THE HIDDEN LAYER SERIES | ARTICLE 4

By Joe Batista | October 2026 | 12-minute read Find the series here at the IdeaVault

The Transformer Bottleneck: How Grid Physics Gates AI

Why energized megawatts, not chips, may decide where AI gets built

This is Article 4 of The Hidden Layer Series. Article 1, "The Map Nobody Has Drawn," laid out the AI Value Map. Article 2, "The Accidental Kingmaker," showed how bottleneck power migrates through the AI stack. Article 3, "Rationed, Not Priced," found that in CoWoS it shows up as control over who gets supply. Article 4 follows the constraint to the layer underneath everything.

CORE INSIGHTS

- AI's next bottleneck may not be compute, or even power generation. It is deliverable power: at the right place, at the right time, with permission to connect.
- Large power transformers are the most visible chokepoint inside that chain. Average power transformer lead times rose from roughly 50 weeks in 2021 to 120 weeks in 2024.
- Utilities plan in decades. Tech plans in quarters. PJM's 2027/28 capacity auction was the first in its history to fall short of its reliability requirement. The 2028/29 auction fell short again.
- In CoWoS, a supplier chose to ration. In constrained grids, no one chooses access is rationed through queues, regulatory decisions and available system capacity.
- The scarce asset isn't the data center. It's the megawatt that can be delivered to it. Our hypothesis: AI capacity will concentrate where those megawatts exist.

The Constraint Nobody Named

Power is the new GPU. That is the slogan. The reality is more specific, and more useful.

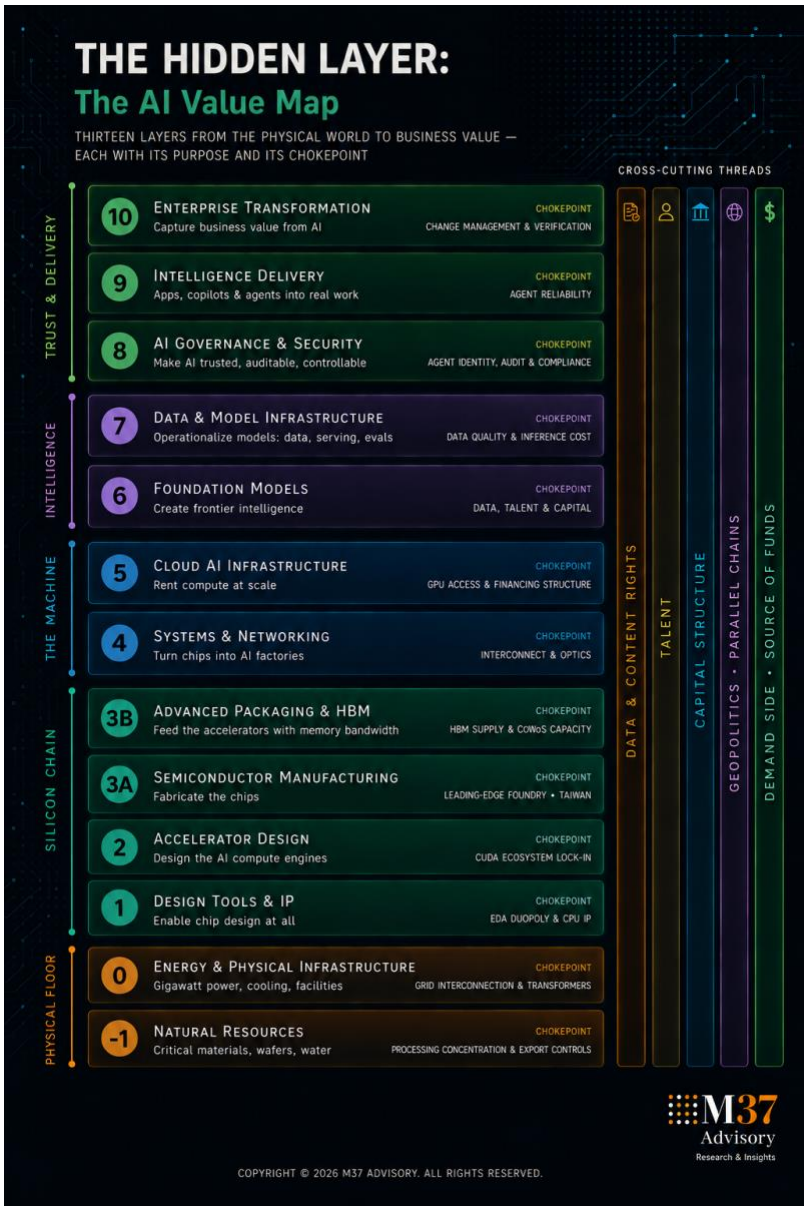
In November 2025, Microsoft's CEO said the quiet part out loud. "It's not a supply issue of chips," Satya Nadella said. "It's the fact that I don't have warm shells to plug into." The result, in his words: "a bunch of chips sitting in inventory that I can't plug in."

A warm shell is a building ready for equipment. But a warm shell is useful only if it can be energized. Nadella was direct about it: the biggest issue, he said, is power, and getting builds done fast enough close to power. Eight months later, on Microsoft's July 2026 earnings call, its CFO said "customer demand continues to exceed available capacity."

The gate is no longer simply compute. Nor is it simply power generation. The U.S. is adding generation, but new supply doesn't necessarily reach the places and dates where data centers need it. It is deliverable power: at the right place, at the right time, with permission to connect. And the visible constraint is often downstream of the real one.

This is Layer 0. The layer below everything, beneath the silicon everyone watches.

Where Layer 0 Sits



The Hidden Layer: The AI Value Map. 13 layers, each with its purpose and its chokepoint.

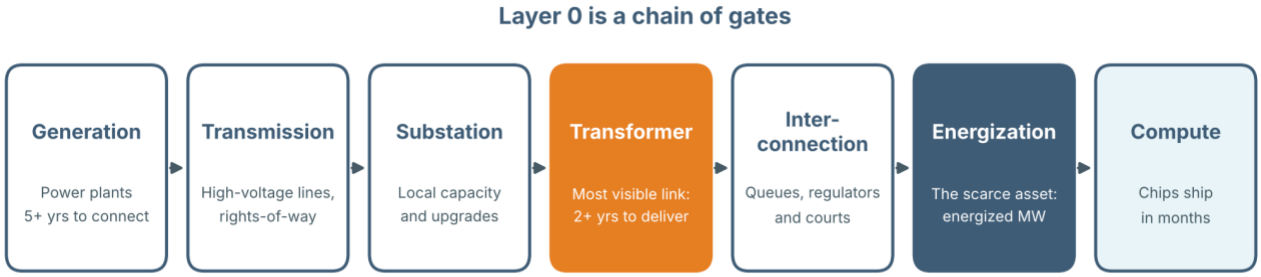
ON THE MAP: Layer 0 — Energy & Physical Infrastructure | Chokepoint: Grid Interconnection & Transformers

Everyone watches chips. Chips have earnings calls, launch events and benchmark wars. Electrons have rate cases.

That is why Layer 0 stays invisible. It has no product launches. It moves through utility commissions, grid operator filings and factory order books that few strategists read.

But constraints collapse downward. When Layer 0 prevents energization, investment in every layer above it can't deliver its intended value. A perfect accelerator in a dark building produces nothing. The four largest hyperscalers plan roughly $725 billion in 2026 capex, by a Financial Times tally of the companies' early-2026 spending plans. That is planned spending, not money earmarked for power, but much of it depends on this same gate.

Layer 0 is not one bottleneck. It is a chain of them.



A delay at any link holds up everything to its right. The transformer is the most visible link, not the only one.

The transformer is the most visible physical constraint inside that chain, which is why it gives this article its title. But it is one link, not the whole gate.

Three Examples That Show the Gate

① The Transformer Timeline

Most people think the building is the long pole. It isn't.

In Virginia, the world's largest data center market, Dominion Energy told Bloomberg in 2024 that new large data center connections could take up to seven years. JLL's 2026 outlook finds that in many major markets, average grid interconnection timelines now extend beyond four years.

Transformers are one reason. Wood Mackenzie found average power transformer lead times rose from roughly 50 weeks in 2021 to 120 weeks in 2024.

Money doesn't fix it fast. Wood Mackenzie also reports power transformer prices up about 77% since 2019, with about 80% of large power transformers imported.

Illustrative timelines, not directly comparable measures: each row describes a different kind of wait.

Step	Kind of wait	Typical or reported time
AI accelerators	Procurement	Months
Data center shell	Construction	Often about 2 years
Large power transformer	Order to delivery (Wood Mackenzie)	1.5 to 4 years; average over 2
Large grid connection, Virginia	Interconnection, worst case reported (Dominion, 2024)	Up to 7 years

READER RESEARCH

Rich McGrath, Systems Engineer at Everpure, responding to Article 1, showed the constraint I had oversimplified. It isn't just "power scarcity." It is transmission build times plus transformer manufacturing plus multi-year approvals. That is decade-scale, not year-scale. Their comment shaped this article.

② Utilities Plan in Decades. Tech Plans in Quarters.

Investor-owned electric companies project $1.4 trillion of total investment from 2026 through 2030, with 2026 their 14th straight record year, according to EEI. That covers all their generation, grid and other spending, not just what serves data centers.

It still isn't keeping pace. Utility capital moves on rate cases and five-year plans; tech capital moves on 18-month model cycles.

PJM, the grid operator for 13 states and Washington, D.C., shows the mismatch in its capacity auctions, which buy the generation the grid will need. Its 2027/28 Base Residual Auction cleared at the price cap and fell about 6,500 MW short of its reliability requirement, the first shortfall in PJM's history. The 2028/29 auction cleared at the cap and fell short again, by 6,831 MW. Both figures come from PJM's July 2026 results release, which counts auction and self-supplied (FRR) capacity together. PJM points to large data center loads in its forecast. A shortfall doesn't mean the lights go out. It means the forward auction didn't secure enough capacity to meet the reliability target without further action.

Dominion's contracted data center pipeline has more than tripled in three years, to 53.8 GW in July 2026. Only 12 GW of it sits under firm electric service agreements, so much of it may never connect. That is exactly what queue audits are now testing.

The result: capital can be ready years before the power is.

③ Regional Bifurcation

Power doesn't spread evenly, so AI won't either.

Texas shows what a rush looks like. By August 2026, ERCOT's large-load requests reached about 474 GW. Those are requests, not committed demand, but Governor Abbott called the total more than five times Texas's record peak demand, with about 90% from data centers. He ordered an audit. In its Batch Zero update for its September 2026 board meeting, ERCOT stated that it "has paused the energization of new Large Load data centers and crypto mining facilities" of 75 MW or more, until verification, audit and a community impact review are complete. The pause applies to facilities not yet connected.

The constraint is not uniquely American. In highly constrained grids, regulators and courts increasingly decide which loads get connected, and when. In the Netherlands, TenneT's cleaned-up high-voltage waiting list still holds about 38 GW of requested consumption capacity. In April 2026, a Dutch court ruled that TenneT was not obliged to resume the connection process for a developer's planned data center at Vijfhuizen, near Amsterdam. The developer had no signed connection contract, the grid there is overloaded and the requested capacity is unavailable, so the project stays on the waiting list. It is one dispute, not a general rule for Dutch data centers. But it shows who decides when the grid is full.

This is where Layer 0 becomes a business story. When AI follows power, data center clusters can pull capital, suppliers, jobs and tax bases with them.

> ***"AI capacity increasingly concentrates where energized megawatts exist, not only where customers, talent or tax breaks are."***

Virginia, with its customers and tax breaks, is still the largest market. But training workloads are generally more geographically flexible than latency-sensitive inference. That gives some training capacity more freedom to follow available power, although network connectivity, cooling, cost and workload requirements still constrain location. For a growing share of AI capacity, the strategic question changes from "Where should we build AI?" to "Where can AI actually be energized?"

Why This Pattern Keeps Repeating

We have seen a version of this before. As an analogy, not proof: in the late-1990s telecom boom, optical gear improved every year, while many binding constraints were physical: rights-of-way, permits and the last mile. Some of the most durable advantage went to companies that controlled scarce physical paths, not simply to those making the fastest equipment.

Today demand is ahead of the physical layer, not behind it. But the lesson carries over: when a physical layer gates a technology wave, value tends to flow to whoever secures access to the scarce path.

Three things make Layer 0 harder than the bottlenecks earlier in this series:

- **It is regulated as a commons.** Grid capacity is allocated by queues, commissions and courts, not sold to the highest bidder.
- **It moves at the speed of approvals.** One audit order in Texas paused energization of new large data centers until each is verified.
- **It is tied to place and permission.** Adding HBM or CoWoS capacity is hard too: equipment, cleanrooms, qualification and yield all take time. But a new packaging line can serve customers anywhere. Layer 0 combines physical scarcity with geographic dependency, manufacturing lead times and public approvals. Capital alone can't move the grid on a tech timeline. It can build around it with on-site generation, at a cost and with its own permits.

WHAT WE MEAN BY "STRUCTURAL"

Our hypothesis: this constraint is structural, not cyclical. It will not clear within a normal two-to-three-year tech investment cycle, because its slowest links (transmission, approvals, transformer factories) run on five-to-ten-year cycles.

What would make us wrong: inflated queues. In power plant queues, about 75% of requested capacity has historically withdrawn. Texas's audit is now testing how much of ERCOT's queue is real. If large-load queues shrink the same way, the gate may open faster than we expect. Transformers may also prove the cyclical link, because new factories are coming. More than 20 manufacturers have announced 27 new or expanded U.S. transformer plants in two years, according to Newton-Evans Research, including a Siemens Energy factory for large power transformers in Charlotte, announced in 2024 with production planned from 2026. And if AI demand cools, queues shrink on their own. Transmission and approvals are the links we expect to stay structural.

The prediction below is how we will test it.

Constraint ≠ Bottleneck Power, at Layer 0

We believe the constraint and the urgency at Layer 0 are among the highest on the map. What makes it different is market structure: high constraint, but less leverage for any single supplier.

THE M37 LENS

Three questions from Article 3: **Constraint** (can the market get enough supply?), **bottleneck power** (does someone control what customers can't substitute?) and **value capture** (does that control become price, margin, allocation or strategic advantage?). Bottleneck Power = Constraint × Market Structure × Customer Urgency is a heuristic, not an equation.

Layer 0 splits into two very different markets.

Transformers are a supplier market. A small group of manufacturers build the largest units, factory slots are rationed, and prices are up about 77% since 2019. That is classic bottleneck power, and it scores like one.

Grid access is a regulated commons. No one can bid their way to the front of the queue. Even where capacity is priced, the price is capped. PJM estimates its 2027/28 price would have been about 60% higher without the cap, adding nearly $10 billion to the bill.

The queue isn't the only path. Developers can build their own generation. JLL expects operators to increase behind-the-meter power arrangements, and Virginia data centers are already proposing on-site gas turbines. That route runs into a different set of gates: air permits, gas supply and state law.

In CoWoS, TSMC held the leverage and chose to ration. For grid connections, no one chooses.

"For grid access, rationing isn't a strategy. It's the rule."

Bottleneck Power Score: Large Power Transformers

M37 Directional Research Index: the same six dimensions used for HBM in Article 2 and CoWoS in Article 3. The band is set by the simple average of six equally weighted dimensions: High (8.0 and above), Elevated (7.0–7.9), Moderate (5.0–6.9), Low (below 5.0). Equal weights are a deliberate judgment: we have no sound basis yet for saying one dimension matters more than another, and equal weights keep the bands comparable across articles. We report the band, not the decimal. Directional, not statistically calibrated: a one-point difference between markets is a judgment, not a measurement.

Dimension	Score (0–10)	Why
Supplier Concentration	7	Three global leaders (Hitachi Energy, Siemens Energy, GE Vernova) plus a handful of specialists can build the largest units; most are imported
Qualification Time	8	Custom engineering per unit; lead times over two years
Replication Cost	7	New factory lines are announced, but take years to reach output
Switching Cost	6	Units are site-specific, but suppliers are interchangeable at order time
Customer Expected-Return Urgency	9	Every idle month strands chips already bought
Capital Availability	9	Measures buyers' ability to absorb higher prices and long waits, not suppliers' access to capital. High today, which strengthens supplier leverage
Band	**Elevated**	**One step below HBM 2024–25 and CoWoS 2026 (both High)**

Why only Elevated? Buyers can switch suppliers at order time, and capacity can be added, slowly. Capital cuts both ways here: today it lets buyers pay up and wait, which strengthens suppliers' hand; over time the same capital funds new factories, which weakens it. We score today. Grid access is not scored at all: it is not a supplier market, so the index doesn't fit. That gap is itself the finding. The most binding part of Layer 0 can't be measured with a supplier-power tool. Regulated commons need their own lens. Article 5 looks at how that rationing works.

The Scarce Asset: The Energized Megawatt

If grid access can't be bought at the front of the queue, value goes to whoever already holds it.

The land market is already pricing this, according to Altus Group. Its August 2026 analysis found that sites within half a mile of a substation trade at about twice the price of comparable sites two to five miles away, and that committed data center load is the strongest variable in land pricing. One caution: being near a substation is not the same as having deliverable power. A site can sit next to one and still lack transmission, transformer capacity or approval. So, land prices support the energized-megawatt thesis; they don't prove it. What they show is that the market is already paying for the chance of power.

"Proximity to power is not access to power."

Hyperscalers are increasingly talking about capacity in gigawatts, not just dollars. Amazon's CEO said AWS added 3.9 GW of power in 12 months and expects to double it again by end-2027. Microsoft said in July that it added another gigawatt in the quarter and is on track to roughly double capacity in two years.

Together, these point to an emerging strategic constraint. How far it will decide where AI capacity concentrates is still a hypothesis, and it's the one our predictions test.

> ***"The scarce asset isn't the data center. It's the megawatt that can be delivered to it."***

If power is the gate, secured access to power becomes a strategic capability. Vertical integration is one route. Others include long-term power purchase agreements, co-locating with generation, behind-the-meter generation, buying already-energized sites, utility partnerships, and flexible-load agreements that trade some uptime for an earlier connection.

Three Questions You Should Be Asking

1. **Where is your AI compute's power coming from?** Is it already energized, or sitting in a queue?
2. **Does your utility's capex cycle match your AI roadmap?** If their substation lands in 2030 and your plan ships in 2027, you have a gap no budget can close.
3. **Which path to power fits your strategy?** Contracts, co-location, onsite generation or energized sites: each trades cost, speed and control differently. Watch who is choosing which.

The Prediction

Same rules as Articles 2 and 3: checkable by anyone, marked in public in October 2028, including where we were wrong. The first two are direct tests of the structural hypothesis. The last three are supporting indicators: they track the story but can't prove it on their own.

Direct tests

Prediction (by Oct 2028)	Confirms the hypothesis	Falsifies it	Status
Grid connection waits stay above four years	JLL's published average grid connection wait for its primary data center markets stays above four years in each annual outlook through 2028 (2026 baseline: more than four years)	The average falls to four years or less	Open
Large power transformer lead times stay above two years	Wood Mackenzie or an equivalent published survey shows over 104 weeks	Lead times fall below 104 weeks	Open

Supporting indicators (these can reinforce the thesis, but they don't test it)

Prediction (by Oct 2028)	What it would show	What would weaken it	Status
PJM's next two capacity auctions also clear short of the reliability requirement, or PJM uses emergency or backstop procurement to close the gap	Supply is still not keeping pace with forecast load	Either auction meets the requirement on its own	Open
Alphabet and Meta join Amazon and Microsoft in reporting company-wide capacity in GW (online or added) on two consecutive earnings calls	Energized capacity has become the metric investors expect	One or both keep reporting only capex dollars or single sites	Open
Powered land keeps its premium	Altus Group or equivalent data show substation-adjacent sites still at 1.8× or more comparable sites farther away (about 2× in 2026)	The premium falls below 1.8×	Open

Why 1.8×? We set the bar just below today's roughly 2× premium, so ordinary noise doesn't count as a change, but a real easing of power scarcity would.

How we'll read the result: these two tests show whether the constraint persists, not exactly which link is binding. If grid connection waits stay above four years, the constraint persists. If they stay long while transformer lead times fall, the transformer wasn't the binding link: we were right about the grid and wrong about the transformer. We would expect transmission and approvals to be what remains, as we argued above, but these two indicators can't prove that on their own. If grid waits fall to four years or less, our structural hypothesis was wrong. Either way, we'll say so in public.

THE BOTTOM LINE

Today, power is becoming the constraint that can stop everything else. The transformer is the visible chokepoint. The business problem is the whole path to energization.

Unlike chips, power can't simply be ordered. It is negotiated with utilities, grid operators and regulators on multi-year cycles, and in the most constrained grids it is rationed, not sold.

When power becomes the gate, energized capacity becomes a strategic asset. The next AI winners may not be decided only by who has the best models or chips, but by who secured access to the physical path that lets them run, first.

So, ask yourself: Where is your compute getting power, and who already holds the megawatts you will need?

What constraint am I missing? What does your region look like? Show me your power constraint. The map improves when you do.

How This Article Was Tested

Before publication, each figure in this article was checked against the source linked beside it, and against the original wherever one is published: PJM's auction results, ERCOT's board materials, the Dutch court's own ruling summary, Berkeley Lab, EEI, Dominion and Microsoft, with Wood Mackenzie, JLL, Altus Group and Newton-Evans for industry data. Where we relied on a governor's letter or trade reporting, the article says so. The draft then went through a claim-by-claim devil's-advocate review and two rounds of editorial factchecking. Several figures and claims changed as a result, including our description of the Dutch court case. Special thanks to Rich McGrath, whose research on transformers and transmission shaped the argument; to John Ellis, founder and editor of News Items, whose post "Power: The physical side of AI and AI-adjacent technologies" tracks the same constraint; and to Chris Coomey, Senior Vice President and Chief Information Officer at St. Mary's Credit Union, for engaging with the series.

Practitioners in utilities, grid planning, transformer manufacturing and data center development: if you see where the evidence is thin, say so in the comments. Challenges that change the analysis will be credited by name in the next update.

About the Author

Joe Batista advises boards, C-suite executives, and go-to-market teams on technology value creation, growth strategies, and the opportunities hidden beneath major technology shifts. He has spent more than 40 years in the technology industry, guiding companies on growth and technology decisions, spanning the minicomputer era at Digital Equipment Corporation (DEC) through client-server, cloud, virtualization, and now AI, with executive experience at Dell Technologies and Hewlett Packard Enterprise.

Through M37 Advisory, Joe helps leadership teams recognize recurring patterns, uncover second-order effects, find the opportunities others overlook, and identify the strategy worth betting on. His work is grounded in a simple belief: **What if the biggest obstacle to your growth isn't the competition, but the way you think?**

> *"Most executives have been bombarded with 'theories' about what might work. Based on our research, Joe's extensive playbook is rooted in creating alignment of interest and business outcomes that leaders should take notice of."*
>
> — Jim Dickie, Research Fellow & Sales Mastery Founder, CSO Insights

Author Contact

Email: Chief.Creatologist@M37-Advisory.com
Website: https://www.m37-advisory.com
Idea Vault: https://www.m37-advisory.com/idea-vault

NEXT IN THE SERIES

Article 5: The Price of the Queue — Same layer, the other side of the gate: the rules. How regulation and policy decide who gets energized first, and who pays for it.

To subscribe to the series, complete the form

Sources

Grid operators, utilities and public agencies

- PJM: 2027/2028 Base Residual Auction report (Dec 2025)
- PJM: 2028/2029 capacity auction results (Jul 2026)
- ERCOT: Batch Zero update to the board (Sep 2026)
- Berkeley Lab: Queued Up, 2026 Edition (2026)
- EEI: $1.4 trillion investment, 2026–2030 (May 2026)
- Dominion Energy: Q2 2026 earnings call slides (Jul 2026)
- Rechtbank Gelderland: TenneT not obliged to resume connection process (ECLI:NL:RBGEL:2026:3358) (Apr 2026)
- Dutch parliamentary letter on grid congestion, via Stibbe (Feb 2026)

Company statements

- TechCrunch: Nadella on power and chips in inventory (Nov 2025)
- Microsoft: FY26 Q4 earnings call (Jul 2026)
- The Register: Amazon Q4 2025 earnings call (Feb 2026)

Industry research

- Wood Mackenzie: Transformer lead times (Apr 2024)
- Wood Mackenzie: Constraints hit US transformer supply (Aug 2025)
- JLL: 2026 Global Data Center Outlook (2026)
- Altus Group: Where value is created in data center land (Aug 2026)
- Siemens Energy: New U.S. large power transformer factory (Feb 2024)
- Newton-Evans Research: March 2026 assessment of the U.S. power transformer industry (Mar 2026)

Reporting

- Bloomberg: Virginia data centers face seven-year wait for power (Aug 2024)
- Virginia Mercury: Data centers want to build their own gas turbines (Jul 2026)
- Tom's Hardware: FT tally of 2026 hyperscaler capex (2026)
- John Ellis, News Items: Power. The physical side of AI and AI-adjacent technologies (Oct 2026)
- Utility Dive: Texas pauses data center interconnections (Aug 2026)

Keywords: AI infrastructure, Layer 0, transformers, grid interconnection, energized capacity, PJM, ERCOT, AI Value Map, M37 Advisory