# From niche to necessity

## The state of the data center industry in 2026

**Kristina Metzger**
CBRE

**Ben Wobschall**
CBRE

Explosive AI and cloud demand have pushed data centers from niche assets to essential infrastructure — driving record absorption, sub-1.5% vacancy, and sweeping changes in how projects are powered, built, leased, and capitalized. Institutional capital has entered the sector, but bottlenecks are impeding supply as execution is harder than ever.

▶ Today, data centers are foundational to the global economy

iStock.com/adventtr

**KEY TAKEAWAYS**

**FOUNDATIONAL INFRASTRUCTURE**
AI, cloud, and enterprise digitization have pushed data centers to the forefront, with the sector operating at sub-1.5% vacancy and experiencing sustained double-digit rent growth.

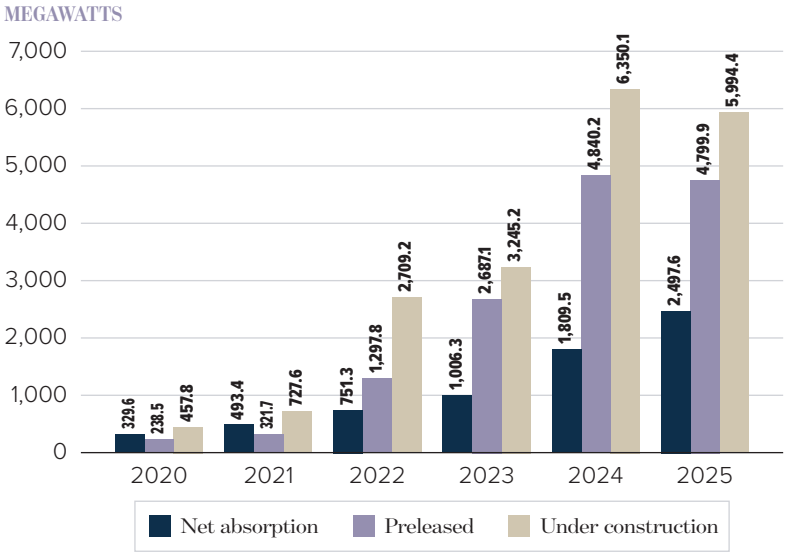
**LEASING STRUCTURES ARE SHIFTING**
Net-lease frameworks now dominate large-scale developments, reducing operator exposure, and creating bond-like income streams attractive to long-duration capital. Turnkey, powered shell, and hybrid delivery models allow tenants to balance customization, speed, and capital efficiency.

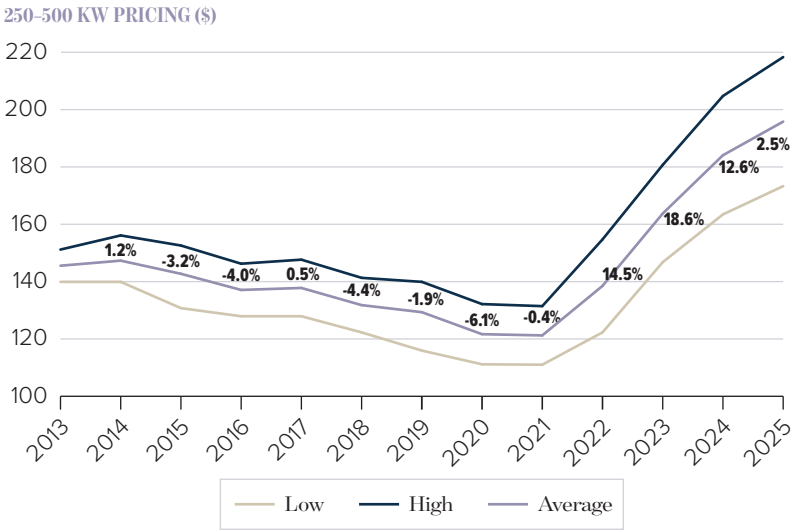
**INSTITUTIONALIZED CAPITAL MARKETS**
Pensions, sovereign wealth funds, infrastructure managers, private equity, and insurance capital are major players. Forward sales, structured JVs, long-tenor construction financing, and record data-center securitization volumes are enabling unprecedented scale.

1

## Demand for data centers is explosive

In primary markets, inventory has expanded by more than 25% annually for the past five years. Net absorption grew by 38% in 2025. Limited supply and strong demand growth has led to rents growing over 10% annually for the past 5 years.




Source: CBRE Research, CBRE Data Center Solutions, 2H2025.

A decade ago, data centers were a specialized niche asset class, necessary to a handful of cloud pioneers, content platforms, and forward-leaning enterprises. Today, they are foundational to the global economy. Every click, stream, swipe, transaction, and inference runs through a rapidly expanding network of digital infrastructure. And with artificial intelligence shifting from experimental to essential, the physical footprint of the internet — power, fiber, buildings — has become a central constraint, and a central investment theme.

The growth numbers tell the story. According to CBRE research,[1] just five years ago, the top eight primary data center markets collectively offered about 2,940 megawatts (MW) of leased operational critical capacity. Today, that inventory has more than tripled to over 9,400 MW, expanding by more than 25% annually for the past five years — a compounding trajectory that few sectors can match (see 1). Coupled with preleasing and development outside of the historical primary markets, the growth story is even more impressive. According to Datacenterhawk, US absorption jumped to 16 GW in 2025, up from 7 GW in 2024.

Yet even that torrid growth hasn't satiated demand. By the end of 2025, vacancy in those primary markets averaged under 1.5%, an almost unheard-of level in commercial real estate. Unsurprisingly, rents have responded: five-year average rent growth exceeds 10% annually.[2] While headline numbers fluctuate by market, the trend is unmistakable — supply cannot keep up with the digital world's voracious appetite for compute.

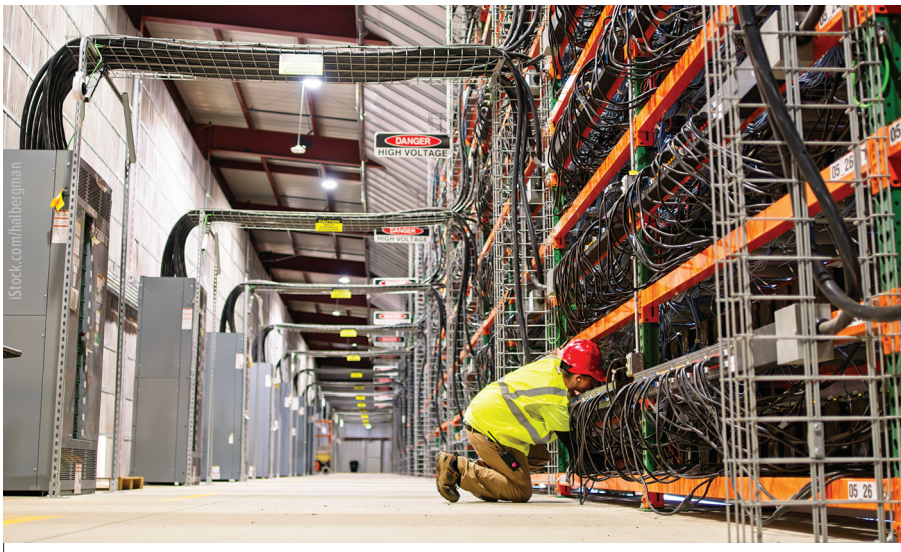
This article explores what is driving that appetite for capacity, the bottlenecks impeding supply, how leasing and development structures have evolved, why densification is redefining project sizing, and how capital markets have institutionalized and adapted to fund this growth — from powered shells to net-lease constructs to

megaproject 'AI factories.' In short: the industry is maturing rapidly, but with power at the center, execution is harder — and more valuable — than ever.

## Demand drivers

The foundation for this demand supercycle began in the 1990s with enterprise adoption of digital processes and channels. As companies embraced e-commerce, online services, remote and hybrid work, data analytics, cybersecurity, and always-on global operations, they recognized that data storage, processing, and connectivity were mission-critical. But these capabilities also required large amounts of capital and expertise to build and maintain in-house at scale. Early enterprise data centers proliferated, but the economics pushed a turn toward outsourcing.

With the 2006 releases of Amazon's Elastic Cloud Compute and Google Docs, cloud services answered that call. Cloud platforms offered flexible compute capacity, speed-to-deploy, global availability zones, and world-class resiliency. Outsourcing infrastructure allowed enterprises to redeploy capital to core businesses while offloading operational complexity.




Data center in a remote location in Stutsman County, North Dakota

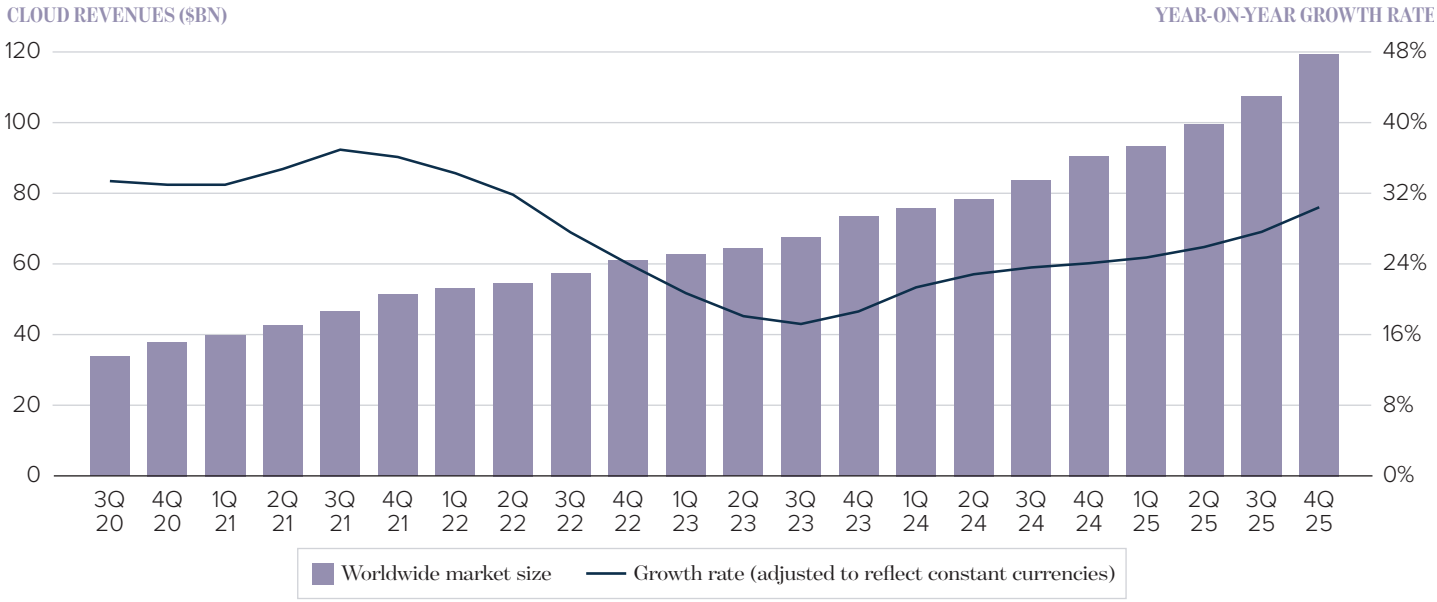
This shift introduced the modern colocation and wholesale data center market structure: developers and operators build and operate the facilities; cloud and enterprise tenants contract for capacity.

Cloud hasn't lost momentum (see 2). Market leaders such as Amazon Web Services, Microsoft Azure, and Google Cloud continue to report consistent cloud

2

### Cloud revenue growth is accelerating again

Global cloud revenue is growing at an accelerating rate, totaling $419bn in 2025, up 30% year over year, as breakthroughs in computing architecture and model performance brought artificial intelligence into the mainstream.




Source: Synergy Research Group.

demand. This persistent baseline growth from cloud adoption remains a primary pillar of sector demand.

In 2023, a step-change occurred. Breakthroughs in computing architecture and model performance brought artificial intelligence into the mainstream. Training large-scale AI models requires clusters of high-performance processors (notably GPUs) with dense power and cooling. The capacity needed to train foundation models is enormous, and rising, creating a wave of hyperscale campus development focused on power availability and efficient cooling.

But AI demand isn't only about training. As models move from labs to production, value shifts to inference — running trained models for end users. Inference needs to be close to the demand — near users, devices, and enterprise networks — to minimize latency and accelerate monetization. That's the long-anticipated 'edge': distributed compute located across regional and metro markets to deliver lightning-fast response times for AI-augmented software, content delivery, and industrial/IoT applications. AI's two-headed demand — training at massive scale and inference at the edge — is now a durable, multi-year force reshaping the national (and increasingly global) map of capacity deployment.

## The supply constraint

Today's capacity bottleneck is fundamentally about electricity. In the early days of data center expansion, the problem was often transmission. Those challenges have not disappeared, but the core issue has shifted to generation.

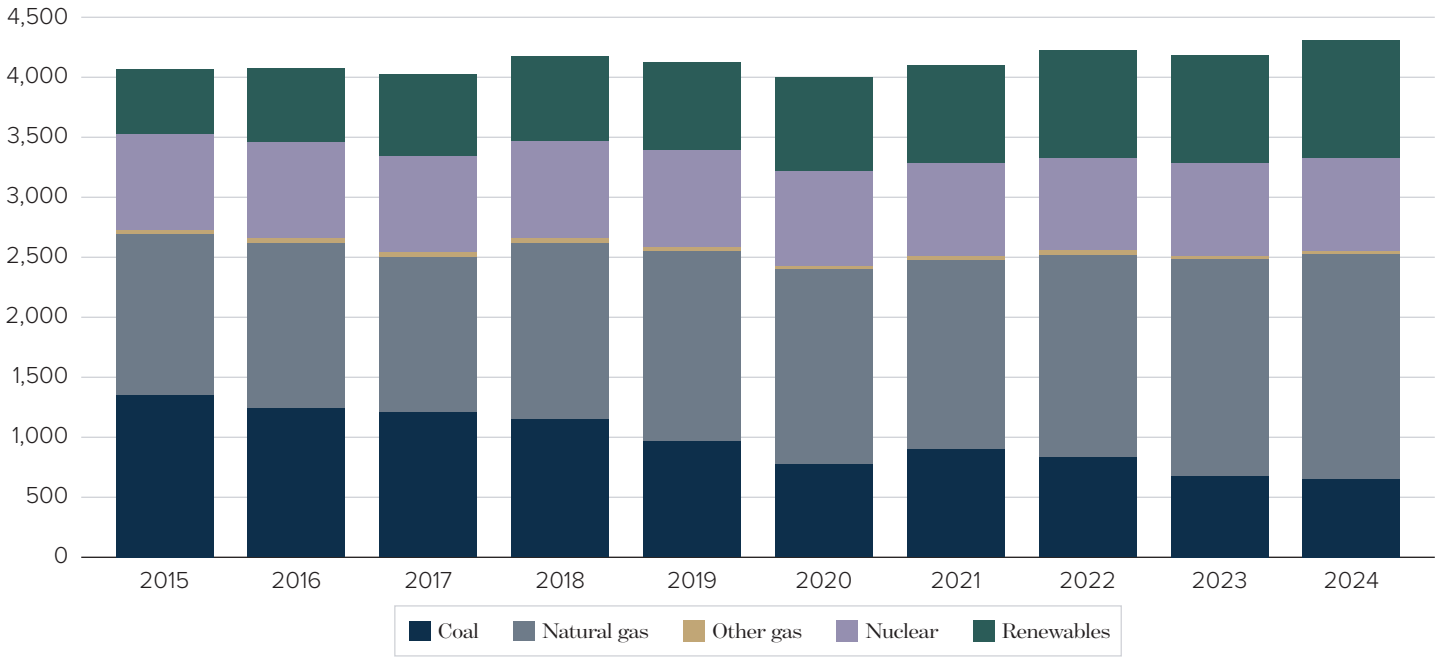
Despite significant investment in renewable generation, total electricity generation in the US has grown a meager 0.6% annually over the past decade, while base-load capacity from fossil fuels and nuclear energy has declined (see 3). Intermittent renewable resources do not perfectly align with 24/7 compute demand

3

### Electric generation growth has been anemic

Despite increasing demand for electrification and a push for renewable energy, total electricity generation in the US has grown a meager 0.6% annually for the past 10 years. Steady baseload generation is decreasing as variable renewable energy makes up a larger portion of total generation.

US NET ENERGY GENERATION BY SOURCE (TWH)




Source: US Energy Information Administration.

without massive storage (still costly) or complementary dispatchable capacity.

Promising long-term solutions like small modular reactors (SMRs) are advancing but remain several years away from commercial deployment. In the interim, the energy industry increasingly views new natural gas generation as a necessary bridge. Developers and hyperscalers are working with utilities, independent power producers (IPPs), and grid operators to align multiyear power roadmaps with phased campus build-outs for both utility capacity and, increasingly, behind-the-meter, on-site generation.

At the same time, tariffs and ongoing trade frictions have compounded an already strained supply chain for large electrical components — transformers, switchgear, UPS systems, and generator sets. Lead times have extended, costs have escalated, and timing uncertainty has risen. For projects dependent on synchronized equipment deliveries, this creates additional schedule risk that must be managed in contracts and contingency budgets.

Power is not the only front-end constraint. Entitlements and public perception are increasingly divisive. Communities have grown more vocal about a project's water and power footprint, and more skeptical about financial incentives. For sites not entitled 'by right' — lacking zoning and infrastructure approvals aligned with data center use — timelines stretch and risk premiums rise. Developers who arrive with utility readiness, environmental diligence, and proactive community engagement gain a significant advantage. In many metros, entitlement risk is now comparable to construction or procurement risk in determining a project's bankability.

These pressures have resulted in tight markets. Even as new construction surges, absorption outruns supply, which naturally pressures pricing and elevates competition for every megawatt.

## Engineering for density

If there is one physical trend that defines the last several years, it's densification. Power per square foot in critical space and power per rack has climbed sharply due to higher-powered servers. The electrical backbone — utility feeds, substations, switchgear, distribution pathways — has scaled to match. Cooling systems are more robust, more targeted, and

## The new lease landscape

Over the past five years, average rent growth has exceeded 10% annually. Behind that headline is a repricing of risk and a reallocation of responsibilities within leases. Put simply, scarcity and operational complexity have increased the value of delivered, dependable power, and reduced the willingness of owners to accept unbounded operating exposure. Simultaneously, tenants increasingly value control and flexibility to tailor electrical and mechanical systems to their workloads. These forces have converged to reshape leasing.

Historically, data center agreements — often structured as master services agreements (MSAs) and service orders — reflected primarily a gross reimbursement model. Tenants paid for direct electrical usage and a pro rata share of building and mechanical utilities (frequently subject to caps), while the operator bore substantial operational obligations and risk. Service level agreements (SLAs) within the contracts typically guaranteed uptime, environmental parameters (temperature/humidity), and security standards. Material SLA breaches could trigger credits or penalties, reinforcing the operator's 'service provider' role and corresponding risk profile. This structure remains prevalent for multi-tenant colocation leasing given the shared personnel and in some cases infrastructure.

For large-scale developments, most typically single tenant in profile, the market has largely shifted to net-leased structures that pass-through costs more comprehensively and curtail operating risk for ownership. Many leases exclude SLAs altogether, or narrow them considerably, further reducing the landlord's exposure to operating risk. Term lengths, renewal options, expense allocation, and repair/replacement obligations increasingly resemble those found in institutional real estate and core infrastructure assets. The net effect is more predictable 'bond-like' cash flows, which has broadened the investor universe to include net-lease and insurance capital that prioritize predictable cash flow.

### Buildout flexibility embedded in leases

Another key evolution is how the facility is delivered and who funds what:

- Turnkey delivery remains common. The developer builds the shell and all critical infrastructure (switchgear, UPS, generators, cooling plants, power distribution). Within turnkey deals, negotiated terms determine who is responsible for repair/replacement during the term and for conditions at lease expiration. These nuances are material to underwriting and valuation.
- Powered shells continue to gain in popularity. The owner delivers the shell with utility power and backbone infrastructure up to agreed demarcation points, while the tenant performs the critical fit out (UPS, generators, cooling, power distribution). This structure reduces owner capex and delivery risk, and enables tenants to tailor the environment for their workload. Tenants typically own the critical infrastructure they install, enabling accelerated depreciation and tax benefits during the lease term.
- Hybrid-powered shells blend the two. Akin to a tenant improvement allowance, the owner funds part of the tenant's buildout. Hybrids allow investors to scale capital deployment into attractive projects without assuming full operating complexity, while tenants preserve flexibility and direct more of their balance sheet toward their core business.

# Once dominated by public REITs and a select group of private operators, the data center sector has now fully institutionalized.

increasingly liquid-based to handle thermal loads that traditional air-cooled solutions struggle to manage.

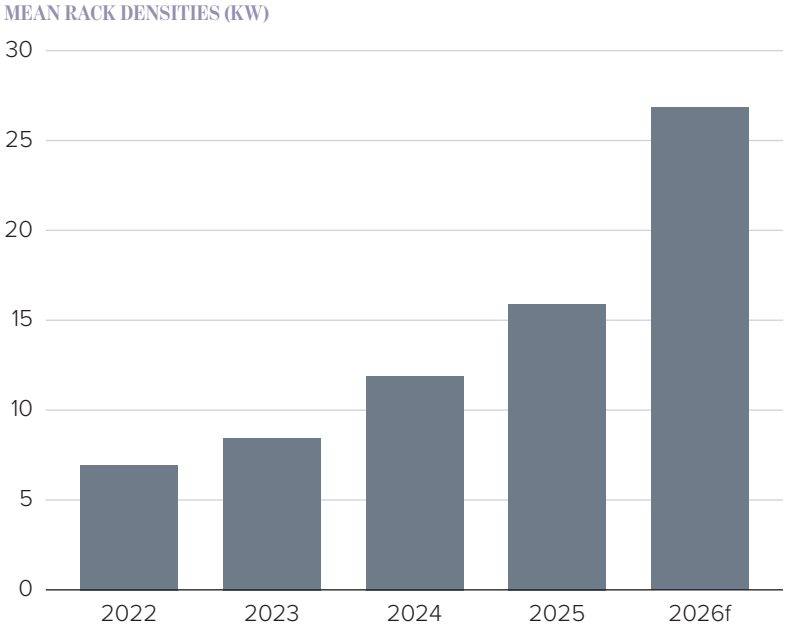
Higher density raises construction costs, but it also increases revenue potential. This is one reason individual asset sizes have grown dramatically (see 4). Not long ago, a 15–20 MW development was considered large. Today, single buildings exceeding 100 MW are increasingly common, and full campuses plan for multiple 100+ MW phases. Construction budgets can surpass $1 billion for a single structure, with stabilized valuations around $2 billion no longer an outlier. This magnitude has forced capital markets to evolve their underwriting, phasing strategies, and partnership models.

Densification also sharpens the focus on power quality, reliability, and resilience. With more value concentrated per rack and per room, the cost of downtime rises. It is no coincidence that leases have shifted to pass more operational risk to tenants that have the deepest insight into their workload tolerance and architecture. Owners, for their part, are incentivized to deliver robust, flexible infrastructure at scale — then allocate risk through lease structures and capital partnerships.

4

### Rack densities are on the rise

Evolution in server technology is concentrating computing power into smaller footprints, increasing the focus on quality, reliability, and resilience. According to AFCOM surveys, mean rack densities in data centers have grown by nearly 4x in the past 5 years.




Source: AFCOM.

## Capital markets

Once dominated by public REITs and a select group of private operators, the data center sector has now fully institutionalized. Pension funds, sovereign wealth funds, infrastructure managers, private equity firms, and multi-asset platforms all hold meaningful ownership positions. The emergence of net-leased, bond-like income profiles has attracted new categories of capital — particularly net-lease investors and insurance balance sheets — that prioritize long-duration, inflation-protected, creditworthy cash flows.

Fundraising continues to gain momentum. According to *PERE*,[3] overall real estate fundraising in 2025 recorded its first year-over-year increase since 2021, with data centers leading all sector-specific strategies. The asset class captured 37% of sector-focused commitments, outpacing traditional categories such as industrial and housing. Capital formation has been diverse in both strategy and sponsorship: institutional managers including PGIM and Principal have raised closed-end, data-center-dedicated vehicles, while operators such as Cloud Capital and Digital Realty Trust have also launched sector-specific funds. Blue Owl added to its closed-end series and expanded its platform with a new evergreen, open-end structure.

With capital being raised across the risk spectrum — but relatively few stabilized assets trading — investors have adjusted their approaches to secure exposure to the sector. Significant pre-leasing and build-to-suit activity has elevated the appeal of forward sales, giving developers efficient capitalization for new projects while offering investors enhanced returns relative to core acquisitions, coupled with contractual, long-term,

durable income streams. Bringing long-term capital in early has also eased concerns around the historic 'wall of takeouts' and the perceived sufficiency of exit capital.

Pairing this expanding equity base with diverse, scalable sources of financing is enabling the industry to grow at a pace that would have been unimaginable only a few years ago. Mirroring trends in the equity markets, data centers are capturing an increasing share of financing activity as well. Traditional bank syndicates, project finance lenders, and private credit providers are absorbing a substantial volume of construction loan demand. In 2025, asset-backed and commercial mortgage-backed securities markets reached record levels for stabilized data center assets, while large-scale structured transactions in the private placement and 144A markets have supported the sector's biggest financings.

Investors with infrastructure DNA are particularly well aligned with the scale and technical complexity of these developments. They underwrite long-term power strategies, phased delivery schedules, offtake certainty, and repair-and-replacement obligations with an emphasis on durability and downside protection. Capital structures have evolved in parallel — longer-tenor construction financing, JV frameworks pairing operator expertise with institutional balance sheets, and multiple forms of securitized takeout capital have all become common. The result is a deeper and more diverse capital base capable of meeting the sector's extraordinary funding requirements.

Even the largest 'AI factories' — multi-phase, power-anchored campuses in nontraditional or tertiary markets — are now financeable when power pathways, credit quality, and delivery sequencing align. Recent examples such as Meta's Hyperion campus in Louisiana and the Stargate projects supporting OpenAI workloads illustrate that individual project budgets exceeding $20 billion can attract appropriately structured capital. These transactions demonstrate a market increasingly comfortable not only with scale, but with the distinctive risk-return profile of AI-driven compute.

Importantly, the market is not limited to megaprojects. Liquidity remains healthy across the full deal-size continuum. Smaller assets, multi-tenant colocation facilities, and incremental expansion phases continue to draw active bidding. Pricing and competition, however, remain most sensitive to four core variables: power visibility, entitlement certainty, tenant credit quality, and lease structure.

## Real estate or infrastructure? The answer is both

Data centers sit uniquely at the intersection of real estate and infrastructure. As real estate, location is critical — proximity to end users and the right entitlements shape viability and speed to market. Yet they also depend on deep infrastructure: resilient power supply, robust network connectivity, and highly engineered mechanical systems. Straddling these categories expands the universe of potential capital, attracting investors that traditionally focus on either property or infrastructure — or increasingly combine both under a broader 'real assets' mandate. Data centers may defy simple classification, but investor appetite for the sector shows no sign of slowing.

## Enduring fundamentals

Forecasting the next wave of data center evolution is inherently difficult in a sector so tightly coupled with technological change. Yet history offers a useful guide. Facilities built in the early 2000s — long before the first iPhone reshaped consumer computing — are today being repurposed to host state-of-the-art AI training clusters and latency-sensitive inference nodes. The fundamentals that defined those early enterprise data centers remain just as critical today: compute requires robust connectivity, consumes significant power, and generates substantial heat. And the end-user appetite has only intensified, with a global population increasingly dependent on instantaneous access to vast volumes of data.

Technology will inevitably evolve and workloads will shift, but the facilities endowed with the essential raw materials — location, power, fiber, and entitlements — will retain enduring value and continue to adapt to each successive generation of compute. These long-term fundamentals underpinning the inherent residual value, coupled with durable income streams, provide an attractive investment profile to deep and diverse pools of capital.

[1] CBRE Research, CBRE Data Center Solutions H2 2025.
[2] CBRE Research, CBRE Data Center Solutions H2 2025 Asking rates for 250+ kW at N+1/Tier III requirements.
[3] PERE Fundraising Report Full Year 2025.

Kristina Metzger is Vice Chairman and Ben Wobschall is Executive Vice President, both with the Data Center Capital Markets team at CBRE.