around distinct data ecologies, so performance, uptime and intensity are matched to each workload and to the energy, water and governance realities of the places that host it.



Early signs of the future in the present

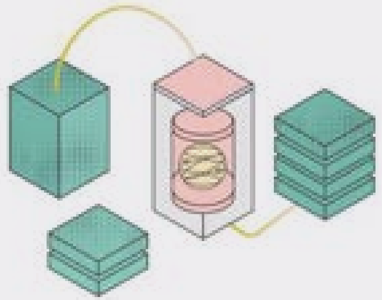
Case studies

Designing a city's neural network – Barcelona, Spain



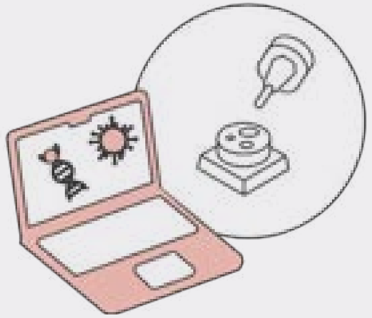
Barcelona has deployed a city-scale edge computing network through 18,000 sensors, allowing some data to be processed closer to where it is generated rather than being sent back to distant central facilities. This matters because it points to a more distributed spatial model for compute, where some capacity is embedded within existing urban infrastructure and buildings. [17] [18]

Social and digital infrastructure – Barcelona, Spain



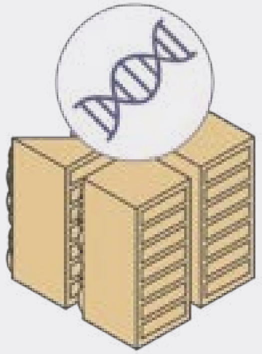
Qilimanjaro Quantum Tech's Barcelona Quantum Data Center is positioned as a multimodal hub that can host up to 10 quantum computers and provide access to both analog and digital quantum systems alongside classical compute for hybrid use cases.[19]

Labs become data centres – Vevey, Switzerland



FinalSpark's Neuroplatform makes 'wetware computing' practical by giving researchers remote access to living neural organoids for continuous electrophysiology experiments, with reported organoid lifetimes of even more than 100 days and datasets of more than 18 terabytes collected from over 1,000 organoids. These organoids are small, lab-grown clusters of human neurons derived from stem cells that mimic key structural and functional aspects of brain tissue. [20] [21]

Material scientists and data centre designers converge –



Catalog is developing DNA-based data storage as a way to store information at far greater density than conventional media, especially for long-term "cold" storage. The implication is that future storage may not always require warehouse-scale physical expansion, opening up the possibility of smaller-footprint facilities.[22] [23]

Recommendations

ACT NOW

The opportunity is not just to build more capacity, but to build it better: more adaptable, more efficient, and more aligned with the places that host it. These recommendations focus on decisions that can be taken immediately to reduce lock-in, manage near-term risk, and improve system resilience under current technological uncertainty, grid constraints and regulatory pressure.

Location strategy

Plan portfolios, not single sites. Define what sits at the edge, what sits in regional aggregation, and what stays in core sites (training, heavy storage), then use that map to guide every new build decision.

Make connectivity and security geographic variables. Treat fibre routes, DCI capacity and Quantum-safe 'secure corridors' as primary siting factors alongside energy, water and land.

Pick a small number of 'control points' to own. These are places where you can be hard to displace, for example key metro interconnection zones, critical industrial corridors, or a capability hub cluster with talent and supply chain depth.

Facility design & architecture

Lock in optionality in design and avoid irreversible choices. Build in heterogeneity from the start, with zoned densities, multiple cooling paths, upgradeable power and space for new interconnect. Set clear rules for when to build specialised hubs versus general-purpose halls.

Design for reconfiguration, not a single ideal layout. Use modular systems, clear upgrade paths and layouts that can be repurposed without major rebuilds.

Build crypto agility into physical and network design. Plan for upgrades to encryption and key handling without disrupting the facility, and create clear internal security zones.

Separate facility types by role. Define 'hot compute', 'warm regional' and 'cold archive', and avoid defaulting to one high-intensity model.

Prepare now by changing how teams work. As compute diversifies from high-density AI to quantum-adjacent and bio environments, shift from sequential handovers to integrated, cross-functional teams.

Operations, risk & governance

Build portfolio-level operations ('fleet management'). Put in place the tools and teams across many sites, not only within one facility (lifecycle management, incident response).

Start workload flexibility with customers. Identify workloads that can be scheduled or paused, then create commercial and technical mechanisms to shift load when grids are constrained or renewables are abundant.

Reduce lock-in where it matters most. Use multi-supplier strategies, modular procurement, and contract structures that keep options open (especially around accelerators, cooling and network fabrics).

Partnerships & innovation

Secure enabling partnerships early. Prioritise grid and energy partners, telcos and fibre providers, interconnection ecosystems, and local authorities.

Build capability hub ecosystems deliberately. If you want to lead in advanced compute, invest in the surrounding system (skills, vendors, maintenance, specialised cooling support, supply chain resilience).

Choose an 'orientation' per bet. Decide where you will act as a first mover (only where you can own a control point), where you will be a fast follower (with clear triggers), and where you keep a 'right to play' (small experiments with learning goals).

Create a small frontier pipeline with governance. Keep a structured way to run experiments (edge models, new storage media, security tech, frontier ideas) with clear success criteria and a path to scale.

Recommendations

MONITOR

The sector's next chapter is less about doing more of the same, and more about doing it with greater precision, matching compute to context, and performance to place. These recommendations focus on signals that will shape the next phase of investment, structural shifts in technology, governance and markets that may justify step-change decisions rather than incremental optimisation.

Location strategy

Where edge growth becomes structural versus opportunistic. Watch whether certain sectors (industry, mobility, healthcare, defence) start demanding local inference as a baseline, or whether edge stays limited to pockets.

Convergence toward hubs versus wider ladders. Track whether the market concentrates into fewer, denser interconnection hubs for control, or spreads into many smaller nodes due to sovereignty and resilience requirements.

Remote and extreme siting options for non-latency workloads. Monitor regulatory, insurance, and serviceability shifts that could make remote archival, harsh climates, or other unconventional geographies more viable.

Facility design & architecture

Quantum-related facility requirements. Monitor what becomes 'must-have' for quantum-adjacent operations (security zones, environmental stability, integration patterns), and only harden design specs when the ecosystem signals stability.

Deep archive media readiness (DNA, glass, other). Watch for real procurement signals: cost curves, standards, retrieval times, and compliance acceptance, then be ready to scale archival facility typologies.

Frontier compute substrates (neuromorphic, bio). Track power and cooling profiles, operational requirements, and governance constraints, treat this as 'right to play' experimentation until there is a clear path to mainstream use.

Operations, risk & governance

Whether flexibility becomes paid-for. Watch for customers, regulators, and grid operators starting to reward schedulable demand and demand response. That is the trigger to turn flexibility from pilots into a mainstream product line.

Shift from site-level permitting to system-level scrutiny. Monitor whether regions move toward cumulative impact assessment (power, water, air quality, noise) across multiple facilities and operators, and prepare operating models accordingly.

Security and sovereignty tightening. Track where requirements shift from 'best practice' to 'mandatory baseline', especially around post-quantum readiness and key management practices.

Partnerships & innovation

Edge market consolidation and shifting control points. Watch whether telcos, hyperscalers, colocation players, or industrial platforms capture the aggregation layer, that will change bargaining power and margins.

Space-based data centres and other extreme tiers. Treat these as learning and partnership optionality, track launch costs, regulation, and credible demonstrations, and avoid building core strategy on them until economics shift.

Standards and interoperability maturity. Monitor whether orchestration, interconnect standards, and tooling make heterogeneous fleets truly repeatable, that is the trigger for faster scaling of mixed environments.

Glossary

Generation & collection-

Creating or acquiring raw data and capturing it with essential context and metadata.

Bio-integrated sensors: Soft, body-adjacent sensors that continuously measure physiological or environmental signals such as chemicals, temperature, pressure, or motion.

Quantum sensors: Sensors that use quantum effects to measure physical quantities (e.g., time, gravity, magnetic fields) with extremely high sensitivity.

Quantum sensing: The use of controlled quantum states to achieve measurement precision beyond classical sensing approaches.

Distributed AI (federated learning): AI training or inference performed across multiple devices or locations, sharing model updates rather than centralising raw data.

Self-learning systems (machine learning): Systems that improve performance over time by learning from new data and feedback, rather than relying solely on predefined rules.

Wearable AI: AI models operating on wearables or nearby edge devices, enabling real-time, localised detection, prediction, and personalisation.

Living wearables: Wearables incorporating biological or biohybrid materials (e.g., living cells) to enhance sensing, adaptability, or user comfort.

Wearable ecosystems: Interconnected networks of wearables, companion devices, apps, and services that collectively capture and utilise personal data.

Storage-

The retention, protection, and management of data over time, including archiving and disposal.

DNA-based storage: Encoding digital data into synthetic DNA sequences for ultra-dense, long-term physical storage.

Glass-based storage: Storing data in durable quartz glass using laser inscription, designed for extremely long lifespans and minimal maintenance.

Ultra-dense SSDs: High-capacity solid-state drives using advanced NAND flash to deliver compact, power-efficient, high-performance storage.

Solid-state storage: Data storage using semiconductor memory with no moving parts, offering speed and durability advantages over magnetic media.

Processing & analytics: Cleaning, transforming, and analysing data to generate usable datasets and actionable insights.

Large Language Models (LLMs): Neural networks trained on large text datasets to generate and understand language for tasks like summarisation, translation, and Q&A.

Foundation models: Large-scale models trained on broad datasets that can be adapted to many downstream tasks.

AI computing: The compute-intensive execution of machine learning algorithms, typically using specialised hardware to extract insights from large datasets.

Large Action Models (LAMs): Models that interpret instructions and execute actions (e.g., tool use, workflows), enabling systems that perform tasks—not just generate text.

Personal Large Action Models (PLAMs): User-specific LAMs tailored to individual context, preferences, and permissions to carry out personalised tasks.

Symbolic reasoning integration (Neuro-symbolic AI): Combining neural networks with symbolic logic to enable both learning from data and structured reasoning.

Quantum computing: Computing based on quantum mechanics (e.g., superposition, entanglement) to solve certain problems more efficiently than classical systems.

Hybrid classical–quantum computing: Systems where classical processors coordinate with quantum hardware, allocating tasks to the most suitable compute type.

Biocomputing: Using biological materials (e.g., molecules or cells) to store or process information.

Organoid intelligence: A form of biocomputing using lab-grown brain cell structures (organoids) to study or enable learning and computation.

Neuromorphic computing: Brain-inspired computing using spiking neurons, typically designed for low-power, event-driven processing.

Photonic computing: Using light (photons) instead of electrons for computation or data transfer, improving speed and energy efficiency in certain workloads.

Stochastic computing: Performing computation using probabilistic bit streams, trading precision for simplicity, robustness, and energy efficiency.

Transmission & utilisation-

Distributing data to systems and users for sharing, and reuse under defined controls.

5G networks: Cellular networks enabling high bandwidth, low latency, and massive device connectivity.

6G networks: Next-generation networks (in development) aiming for even higher capacity, ultra-low latency, and integrated sensing and AI.

7G networks: A conceptual research term for post-6G systems; not yet formally defined or standardised.

Quantum communication: Communication using quantum states, typically for ultra-secure data exchange (e.g., quantum key distribution).

Brain–computer interfaces (BCIs): Technologies that translate brain signals into commands for external devices, enabling direct human–machine interaction.

Space-based data centres: Computing and storage infrastructure deployed in orbit to process space-generated data and reduce Earth-based transmission demands.

Embodied AI / Robotics: AI systems embedded in physical machines that perceive, decide, and act in the real world, enabled by robotic hardware.

References

1 Noffsinger, J., Patel, M. and Sachdeva, P. (2025) The cost of compute: A $7 trillion race to scale data centers. McKinsey & Company.
2 AlphaTarget (2026) 'Data centres in space', AlphaTarget Blog.
3 International Data Corporation (2021) 'Data creation and replication will grow at a faster rate than installed storage capacity, according to the IDC Global DataSphere and StorageSphere forecasts', Business Wire.
4 Monroe, J. (2022) The escalating challenge of preserving enterprise data. Furthur Market Research.
5 Pardhi, K., Gaikwad, N., Kamble, M. and Baheti, J. (2025) 'Unlocking the future: A comprehensive review of DNA as a data storage medium', International Journal of Pharmaceutical Sciences, 3(6), pp. 2457–2463.
6 Raza, M.H., Desai, S., Aravamudhan, S. and Zadegan, R. (2023) 'An outlook on the current challenges and opportunities in DNA data storage', Biotechnology Advances, 66, p. 108155.
7 Craske, B. (2026) 'How will data sovereignty impact data centre design?', Data Centre Magazine.
8 National Security Agency (2024) The Commercial National Security Algorithm Suite 2.0 and quantum computing FAQ (Version 2.1).
9 Hennessy, J.L. and Patterson, D.A. (2019) 'A new golden age for computer architecture', Communications of the ACM, 62(2), pp. 48–60.
10 Intel (2022) Compare benefits of CPUs, GPUs, and FPGAs for different oneAPI compute workloads.
11 Uptime Institute (2026) Uptime Institute announces five data center predictions report for 2026.
12 Parsons, R. et al. (2025) 'Highly uniform thermally undercut silicon photonic devices in a 300 mm CMOS foundry process', Scientific Reports, 15, p. 29906.
13 Goldwasser, I., Hsu, M. and Petty, H. (2025) 'Building the modular foundation for AI factories with NVIDIA MGX', NVIDIA Developer Blog.
14 Murphy, M. (2024) 'How to test a quantum computer chip', IBM Research Blog.
15 NVIDIA (no date) MGX platform for modular server design.
16 Miller, B. (2024) 'CXL 3.0 and the future of AI data centers', Keysight Blog.
17 Law, M. (2025) 'Why the edge revolution is reshaping data centre strategy', Data Centre Magazine.
18 Sentilo (no date) What is.
19 Qilimanjaro Quantum Tech (2025) Qilimanjaro launches world's first multimodal quantum data center, paving the way for the computing of the future.
20 Jordan, F.D., Kutter, M., Comby, J.-M., Brozzi, F. and Kurtys, E. (2024) 'Open and remotely accessible neuroplatform for research in wetware computing', Frontiers in Artificial Intelligence, 7, Article 1376042.
21 FinalSpark (no date) Home.
22 Biomemory (no date) Pioneering DNA data storage appliances for data centers.
23 CATALOG (no date) The next leap in computation: DNA.

Read more perspective pieces from Arup

Quantum computing in the built environment - Arup

Accelerating the next generation of modular data center design - Arup

Can we design data centres that are fit for the future? - Arup

Get in touch with our experts to discuss further.

Arup University is our corporate university function that prepares our clients, collaborators, and global membership for the future. We are committed to raising technical standards across the natural and built environment sectors.

Contacts
Arup University Foresight
foresight@arup.com

Arup Digital Services
https://www.arup.com/services/digital/

Arup and Data Centres
https://www.arup.com/markets/data-centres/

Data Centre Futures team
Bree Trevena
Charlie Warwick
Olivier Woeffray
Andrea Pavon
Arlind Neziri
Julien Clin
Lauren Smith
Belen Palao
Felipe Sarachaga
Sagar Bhat

Arup expert contributor
Jimmy Chan

External expert contributor
Marina Otero

ARUP