Everything You Wanted To Know About Data Centres (But Were Too Afraid To Ask)

Petr Rojicek
Chief Investment Officer
Amundi Alpha Associates

Data centres are rarely out of the headlines. From the scale of demand to the challenges associated with construction, the buildings which power our digital world are at the forefront of the conversation regarding the future of our economies, and even our societies.

Is AI a bubble? Or the second industrial revolution? And what really is a data centre anyway?

Read on to find out everything you ever wanted to know about data centres, but were too afraid to ask.

Rapid digital growth

As the internet has matured and digital services have expanded, the demand for data centres - a building which houses a company's IT infrastructure and stores its data - has boomed.

In 2010, the world consumed just 2 zettabytes (the equivalent of roughly 2 trillion megabytes) of data. In 2024, that had grown to 173 zettabytes, and by 2029 consumption is forecast to more than triple to 528 zettabytes.[1] The rapid build-out of data centres has been, and will continue to be necessary to service this exponential growth.

Is this all down to AI?

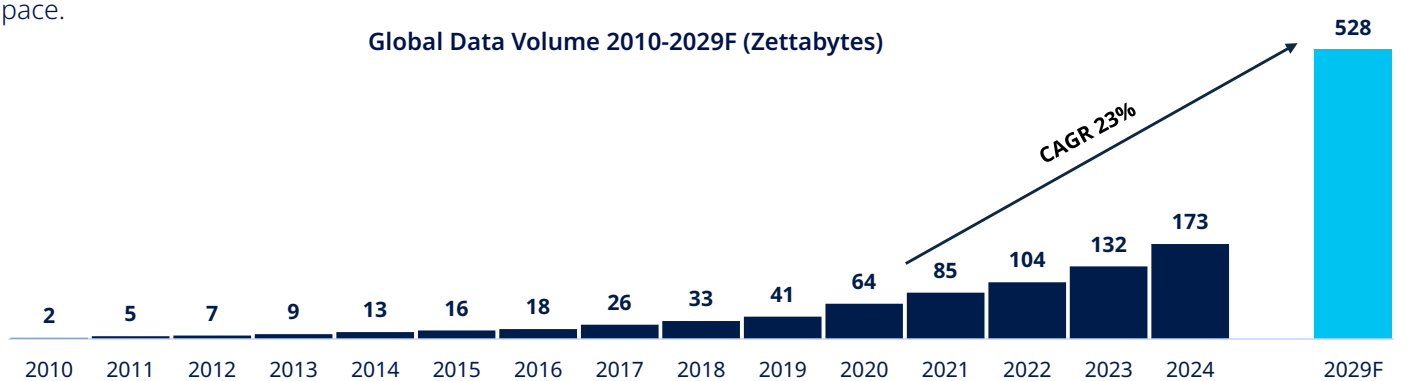
To date, no. The sector has been in a three-decade cycle of expansion driven by smartphone adoption, video streaming, e-commerce, digital payments, mobile apps and companies moving to the cloud. Until around 2022, cloud migration was the single biggest structural driver behind the growth in data centres. Now, however, demand for new data centres is accelerating dramatically because of AI.

Hyperscalers and colocation

For many years data centres were simply a server room in a company's own building, but as the cloud migration shift got underway, companies began moving their data to remote cloud service providers such as Amazon, Microsoft and Google. These tech giants, known as hyperscalers, manage data in two ways: they build their own campuses hosting hundreds of thousands of servers and also lease space and power from specialised data centre operators. These facilities, known as colocation data centres, are owned by specialised third-party providers with the space either leased by one hyperscaler, or shared among multiple enterprise customers.

While very large organisations such as governments, banks and telecommunications companies still own and run their own enterprise data centres, many small and medium-sized companies have replaced owned facilities with a colocation space.

Global Data Volume 2010-2029F (Zettabytes)



Source: : IDC Global Datasphere Forecast

AI is changing the data centre story

Until recently, the data centre industry, fuelled by demand for cloud services, was largely optimising for space and storage. The constraint was access to real estate, not computational power, energy availability and grid connectivity.

But a piece of silicon just a few centimetres in width has changed all that.

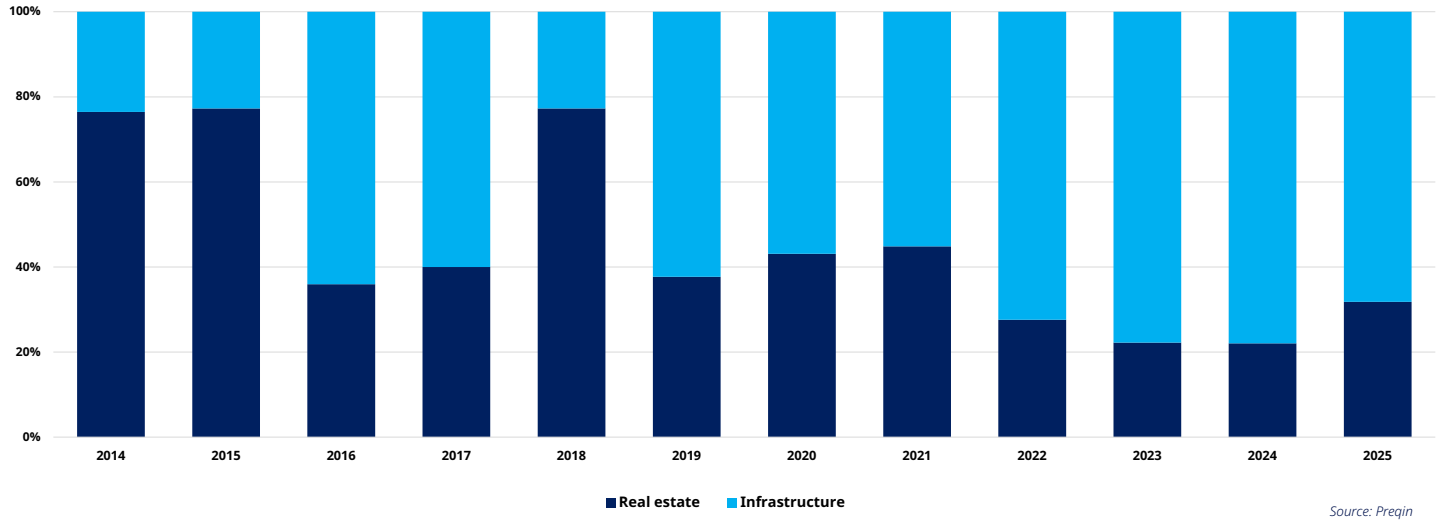
A Graphics Processing Unit (GPU) is the processing chip which underpins the training of AI models. And GPUs, which process data at very high speeds and in parallel with many other data requests, are replacing slower Central Processing Units (CPUs) previously used to power cloud computing. The capabilities of GPUs make them ideally suited to training big AI programs known as Large Language Models (LLMs), and to fulfil the end tasks (such as chatbot requests) performed by LLMs after they have been trained.

But GPUs come with significant challenges: they are expensive, power hungry, and get very hot.

These challenges compel a more rounded investment approach in which the data centre is treated not as a stand-alone asset but as one piece of the puzzle within a broader infrastructure ecosystem. The data centre story then, once a bricks and mortar play, increasingly resembles an infrastructure opportunity encompassing fiber connectivity, substations, transmission networks, renewable energy, gas generation, cooling, and long-term electricity pricing.

And with all these considerations at the forefront of the next phase of the AI data centre build-out, infrastructure players are moving in.

Proportion of data centre deals by asset class – infrastructure and real estate[2]



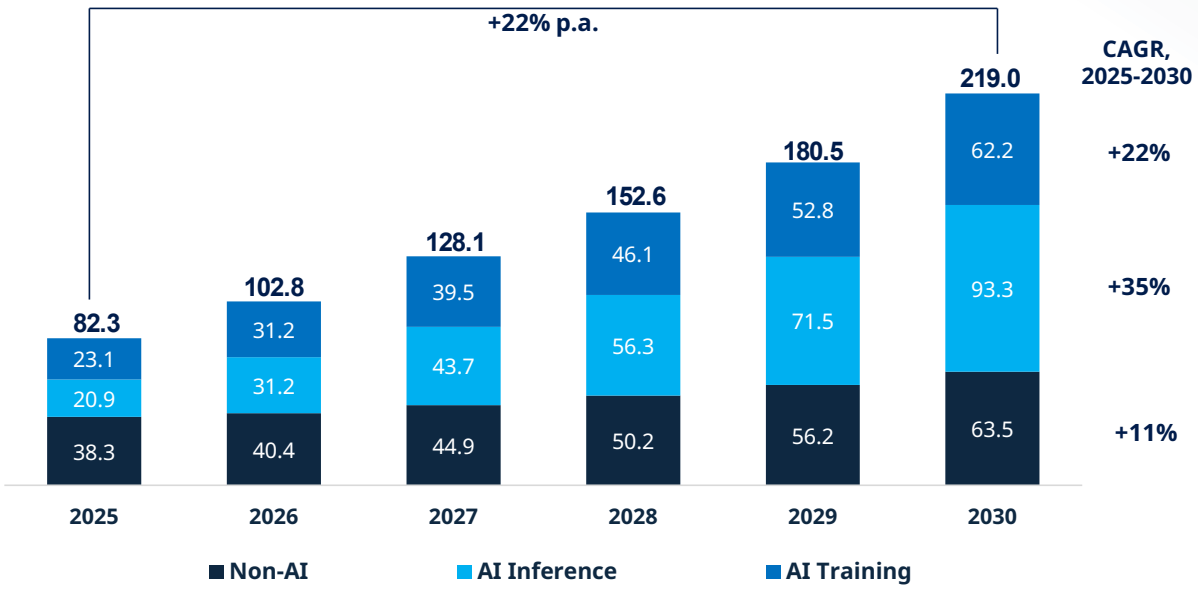
Source: Preqin

Training vs inference

AI expansion to date has largely been focused on data centres where LLMs are trained to perform tasks. But once the AI model is trained, it needs to be put to use generating real-time outputs for users, commonly via chatbots or image generators. These tasks require a different type of data centre known as an inference data centre. Inference data centre construction is forecast to grow 79% annually to 2030 and by the end of the decade could represent 40% of total data centre demand.[3]

A third type of AI data centre centred around GPU processing is an edge data centre. Smaller than the campus of a hyperscaler, an edge data centre is located close to the end customer to support applications which require high latency such as streaming services, gaming or chat bots.

Global data centre demand by workload, 2025-2030 (gigawatts)



Source: McKinsey Data Center Demand Model

Investment need

While inference data centres are forecast to be the fastest-growing workload type, training centres, edge centres and even traditional cloud centres are all on an upward trajectory, and growth across all sub-sectors will necessitate USD 1.6t of investment by 2030.

Given the shift in the sector towards infrastructure-like projects, specialist managers are well-placed to contribute to this investment need, seize the rewards and overcome the challenges.

And the key challenge is energy.

Considerations and challenges

Because GPU-powered AI data centres have much greater power requirements than CPU data centres (some hyperscale projects are as large as entire cities), the most valuable sites are not where land is cheapest or buildings highly rated, but where grid power is available.

Additionally, because high-density GPUs generate extreme heat and require advanced cooling systems, securing access to water is critical. These challenges, grounded in the availability of resources essential to the life of the surrounding community, raise additional social and geopolitical risks.

In the US for example, community opposition to data centres because of resource and energy concerns has seen projects worth USD 64b blocked or delayed since 2023, with at least 25 cancelled outright in 2025 alone.[4]

Substantial rewards

The rewards on offer for those who manage to navigate the challenges are substantial. Difficult-to-replicate infrastructure and development pipelines, long-lived resilient assets, and structurally rising demand supported by supply constraints. Projects also benefit from long-term contracted revenues, inflation-linked contracts, pass-through energy costs, and high switching barriers which create strong customer retention. And these customers tend to be creditworthy, high-value tech enterprises.

Recent data suggest that these factors have converged to make data centres an exceptionally well-performing asset class: between 2011 and 2025, annualised returns have averaged 23.8%.[5]

Going it alone

For an individual investor to navigate this landscape and claim these rewards alone, however, is a tall order. Such a complex sector requires serious expertise, deep networks and the ability to construct a diversified portfolio to hedge against material risks. For example, different managers pursue markedly different strategies: Some focus on hyperscale data centres with long-term contracts and high-credit-quality tenants, albeit often at elevated valuation levels. Others prefer colocation models with a broader customer base, but shorter contract durations. Some managers focus on niche segments such as edge data centres or carrier hotels.

Approaches also differ significantly on the execution side, ranging from greenfield development strategies to brownfield approaches involving M&A transactions or corporate carve-outs. This highlights that, even within a structurally attractive market, risk-return profiles can vary considerably.

Multi-manager territory

The complex, multifaceted nature of data centre investing, and the whole-ecosystem perspective it demands, is ideally suited to a multi-manager approach.

Through the careful selection of underlying GPs with different sectoral, project and geographic focuses, a multi-manager can achieve diversification across data centre types, from hyperscale, to colocation to edge. Diversifying across managers with a range of strategies can protect against overheating in any one subsector and provide downside protection on the energy & connectivity infrastructure supplying projects.

Given the similarities between data centres and traditional infrastructure, a multi-manager can also draw on a broad network of dedicated infrastructure managers experienced in navigating the intricacies of the sector, from project viability, to energy contracts and pricing. These skills and experiences are especially crucial for evaluating and managing higher-risk greenfield projects.

A broad network of managers with local knowledge is also especially critical in infrastructure where regulatory landscapes are distinct across geographies and socio-political considerations can be an important factor.

Not all types of GPs with data centre exposure will be the right fit for all investors. A multi-manager can therefore help an investor take an informed view on constructing a portfolio of funds, and if appropriate can also offer an investor direct exposure to underlying assets via co-investments.

Diversification, diversification, diversification

The data centre story is complicated and layered: a new growth phase is emerging, but the old growth phase hasn't ended. All types of assets can offer exceptional gains, but complexity abounds and the risks are real.

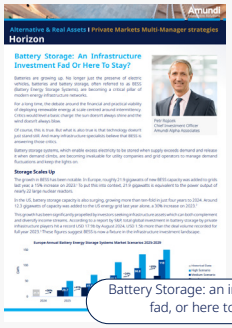
In such an environment, fraught with challenge but abundant with opportunity, the importance of experience, access and above all, diversification, cannot be overstated. The rewards for investing in our digital future are clear. And to capture these rewards, true diversification is the key.

About Amundi Alpha Associates

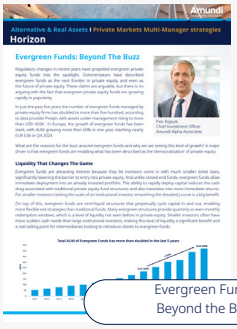
Amundi Alpha Associates is the private markets multi-manager business line of Amundi Alternative & Real Assets, a trusted European partner with over 40 years of investment experience. Amundi Alpha Associates manages over EUR 23b in AuM across private equity, private debt, and infrastructure funds-of-funds and segregated accounts, primarily on behalf of a global, institutional client base. The business line employs over 80 professionals across its offices in Zurich, Paris, and Barcelona.

Related Insights

Unlocking private markets: The rise of evergreen funds



Battery Storage: an infrastructure fad, or here to stay?



Evergreen Funds: Beyond the Buzz

Important Information

1. IDC Global Datasphere Forecast
2. Preqin Data Centers Report Q1 2026
3. https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/the-next-big-shifts-in-ai-workloads-and-hyperscaler-strategies)
4. https://www.datacenterwatch.org/report
5. https://www.msci.com/research-and-insights/blog-post/the-rise-and-concentration-risk-of-data-centers-in-private-markets)

Amundi's private market multi-management business, based in Paris, was launched in 1998 under the legal entity Amundi Private Equity Funds (PEF). Zurich-based Alpha Associates AG was founded in 2004 as a spin-off from Swiss Life. These two entities have been working together since 2024 to create the Amundi Alpha Associates platform.

Date of first use: 26/05/2026

Doc ID no. 5517243

Document issued by Amundi Alpha Associates AG, a "société anonyme" (Swiss public limited company) with a share capital of 250,000 Swiss francs, authorised by the FINMA (Swiss Financial Market Supervisory Authority) as a collective wealth manager. Registered office: Bahnhofstrasse 13, 8001, Zürich, Switzerland. Zefix (business registration no.): CHE-110.534.686, Commercial Register of the canton of Zürich.