Not All AI Data Center Demand Is Bankable

Why private capital needs to distinguish AI exposure from bankable infrastructure cash flow

By Richard Nohe
Digital Infrastructure Strategy & Risk Advisor
May 2026

Not every dollar of AI-driven data center demand will become infrastructure-grade cash flow.

That is the point private capital needs to hold in mind as more money moves into the sector. AI demand is real. Power-constrained capacity is scarce. Data centers are increasingly central to the economy. But in a crowded market, those facts can blur an important distinction: not every AI-exposed data center asset has core infrastructure risk, and a signed lease does not always convert AI demand into durable cash flow.

We can learn from similar versions of this cycle in the past. During the late-1990s and early-2000s internet infrastructure cycle, the internet thesis was right. The world needed more bandwidth, more fiber, more hosting, and more network infrastructure. The problem was not the long-term demand curve. The problem was that many of the companies filling data centers, buying network capacity, and building business models on top of that infrastructure did not survive long enough to support the capital structures underneath them.

One example that I was directly involved in was AboveNet's predecessor, Metromedia Fiber Network, which went bankrupt in 2002 after building one of the more ambitious fiber and data center platforms of that era. The company had real assets and real infrastructure, but the market around it changed faster than the capital structure could absorb. It was not alone. Global Crossing also filed for bankruptcy in 2002 with more than $12 billion of debt, and WorldCom did as well in one of the most significant collapses of the telecom cycle.

That history should not be reduced to a simple claim that AI is another bubble. The better lesson is more useful and more uncomfortable: **a technology thesis can be right while the cash flows financing the infrastructure are wrong.**

That is the issue now.

McKinsey projects U.S. data center power capacity rising from roughly 30+ GW in 2025 to 90+ GW or more by 2030, with hyperscalers expected to capture about 70% of forecast U.S. capacity through owned or leased options. It also expects inference to become more than half of AI workloads by 2030, representing roughly 30% to 40% of total data center demand.

Those are powerful numbers. But numbers like that can create a false sense of precision. They describe aggregate demand. They do not answer the investor's harder question: **which tenants, workloads, assets, and contract structures will produce durable cash flow through the cycle?**

Infrastructure pricing requires infrastructure cash flow.

Brookfield framed the issue well in a June 2024 paper: *Not all (core) infrastructure is created equal*. Core infrastructure is not just something essential or capital-intensive. It is expected to have long-term cash flow visibility, resilient demand, creditworthy counterparties, operating maturity, capital structure stability, and a useful life that extends beyond a single customer or business model.

That framework is exactly what should be applied to some, but not all, AI data centers.

A stabilized colocation facility with diversified enterprise customers is not the same investment as a single-tenant hyperscale build-to-suit. A powered shell is not the same as an operating campus. A lease to a hyperscaler is not the same as a lease to a venture-backed GPU cloud provider. A data hall designed around one workload may not have the same residual value as one that can serve several.

Yet the market often compresses these distinctions into one phrase: "data center exposure."

That phrase is becoming too broad to be useful.

The better question is not whether the asset has AI exposure. The better question is whether it has **bankable AI exposure**.

Tenant credit is not enough

In traditional infrastructure underwriting, counterparty credit matters. It still does. But in AI-driven data centers, tenant credit is only the starting point.

A data center lease is only as durable as the economic model sitting behind the tenant.

If the tenant is AWS, Microsoft, Google, or Meta, the payment-risk analysis is very different from a lease to a younger AI infrastructure company. But even with a strong hyperscaler, investors still need to understand control rights, expansion rights, renewal dynamics, workload flexibility, and residual-use risk. Hyperscalers are excellent counterparties in many respects, but they are not passive tenants. They are sophisticated ecosystem controllers.

The harder underwriting problems sit at the layers below hyperscale.

That is where private capital may find more accessible opportunity: AI-oriented colocation, GPU-as-a-service providers, neoclouds, managed AI platforms, powered shell delivery, structured credit, and retrofit plays. These are also the places where tenant economics may be least settled.

If a tenant is selling AI infrastructure services to third parties, the investor should not stop at lease term and rent coverage. The questions need to go deeper:

- What service is the tenant actually selling?
- Who are its customers?
- Are revenues contracted, recurring, or usage-based?
- Are margins durable after GPU cost, power, cooling, bandwidth, software, financing, and customer acquisition costs?
- Is demand tied to production workloads or experimentation budgets?
- Does the tenant have customer concentration risk?
- What happens if inference pricing compresses?
- What happens if model efficiency improves faster than demand expands?
- If the tenant fails, who can use the space next?

Those are not technical questions for the sake of being technical. They are cash-flow questions.

The inference-margin question

Inference is likely to be central to the next phase of AI infrastructure. Unlike training, which is capital-intensive and often harder to connect directly to revenue, inference is more recurring and more closely tied to commercial usage. It powers search, chatbots, recommendation engines, enterprise AI applications, and other real-time services.

That makes it attractive. It also makes the margin question unavoidable.

Who captures the economics of inference?

Is it the model owner? The hyperscaler? The GPU cloud? The enterprise application provider? The chip supplier? The data center operator? Or the capital provider financing the facility?

The answer is that each captures part of the economics, but not necessarily in proportion to the capital intensity, operating risk, or duration of the commitments being made. For data center investors, the underwriting issue is whether the tenant's share of those economics is sufficient, durable, and predictable enough to support long-term rent and infrastructure financing.

McKinsey notes that the growth trajectory for inference remains uncertain because hardware efficiency, software optimization, smaller fine-tuned models, improvements in precision formats, permitting timelines, regulatory limits, and grid constraints may all affect the pace and shape of demand.

That does not make the sector unattractive. It makes the sector interesting.

A strong investor does not need to know exactly how AI demand will evolve. But the investor does need to know whether the asset, tenant, and capital structure can survive more than one version of the future.

If inference becomes a large, profitable, recurring service layer, many AI-oriented data center investments will perform well. If pricing compresses, utilization disappoints, or enterprise adoption takes longer than expected, some tenants may find they leased long-lived infrastructure against short-lived revenue assumptions.

That is where the telecom lesson matters. The lesson was broader than timing: infrastructure can be directionally right and still fail financially when capacity is built ahead of monetizable demand and supported by tenants whose own balance sheets cannot survive a correction.

The question is not whether demand ultimately arrives. It may. The question is whether the tenant can finance the bridge between today's costs and tomorrow's durable revenue.

Crowded markets reduce the margin for error

One of the more telling comments from an interview with Mike Armstrong of GI Partners was not that data centers are attractive. Many people already believe that. It was that the market has become dramatically more crowded. He described typical marketed transactions moving from roughly five bidders in 2014 to 20 or 30 today, with more well-capitalized participants chasing the same product. He also noted that while more market participants can validate pricing over the long term, the short-term effect is increased competition and a reduced margin of safety.

That is exactly when underwriting discipline matters most.

Low capitalization rates and crowded bid sheets may show confidence in the sector. They do not prove that every asset is priced correctly for tenant risk, re-leasing risk, power risk, retrofit risk, or workload risk. Those risks sit below the cap rate, and they are often where the difference between infrastructure income and value-add exposure actually appears.

In a less crowded market, investors may have more room to be generally right. In a crowded market, they need to be specifically right.

Asset reusability depends on the workload — and on what capital is funding

In strong markets, investors focus on growth. In corrections, they rediscover reuse. But in AI data centers, "reuse" is not a generic concept. It depends heavily on the workload the facility was built to support.

The first distinction is between training and inference.

Training facilities are more likely to be large-scale, power-intensive, water- and cooling-sensitive, and less latency-constrained. McKinsey describes training workloads as requiring high-density campuses with advanced MEP systems, specialized hardware integration, and, in some cases, very high rack densities. Because training can tolerate more latency, these facilities can be located in remote, power-rich markets rather than close to dense end-user demand.

But power availability is only part of the story. Permitting and local politics are also part of the critical path. Large AI training campuses can create visible demands on grid capacity, water, land use, and local infrastructure, so community acceptance, environmental approvals, and utility coordination can affect timing, cost, and exit value even in markets that look attractive on power availability alone.

Those same characteristics also affect residual value. A large AI training campus may be valuable, but its likely customer universe is limited: hyperscalers, frontier AI labs, very large GPU cloud platforms, and a small number of other highly capitalized players. It may not be easily relet to ordinary enterprise or colocation demand. Its power, cooling, water, network, and campus configuration may be too specialized, too large, or too remote for a broad replacement-tenant market.

That does not make training facilities unattractive. It means they should be underwritten as highly specialized infrastructure with concentrated customer and residual-value risk.

It also helps explain why much of the largest-scale training infrastructure is likely to remain closely tied to hyperscalers, major AI labs, strategic technology suppliers, and balance-sheet-heavy financing structures. Recent market examples show how chips, cloud capacity, financing, and data center expansion can become intertwined. Nvidia and OpenAI announced a letter of intent under which Nvidia intends to invest up to $100 billion as OpenAI deploys at least 10 GW of Nvidia systems for next-generation AI infrastructure. Nvidia has also invested $2 billion in CoreWeave, with Reuters reporting that the investment is intended to help CoreWeave accelerate land and power procurement for data center expansion. Those structures may be rational for the participants, but they reinforce the point that the largest training platforms are not ordinary real estate assets. They are part of a strategic compute ecosystem.

Inference is different.

Inference workloads are more closely tied to recurring commercial usage: search, chatbots, recommendation engines, enterprise AI tools, and other applications that require responsiveness. McKinsey expects inference to become more than half of AI workloads by 2030 and notes that inference is more latency-sensitive and often needs to be closer to applications, storage, networking zones, and end users.

That makes inference-oriented data centers more accessible and potentially more attractive for private capital. They may be smaller, more distributed, more network-sensitive, and more capable of serving multiple customers or use cases. They also fit better with the traditional PE playbook: fund capacity, lease it to growing tenants, protect downside through reusability, and exit to a broader universe of buyers.

But that opportunity comes with a different risk. Inference assets may be more reusable than training assets, but they are also more exposed to the economics of the tenants selling inference services.

That is the central underwriting question. If PE is financing shell and core, the downside risk is different than if it is financing MEP, data halls, liquid cooling, GPUs, or tenant-specific improvements. Each layer has a different recovery profile if the tenant fails or the workload strategy changes. Shell and core may have a broader residual market. Specialized MEP may be reusable only within a narrower density band. GPUs may have a shorter economic life and greater technology-cycle exposure. Tenant improvements may have little value outside the original use case.

So the right question is not simply, "Can the data center be reused?"

The more nuanced questions are:

- Is this a training asset or an inference asset?
- Is PE funding the shell, the powered shell, the MEP, the data rooms, the cooling systems, the GPUs, or some combination of these?
- How broad is the replacement-tenant universe?
- Does the asset depend on one AI lab, one GPU cloud, one hyperscaler, or one workload profile?
- Are the tenant's service margins sufficient to support long-term rent after GPU, power, cooling, network, and financing costs?
- If the tenant fails, what exactly does the investor own, and who else can use it?
- Are exit options backstopped by a broader market, or dependent on another buyer believing the same AI thesis?

That is where the underwriting becomes more than real estate and more than technology. It becomes ecosystem economics.

The key for PE is to understand where its capital sits in the AI infrastructure stack, what margin pool supports that layer, and whether the exit is protected if the original tenant or workload strategy changes. A lease matters. But it is not enough. The investor needs to know whether the tenant's business model, the asset's reuse profile, and the funded infrastructure layer all support the same investment thesis.

A better test for private capital

The objective is not to avoid AI data center investment. The objective is to classify it correctly.

By the end of diligence, private capital should be able to answer four questions:

1. **What workload is the asset really serving?**
   Training, inference, general compute, or hybrid — and does the asset remain valuable if the workload mix shifts?

2. **What tenant economics support the lease?**
   Does the tenant capture durable margin after GPU, power, cooling, network, software, and financing costs?

3. **What layer of the stack is being funded?**
   Shell, powered shell, MEP, data halls, liquid cooling, GPUs, tenant improvements, or a blended exposure — and what is the recovery profile if the thesis changes?

4. **Who is the next user or buyer?**
   Is residual value supported by broad infrastructure utility, or by another buyer accepting the same AI growth narrative?

The practical test

For PE and infrastructure investors, the diligence should ultimately come back to a small set of harder questions:

Dimension	Question
Workload	Is this training, inference, general compute, or a hybrid — and does the asset remain valuable if the workload mix shifts?
Tenant economics	Does the tenant's business model generate durable margin after GPU, power, cooling, network, software, and financing costs?
Funded layer	Is the investor funding shell, powered shell, MEP, data halls, cooling, GPUs, tenant improvements, or a blended exposure?
Power	Is power contracted, deliverable, and expandable — or merely part of the development narrative?
Control	Who controls upgrades, expansion, workload changes, equipment rights, and re-leasing flexibility?
Residual value	If the tenant fails or changes strategy, who is the next user of the asset?
Capital structure	Does the financing assume core infrastructure stability where the exposure is actually development, retrofit, tenant, or technology-cycle risk?
Exit	Is the buyer universe broad, or dependent on another investor accepting the same AI thesis?

The objective is not to avoid AI data center investment. The objective is to classify it correctly.

Some AI data center exposure will deserve core infrastructure pricing. Some will be value-add. Some will be development risk. Some will be structured credit. Some will be tenant-credit risk. Some will be closer to venture-linked exposure sitting inside an infrastructure wrapper.

The work is knowing which is which.

Conclusion

AI will require enormous data center capacity. Power-secured sites will remain valuable. Hyperscalers will continue to reshape the market. New AI infrastructure companies will emerge, and some will become durable businesses.

But the investment opportunity is not as simple as "AI needs data centers."

The better question is whether a particular asset, tenant, workload, funding layer, and exit path together produce cash flow durable enough to merit infrastructure pricing. That is a more demanding question, but it is also where private capital can still find edge.

The early internet infrastructure cycle showed that the right technology thesis does not guarantee the right investment outcome. Demand can be real while tenants fail. Assets can be essential while capital structures break. Infrastructure can ultimately matter while the original investors are diluted, restructured, or wiped out.

That is the lesson worth carrying into the AI data center cycle.

The best investors will distinguish between demand and bankability. They will look past the label and ask where the economics actually sit, who controls the critical inputs, what margin supports the rent, and whether the asset remains useful if the original tenant or workload thesis changes.

Not all AI data center demand is bankable.

The capital that performs best will be the capital that knows the difference.

About the Author

Richard Nohe is a senior technology and infrastructure executive with more than two decades of experience across telecom, cloud infrastructure, data centers, regulatory risk, and complex commercial transactions. He has served as regional General Counsel for a global technology and telecom company and has advised on mission-critical infrastructure, digital transformation, and large-scale contracting frameworks. LinkedIn

References

Brookfield. "Not all (core) infrastructure is created equal." June 2024. https://www.brookfield.com/views-news/insights/not-all-core-infrastructure-created-equal#the-infrastructure-universe

McKinsey & Company. Chhavi Arora, Marc Sorel, Pankaj Sachdeva, et al. "The next big shifts in AI workloads and hyperscaler strategies." December 17, 2025. https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/the-next-big-shifts-in-ai-workloads-and-hyperscaler-strategies

Data Center Knowledge. Nathan Eddy. "Global Hyperscale Growth Persists Despite Grid and Land Constraints." December 16, 2025. https://www.datacenterknowledge.com/hyperscalers/global-hyperscale-growth-persists-despite-grid-and-land-constraints

Cool Vector. "Competition for Digital Infrastructure Deals is 'Really Positive'." Interview with Mike Armstrong, Managing Director and Head of Technology and Science Investments, GI Partners. February 2026. https://coolvector.substack.com/p/competition-for-digital-infrastructure

Los Angeles Times / Reuters. "Metromedia Fiber Files Chapter 11 Petition." May 21, 2002. https://www.latimes.com/archives/la-xpm-2002-may-21-fi-techbriefs21.1-story.html

Congressional Research Service. "WorldCom: The Accounting Scandal." Updated August 29, 2002. https://www.everycrsreport.com/reports/RS21253.html

The Washington Post. Christopher Stern. "Global Crossing Files for Bankruptcy." January 28, 2002. https://www.washingtonpost.com/archive/business/2002/01/29/global-crossing-files-for-bankruptcy/fffcad1b-f522-479a-a9d4-659e74165b05/

NVIDIA Newsroom. "OpenAI and NVIDIA Announce Strategic Partnership to Deploy 10 Gigawatts of NVIDIA Systems." September 22, 2025. https://nvidianews.nvidia.com/news/openai-and-nvidia-announce-strategic-partnership-to-deploy-10gw-of-nvidia-systems

Associated Press. "Nvidia to invest $100 billion in OpenAI to help expand the ChatGPT maker's computing power." September 22, 2025. https://apnews.com/article/openai-nvidia-investment-partnership-chatgpt-610d894d93f9be23c46762950997a67f

Reuters. Deborah Sophia. "Nvidia invests $2 billion in CoreWeave to boost data center build-out." January 26, 2026. https://www.reuters.com/technology/nvidia-invests-2-billion-coreweave-expands-partnership-2026-01-26/

Reuters. "Trump's push for more AI data centers faces backlash from his own voters." December 1, 2025. https://www.reuters.com/business/retail-consumer/trumps-push-more-ai-data-centers-faces-backlash-his-own-voters-2025-12-01/

Associated Press. "Trump's AI plan would sweep away key environmental law." July 29, 2025. https://apnews.com/article/national-environmental-policy-act-ai-data-centers-a1efa3f9a8b2022ded585fd8f1524cce