Clarke

2026

The *intelligence layer* for **AI Infrastructure.**

The $500 Billion Question

In October 2024, Meta announced plans for a 1-gigawatt-class data center campus in Louisiana. The same month, xAI began construction on a 100,000-GPU facility in Memphis. Microsoft secured a 10 GW energy framework deal with Brookfield spanning multiple regions. OpenAI and its partners launched the first Stargate data center in Texas and evaluated additional 500+ MW sites across the Southwest and Midwest.

Together, these projects represent over $50 billion in infrastructure investment — announced in a single quarter. **Scale this forward, and the industry is staring at a $500B+ buildout over the next 5 years.**

Yet ask three different analysts what "1 gigawatt of capacity" actually means, and you'll get three different answers. Ask when these facilities will be operational, and you'll find six-month (or multi-year) discrepancies. Ask which markets are constrained by power, and you'll encounter speculation dressed as analysis.

The AI infrastructure market has a fundamental information problem: there's no shared language, no verified timeline data, and no forward-looking framework for capacity planning.

Clarke is a structured data and forecasting platform for AI-oriented data centers—turning public filings and project disclosures into decision-grade capacity intelligence.

Our mission is to become the intelligence layer for AI infrastructure — mapping, verifying, and forecasting the physical compute that powers frontier-scale AI. Clarke provides the neutral, evidence-based infrastructure intelligence that operators, investors, utilities, and policymakers need to make long-horizon decisions under real constraints.

AI Infrastructure Demands a New Approach

> *"Since there is no central clearing house regarding the numbers in operation, under construction, or in the planning stage, it is imperative that some type of future dashboard be developed."*
>
> — The Brookings Institution, 2025

Traditional data center intelligence was built for cloud computing: tracking square footage, PUE metrics, and tenant mix across colocation facilities.

AI infrastructure operates at a completely different scale:

Density. AI facilities run at 50–100 kW per rack versus 5–8 kW for traditional data centers — a 10x jump that transforms cooling architecture, power delivery, and site economics.

Power Scale. Hyperscale sites that once topped out at 50–100 MW are being replaced by multi-GW campuses that rival industrial manufacturing complexes, and the power needs of a small city.

Speed. Construction timelines have compressed from 24–36 months to 12–18 months for shells and initial phases, putting acute pressure on supply chains, grid interconnections, and permitting.

Grid Dependency. These campuses increasingly require utility-scale transmission upgrades, new substations, on-site generation, and multi-year regulatory processes — often just to connect a single load.

Capital Intensity. Meta's Richland Parish campus alone is a $10 billion project, backed by new power infrastructure and long-term energy agreements.

The problem is definitional chaos: "capacity" means different things to different actors, "under construction" spans land-work to energization, and "AI-ready" can describe anything from conventional colo with headroom to liquid-cooled GPU clusters.

Clarke's thesis is simple:
AI-grade infrastructure requires *AI-grade intelligence.* The market needs a neutral layer that turns fragmented figures into comparable, decision-grade insights.

The Market Failure:
A Real-World Case Study

These are not hypothetical dynamics. They're visible today in one of the highest-profile AI infrastructure projects in the United States: Meta's data center campus in Richland Parish, Louisiana. In late 2024, Meta and state officials announced a four-million-square-foot, $10 billion campus on a 2,250-acre site, describing it as the company's largest data center to date and part of a roughly gigawatt-scale buildout.

Energy-sector analysis, including work from RBN Energy, described configurations in which Meta could ultimately draw 2 GW of "compute" load backed by new combined-cycle gas plants and transmission investments, with scenarios stretching toward 5 GW of total generation.

Local and national coverage variously referred to the project as a 2 GW data center, a 2.26 GW gas-plant complex, or a 5 gigawatt AI campus, often without clearly separating critical IT load, site power, and the broader expansion envelope.

Each figure is defensible but describes a different layer.

- Crit IT vs cooling and overhead
- Initial phase vs full build-out
- Committed vs optional or contingent capacity

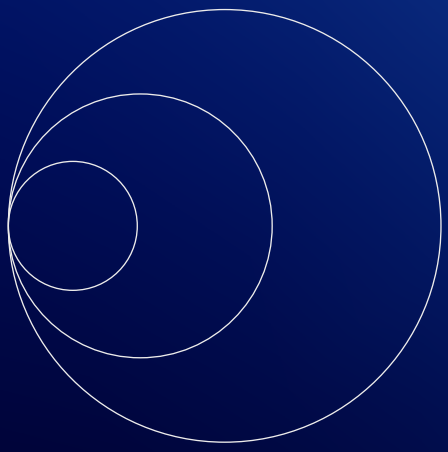
"Meta's Louisiana data center" is 2, 2.26, or 5 GW—with no breakdown of:

- Crit IT vs total site load
- Phase 1 vs total master plan
- Firm capacity vs contingent expansion

For grid planners, regulators, investors, and local communities trying to model demand, this is more than semantics.

Clarke is designed to turn exactly this kind of ambiguity into structured intelligence.

Clarke's Approach to AI-Compute Intelligence



Clarke's vision is to be the neutral intelligence layer for AI infrastructure — the framework that makes the physical AI-compute landscape legible.

We do that by combining three layers:

1. Clarke Standards – an open, versioned vocabulary for AI-oriented data centers.
2. Clarke Unified Facility Schema (CUFS) – a structured data model that standardizes facility intelligence for machine readability and analytical precision.
3. Clarke Intelligence Platform – the layer that delivers actual insight: facility dossiers, regional grid views, and forward-looking capacity forecasts.

The standards and schema aren't the product; they're the foundation under the product. Clarke exists so that investors, operators, utilities, and policymakers can answer a simple question with precision:

What is actually being built, where, on what grid, and how far can it realistically scale?

1.

Clarke Standards

To make the AI-compute landscape comparable, Clarke publishes *Clarke Standards* — an open, versioned vocabulary for AI-oriented data centers. The goal isn't to win an argument over terminology, but to give the market a shared baseline that Clarke then builds intelligence on top of.

At a glance, Clarke Standards pin down:

- **Capacity metrics.** Clear separation between I*T Capacity — Live*, *Site Power Contract*, and *IT Capacity — Planned*.

- **Facility classifications.** A taxonomy that distinguishes AI data centers, AI-native and frontier AI facilities, AI-capable hyperscale sites, and general-purpose data centers — and defines which facilities belong in the *Clarke Index*, our unified list of facilities.

- **Development and cooling context.** Where a facility sits in the entitlement and build phase and how it is cooled, as a gauge for density and AI workload readiness.

These definitions are based on systematic analysis of operator disclosures, utility filings, permitting documents, and regulatory submissions across the U.S., converging around how the market actually uses these terms rather than how marketing copy describes them.

Clarke Standards are public, free, and open for reuse with attribution. They exist so that Clarke — and anyone else — can speak about AI infrastructure in a way that's stable over time.

The full set of definitions is published at *Clarke Standards (v1.0)* clarkeindex.com/standards.

02.

Clarke Unified Facility Schema (CUFS)

If Clarke Standards define the language, the *Clarke Unified Facility Schema (CUFS)* is where that language becomes data.

CUFS is a purpose-built data model for organizing AI-infrastructure intelligence across:

- Facility identity and operator ownership
- Geographic location and regional grid context
- Capacity specs with source and verification tagging
- Development timelines and construction milestones
- Cooling architecture, density, and AI workload profile
- AI classification and cluster-presence signals

Every facility record in CUFS carries:

- **Source provenance.** Operator filings, utility dockets, regulatory documents, or secondary media/analyst sources, tagged field-by-field.
- **Verification level.** Confirmed, probable, or estimated (with Clarke's methodology clearly labeled when used).
- **Temporal validity.** When a data point was applicable, when it was last updated, and which historic projections have since been revised.

CUFS is the backbone that enables Clarke to *normalize* complex projects into a single, structured view. It distinguishes over 70 facility data points and prepares the contents for the intelligence layer to perfom it's synthesis.

03. Clarke Intelligence Platform

On top of Clarke Standards and CUFS sits the Clarke Intelligence Platform — where definitions and schema data turn into decision-grade insights. These capabilities are delivered through the Clarke interface and, over time, programmatic/API access.

1. Facility Intelligence

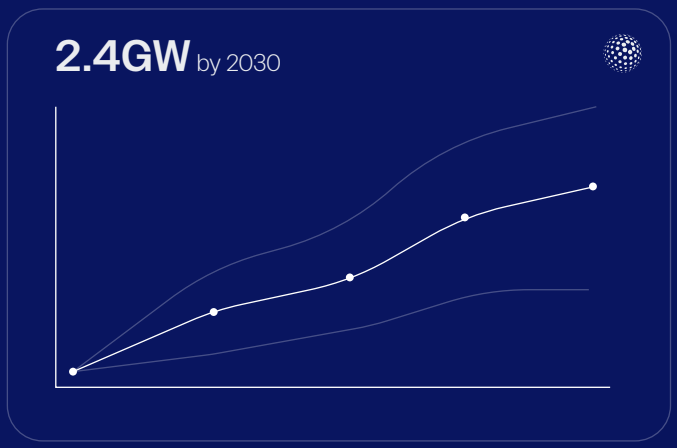
Clarke maintains structured profiles of U.S. AI-related data centers — purpose-built AI campuses, high-density retrofits, and strategic colo sites near critical grid nodes. Each record standardizes capacity, development status, cooling and density, AI classification, and core grid context, all tagged with provenance, verification level, and last-updated timestamps. Detailed breakdowns and regional roll-ups are available in paid tiers.

2. Regional & Grid Context

AI infrastructure is concentrating in a handful of patterns — Southwest solar- and gas-anchored hubs, Midwest legacy-grid corridors, Southeast utility-partnership growth zones, and cloud-dense regions with limited headroom. Clarke tracks these dynamics alongside ISO/RTO conditions, transmission projects, and utility resource plans, turning them into structured regional intelligence.

3. Capacity Forecasting

Clarke produces forward-looking views of AI infrastructure build-out through 2030 — with ranges for major operators and key regions. Forecasts are grounded in project disclosures, filings, timelines, and grid and generation constraints, and expressed as Conservative, Base Case, and Optimistic bands with Forecast Vintage dates and tracked accuracy over time; operator-level projections and detailed confidence bands sit behind paid tiers.



Clarke's Advantage

For Infrastructure Investors

Clarke normalizes operator capacity claims across the market, turning fragmented "2 GW" announcements into comparable intelligence about what's actually being built and when. Regional power availability and transmission constraints become structured inputs for risk/return analysis rather than narrative speculation. Investors can forecast AI-driven data center demand by market and operator using transparent, auditable methodologies — not analyst guesswork.

For Utilities & Energy Developers

Clarke provides structured views of projected data center load by region, timeframe, and operator—grounded in verifiable project data rather than speculation. This clarity helps utilities identify high-value interconnection and generation opportunities while stress-testing resource plans against plausible demand scenarios, turning a reactive posture into a strategy.

For AI Companies

Clarke enables operators to benchmark competitors' physical infrastructure trajectories and identify which markets offer genuine capacity versus power-constrained environments. This intelligence supports multi-year capacity planning with realistic timelines and grid-aware assumptions, helping companies position training clusters where power and interconnection actually exist rather than where press releases suggest they might.

For Equipment Vendors & Contractors

Clarke tracks verified project pipelines across operators and regions, replacing rumor-driven sales strategies with structured intelligence. Vendors can forecast cooling, power, and construction demand in line with capacity projections that carry provenance and confidence bands, enabling sales and partnership around projects that are actually moving forward.

Governance & Data Integrity

Clarke operates under a simple principle: if a number matters for capital allocation or grid planning, it needs provenance and context.

Evidence-Based Intelligence. Every data point is traceable to publicly accessible sources: operator disclosures, utility/regulatory filings, environmental impact statements, dockets, or verifiable media/analyst work.

Verification Tiering. Clarke distinguishes between Decision-grade (primary sources), Probable (multiple independent secondary sources), and Speculative (limited information, directional only) data, and clearly labels any **Clarke Estimate** values at a field level.

Temporal Accuracy. Capacity figures carry timestamps and validity windows. Superseded projections remain in the system to enable forecast-accuracy tracking.

No Proprietary Sourcing. Clarke doesn't rely on leaks or confidential information. The value comes from structure, methodology, and an accurate track record.

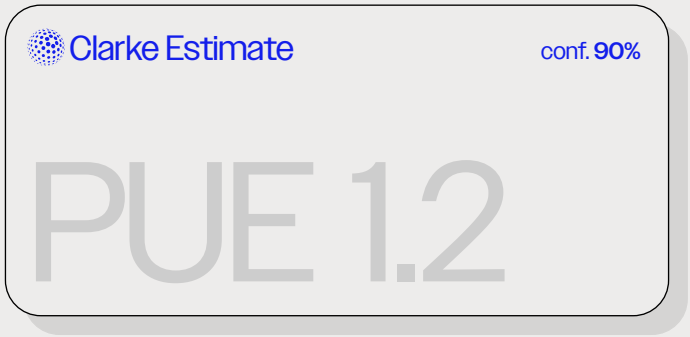
Operator Neutrality. Clarke doesn't advocate for or against specific operators or projects. The platform presents findings and insights and acts a neutral party.

This governance model is what allows Clarke to become a trusted intelligence layer for operators, regulators, and capital providers who may disagree about what should happen but still need to agree on what's actually being built.

Clarke publishes scenario-based forecasts that evolve as new information emerges, building a transparent track record that customers can audit over time.

Clarke Estimates

Structured Modeling When Primary Data Is Missing



In many AI data centers, some of the most important fields—true IT capacity, realistic PUE, or AI-specific build-out — are not disclosed in primary documents. Rather than leaving those fields blank, Clarke uses *Clarke Estimates:* clearly labeled, modeled values that fill gaps where primary data points aren't publicly accessible. These are always presented as Clarke's view, not operator claims, and are tagged in the schema so users can immediately see which fields are estimated versus sourced.

Clarke Estimates are built from a repeatable methodology that combines disclosed signals (i.e. site power contracts, cooling technology) with comparables from Clarke's growing facility dataset. From there, we apply a standardized modeling technique against similar facilities in the same market. Each estimate is associated with a methodology version, confidence level, and "last calculated" timestamp so the basis of the number is transparent and auditable.

As Clarke covers more facilities and accumulates more history, these estimates are recalibrated against what actually gets built. Priors for specific markets, cooling approaches, and AI workloads tighten over time, and updated methodology is applied through the dataset. The result is a modeling layer that becomes *more accurate* and *more useful* as coverage deepens, while always remaining explicitly labeled as an estimate.

The Road Ahead

Clarke's mission is to become the intelligence layer for global AI infrastructure: the standard framework for understanding where, when, and how AI compute capacity is being built, and what it implies for power, policy, and capital.

In the near term, Clarke will:

- Expand coverage of AI-related facilities in the U.S. and refine operator-level views.
- Deepen integration of utility, ISO/RTO, and regulatory data into CUFS.
- Evolve Clarke Standards in collaboration with operators, utilities, analysts, and policy researchers.
- Offer programmatic/API access to its intelligence.

Over time, Clarke plans to:

- Enrich regional and grid-level perspectives, including capacity-availability and constraint mapping.
- Extend coverage to additional geographies.
- Build a multi-year track record of forecast performance that customers can audit.

What doesn't change is our objective: to provide reliable, decision-grade infrastructure intelligence for the AI-energy-compute economy.

Partner early in AI Compute Intelligence

AI is shifting from a software story to an infrastructure story. The next decade of machine intelligence will be constrained not by algorithms or talent, but by power, cooling, land, and steel in the ground.

Clarke cuts through the fog. We map, verify, and forecast the physical infrastructure powering frontier AI—giving operators, investors, and policymakers the forward-looking intelligence they need to make decisions under real constraints.

AI infrastructure is being built at unprecedented scale. Clarke is the intelligence layer to understand it.

Clarke is currently building its initial Capacity Forecasts MVP and forming a small circle of operator, investor, and utility advisors. Qualified partners can request early access at clarkeindex.com

Clarke

Clarke Index © 2026 — All Rights Reserved