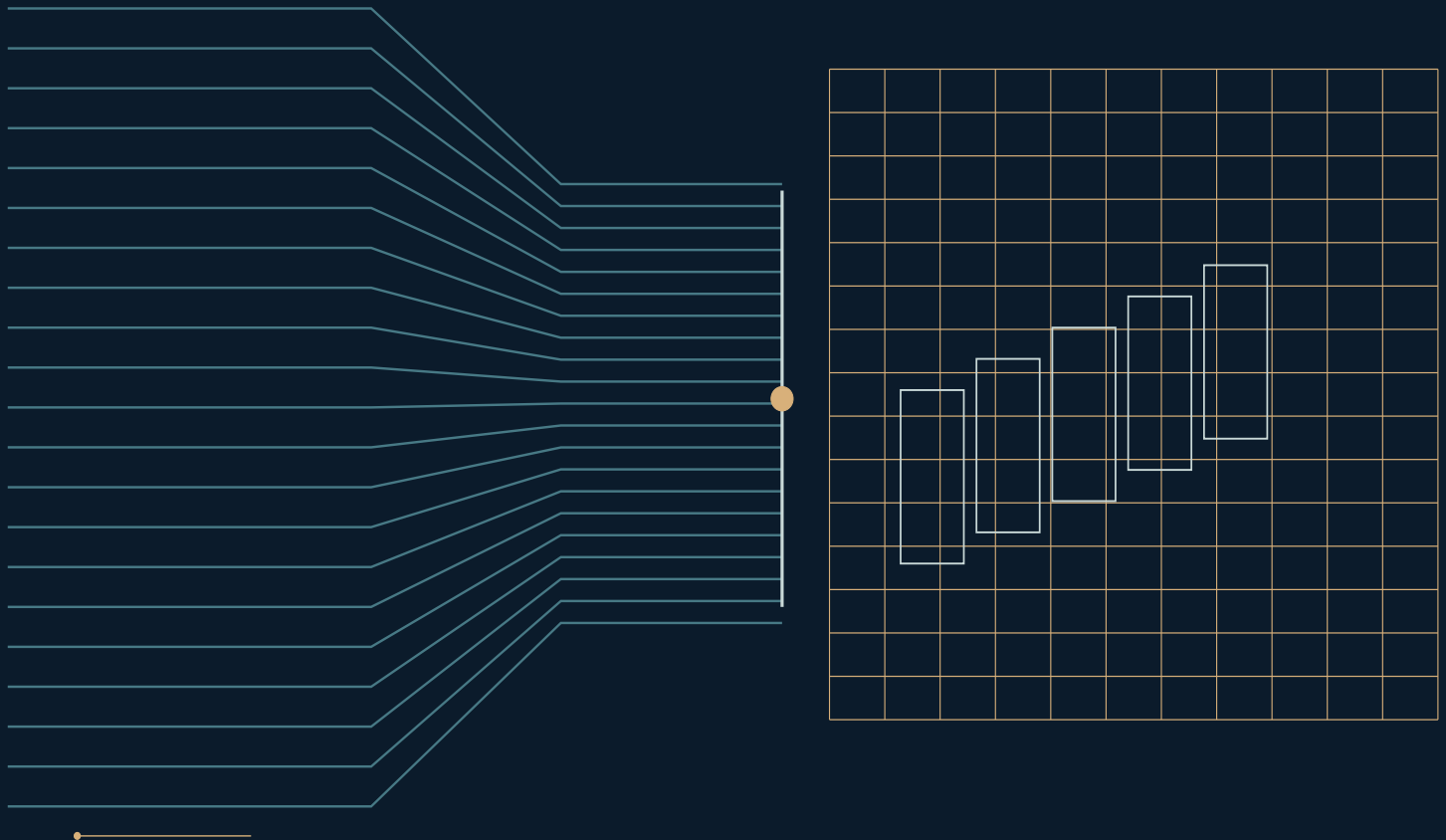




STRATEGY / OCTOBER 2026

# AI, energy and infrastructure choices

Treat compute demand, power access and flexibility as one investment decision



THE DECISION IN BRIEF

AI infrastructure creates value when useful demand, deliverable power and financial commitments arrive together. Invest against that match, not a global demand forecast alone.

- 01

Put the application roadmap, grid connection and cash commitments on one timeline. A capacity purchase can fail commercially even when long-run demand is strong.
- 02

The IEA projects global data-centre electricity use rising from 485 TWh in 2025 to 950 TWh in 2030. This all-data-centre outlook establishes scale; it does not prove that a particular site can be energised on time.
- 03

Reserve dependable capacity for evidenced demand and price flexibility for the uncertain portion. Keep energy, peak power, carbon and water boundaries separate so the investment case exposes its actual constraint.

THE MAYDEN VIEW

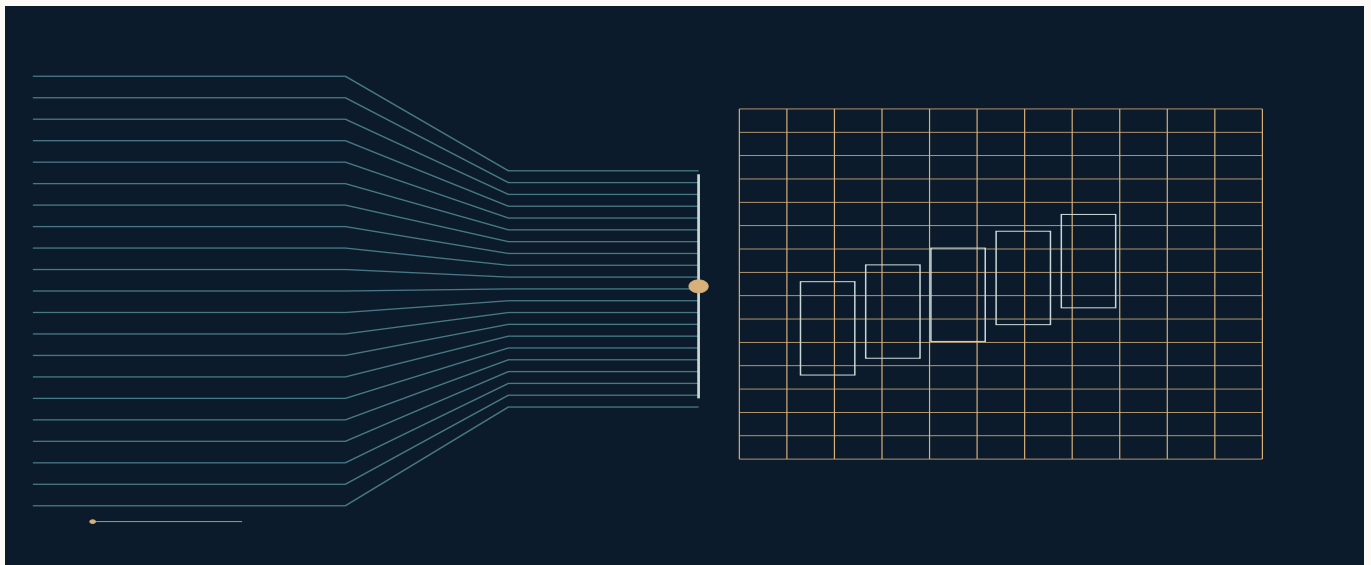
Useful demand, deliverable power and financial commitment must arrive together.

INSIDE THIS PERSPECTIVE

|                                                                       |       |
|-----------------------------------------------------------------------|-------|
| The argument at a glance                                              | 3     |
| 01 / Align the three investment clocks                                | 4     |
| Exhibit 01 / Translate useful work into energy before sizing capacity | 5     |
| 02 / Secure deliverable capacity                                      | 6-7   |
| Exhibit 02 / Global data-centre demand is projected to roughly double | 8     |
| 03 / Price local impact and flexibility                               | 9     |
| Exhibit 03 / Match the commitment to the evidence behind demand       | 10    |
| 04 / Commit against workload evidence                                 | 11-12 |
| The leadership agenda                                                 | 13    |
| Method and references                                                 | 14    |

ONE DECISION, FOUR CONNECTED QUESTIONS

# Connect compute to physical capacity



POWER AND COMPUTE GRID / CONCEPTUAL ARTWORK

## 01 Align the three investment clocks

Align demand, power and the cash commitment

## 02 Secure deliverable capacity

Secure the connection, not just the energy

## 03 Price local impact and flexibility

Separate physical impact from accounting claims

## 04 Commit against workload evidence

Translate useful work into a capacity obligation

Useful demand, deliverable power and financial commitment must arrive together.

## CHAPTER 01

## 01

# Align the three investment clocks

## Align demand, power and the cash commitment

An enterprise can approve an AI application before the grid can supply it. Hardware may arrive before power; power may arrive before useful demand. The resulting idle commitment can disappear in a steady-state cost-per-compute estimate. Leaders need one investment case that aligns application demand, deliverable capacity and the timing of cash - not three plans approved independently.

The IEA's 2026 central outlook projects global data-centre electricity demand rising from 485 TWh in 2025 to 950 TWh in 2030. <sup>[1]</sup> This is an updated projection for the entire data-centre sector. It should not be described as AI-only consumption or a guaranteed outcome. The scale is strategically relevant, but the more immediate investment question is local: which connection, substation, equipment supply chain and cooling design will support the proposed workload?

Start with the workload portfolio. Stable interactive demand, schedulable batch work and experimental use have different tolerances for delay, location and interruption. Treating all three as identical can lock uncertain applications into inflexible capacity. mayden proposes matching each class to a capacity commitment that remains useful as demand and technology change.

## Keep forecast revisions visible

Forecasts are necessary for long-lived assets, but they need version control. A demand outlook contains assumptions about adoption, hardware efficiency, utilization and the mix of tasks. Improvements in efficiency can reduce energy per task while new uses increase total demand. There is no contradiction: the denominator and the activity level both change. A board should see which assumptions drive its own exposure instead of receiving a single external number presented as certain.

Lawrence Berkeley National Laboratory's 2025 Update, published in June 2026 and hosted by the US Department of Energy, estimates 192 TWh of US consumption in 2024, or 4.7% of national electricity use, and revises earlier historical estimates. <sup>[2]</sup> That revision illustrates a wider measurement issue: even the starting point can change as information on equipment shipments and operations improves. Comparing two forecasts without checking their vintage, geography and scope can manufacture a growth story that is partly a change in estimation.

mayden recommends three internal scenarios anchored in business demand rather than arbitrary percentages: committed services, plausible expansion and an upside requiring additional evidence. For each, specify computing hours, service levels and the capacity that must be available simultaneously. Translate that into an infrastructure range using tested assumptions. Use the scenarios to identify commitments that remain sensible across different futures and decisions that should wait for a clearer demand signal.

## ILLUSTRATIVE DECISION TOOL

# Translate useful work into energy before sizing capacity

All inputs hypothetical; energy in kWh, power in kW

01

**Workload and attempts****1,000 accepted jobs; 1,250 total attempts**

Retries and failures remain inside resource use

02

**IT energy** **$1,250 \times 0.04 \text{ kWh} + 20 \text{ kWh shared IT} = 70 \text{ kWh}$** 

Assumed attempt energy plus separate shared activity; no double counting

03

**Facility energy** **$70 \text{ kWh} \times 1.4 \text{ PUE} = 98 \text{ kWh}$** 

Same facility/time boundary; assumed PUE, not a benchmark

04

**Average versus peak** **$98 \text{ kWh} / 10 \text{ hours} = 9.8 \text{ kW average}$** 

Peak capacity still needs a coincident-load assessment

**THE MANAGEMENT IMPLICATION**

**Accepted work supports energy accounting; deliverable peak capacity needs additional engineering evidence.**

Arithmetic illustration, not measured AI efficiency or recommended PUE. Shared IT excludes energy already counted per attempt. PUE must match the facility and period. No peak is inferred from average power. Sources: [1], [2]

## CHAPTER 02

## 02

# Secure deliverable capacity

## Secure the connection, not just the energy

Annual electricity consumption is measured in energy units such as MWh or TWh. A site's connection and equipment are sized around power, measured in MW or GW, with operating constraints and peaks that an annual average can hide. A procurement paper that moves between these quantities without explanation can obscure the real bottleneck. A low annual utilization rate does not necessarily remove the need for a substantial connection if demand is concentrated in particular periods.

The IEA's Electricity 2026 analysis notes that grid infrastructure can take 5-15 years to plan, permit and complete, compared with 1-3 years for data centres. <sup>[3]</sup> These are broad ranges, not project schedules. The investment implication is to verify the specific connection path before treating announced capacity as deliverable capacity. Land, an energy contract and a hardware order are not substitutes for the electrical infrastructure and permissions needed to operate.

The IEA's September 2026 report on modernising grids broadens the response beyond new construction. It examines monitoring and control, grid-enhancing technologies and AI-enhanced applications as

ways to use existing networks more effectively. <sup>[8]</sup> These are complementary routes, with different operational and regulatory requirements. A general technology opportunity does not establish spare capacity at a particular connection point.

For an AI infrastructure investment, ask the grid operator which constraint is binding and which intervention could change it. A thermal limit, a substation bottleneck and an operational security constraint do not have the same remedy. mayden proposes carrying that diagnosis into the capacity contract and project schedule. Until a site-specific assessment supports a change, potential optimization should remain an option in the investment case rather than committed deliverable power.

Gulf site choices need the same local assessment. This report does not assume that all GCC markets share the same electricity economics, network capacity or permitting regime. A site decision requires local evidence about connection timing, reliability, cooling conditions, water and contractual obligations. Ask who bears delay risk and whether the business can use alternative capacity during a shortfall. A staged commitment may preserve more value than an apparently attractive unit price tied to capacity that cannot be energized when required.

## CHAPTER 02

## 02

# Secure deliverable capacity

---

## Reduce the workload before fixing capacity

The business has more choices than build or buy. It can route different tasks to different models, reuse results where appropriate, schedule flexible work, reduce unnecessary generation and choose a service level that fits the task. These are proposed engineering and procurement options; the report does not claim a universal saving from any one of them. Each needs measurement of quality as well as resource use. A smaller model that creates more retries or review may be less efficient across the complete workflow.

Classify workloads by latency, location constraints, confidentiality, recoverability and variability. Interactive customer services may need predictable response and strict data controls. Internal batch analysis may tolerate a wider processing window. Model training,

fine-tuning and inference can have different resource profiles and scheduling options. The right unit of optimization is a successfully completed business task at the required quality, including failed attempts and downstream correction. Tokens generated are a useful engineering measure but an incomplete business outcome.

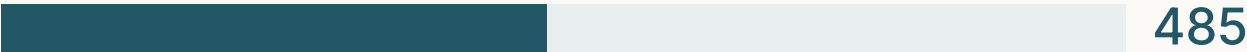
Use this classification to structure contracts. Reserve capacity for demand that is demonstrably stable, retain flexible purchasing for uncertain demand and specify portability where feasible. Test whether the application can move before relying on exit rights. Software dependencies, data-transfer costs, specialized hardware and operating expertise can make a nominally portable workload difficult to relocate. The CFO and technology leader should examine those frictions together, because they determine whether flexibility is real or merely a clause in a contract.

FORECAST COMPARISON · 2025 ESTIMATE AND 2030 PROJECTION

# Global data-centre demand is projected to roughly double

TWh per year

2025 estimate



2030 central projection



THE MANAGEMENT IMPLICATION

The global scale justifies planning, while project commitments still require local evidence.

IEA 2026 Key Questions on Energy and AI. All data centres, not AI alone. 2030 is a projection; use the same forecast vintage for both bars. Sources: [1]

## CHAPTER 03

## 03

# Price local impact and flexibility

## Separate physical impact from accounting claims

The IEA's 2025 supply analysis distinguishes the electricity physically consumed by data centres from operators' contractual energy mix. <sup>[4]</sup> That distinction should remain visible in executive reporting. Purchasing environmental attributes or signing a renewable contract is different from matching a site's demand with local supply at every hour. Both can be relevant, but they answer different questions. A report should state the accounting method and avoid implying that a contractual claim eliminates all local grid impacts.

Water requires the same care. Google's 2026 environmental reporting describes replenishment projects across watersheds and reports replenishment equivalent to 78% of its 2025 freshwater consumption. <sup>[5]</sup> This is a company-reported stewardship measure, not evidence that every facility is water-neutral or that local consumption has disappeared. mayden's broader recommendation is to examine withdrawal, consumption, source, location and season separately. A global balance can conceal pressure in a particular watershed, just as an annual energy balance can conceal hourly constraints.

The European Commission's data-centre energy-performance work establishes a reporting framework under Delegated Regulation EU 2024/1364. <sup>[6]</sup> Its relevance is the need for comparable information; specific obligations depend on scope and jurisdiction. Procurement teams should ask for operational metrics they can reconcile, with clear site and time boundaries. Avoid combining electricity, emissions and water into one unvalidated sustainability score. Keeping the measures separate makes trade-offs discussable and

prevents improvement in one dimension from silently offsetting deterioration in another.

## Price the local constraint and the delay

Global averages can understate the effect of concentrated infrastructure. Ireland's Central Statistics Office reports that data centres used 23% of the country's metered electricity in 2025, compared with 5% in 2015. <sup>[7]</sup> This is a national descriptive statistic for all data centres, not an AI-only measure and not a forecast for other countries. Its strategic lesson is that location matters: a modest global share can coexist with a material local demand on infrastructure.

A credible site case should therefore include the community and system boundary. Examine who funds network upgrades, how costs and risks are allocated, what happens during shortages and whether the site's flexibility can support the system. A promise of future jobs or digital investment does not answer these questions by itself. Nor does local opposition automatically prove a project is uneconomic. The decision needs transparent assumptions and a comparison with realistic alternatives, including phased capacity or a different location.

Stress-test the cash profile across connection, commissioning and application milestones. Name the commitments that become sunk before demand is established and the party bearing each delay. Compare the cost of waiting with the cost of committing too early. A low eventual compute price does not compensate automatically for idle equipment or a missed service launch.

ORIGINAL MAYDEN FRAMEWORK

# Match the commitment to the evidence behind demand

mayden analytical investment framework

| Demand condition            | Possible capacity posture               | Evidence before commitment                   |
|-----------------------------|-----------------------------------------|----------------------------------------------|
| Established, stable service | Reserved or owned capacity              | Measured workload and deliverable connection |
| Variable or emerging use    | Flexible external capacity              | Quality, portability and full unit cost      |
| Credible expansion option   | Conditional reservation or staged build | Trigger, expiry and cost of unused capacity  |

THE MANAGEMENT IMPLICATION

Flexibility has value only when the workload and contract can actually change.

Illustrative options, not financial advice or a measured optimal allocation. Sources: [1], [2], [3]

## CHAPTER 04

## 04

# Commit against workload evidence

## Translate useful work into a capacity obligation

A capacity reservation should begin with an account finance and engineering can reconcile. Exhibit 01 uses invented inputs solely to show the units: 1,000 accepted jobs require 1,250 attempts. At an assumed 0.04 kWh per attempt, plus 20 kWh of shared IT activity not already counted, IT energy is 70 kWh. An assumed facility power usage effectiveness of 1.4 gives 98 kWh for the same period and boundary. Over ten hours that is 9.8 kW average facility power. It says nothing by itself about the maximum simultaneous load, memory or network bottleneck. No figure is a vendor price, observed efficiency or recommended design value.

Before using this account in an investment case, replace the assumed attempt energy and shared load with measurements for the intended configuration, including failed work, evaluation and recovery without counting any activity twice. Match facility overhead to the relevant site and observation period; a portfolio average may not describe the selected service. Separately establish the workload schedule, concurrency, required availability and tested peak load. A batch job that can move within a completion window creates different capacity options from an interactive service whose

demand arrives together. State which requirement drives the proposed reservation and which demand remains flexible.

Then map that record to a delivery schedule. A facility may have a planned connection date, a separate hardware commissioning date and a later date for the application to use the capacity effectively. Show the cash commitments against each milestone and name the party responsible for delay. Test the consequence of power arriving after equipment, or equipment arriving before proven demand. The point of the exercise is to identify exposures that disappear in a steady-state unit-cost estimate, including idle assets, temporary external supply and staff hired before operations begin.

mayden's proposed decision rule is to reserve irreversible commitments for demand supported by the service case, while pricing flexibility explicitly for the uncertain portion. Compare the premium for a cancellable or transferable reservation with the cost of unused capacity and the revenue at risk from delay. These values are project-specific; this report supplies no universal optimal split. A board may rationally accept an early commitment to secure a scarce site. It should do so with the downside stated and a decision owner, rather than treating a global demand forecast as evidence that its own applications will consume the capacity.

## CHAPTER 04

## 04

# Commit against workload evidence

---

## Buy capacity and flexibility for different purposes

Waiting carries a cost where sites and connections are scarce. A staged strategy should therefore price the option to secure capacity, not assume delay is free. Compare committed capacity, conditional expansion and flexible supply against the same workload and service obligations. Each reservation needs a purpose, an exercise trigger and an account of the cost if demand does not arrive.

Bring five questions to the investment decision: which service creates the workload; what quality and availability it needs; when power, cooling and connectivity can be delivered; which environmental and operating boundaries will be

measured; and which commitments remain revisable. A missing answer becomes a condition to resolve or a risk explicitly accepted by the accountable executive.

Rebuild the largest proposed capacity commitment from the application portfolio upward. Test lower demand, delayed energisation and a different workload mix. Show the cost of unused capacity beside the cost of being late. Commit where the service and delivery evidence support the exposure; retain options where they do not. Update the decision when either demand or connection evidence changes.

## FROM INSIGHT TO ACTION

# Decisions to take forward

## PRODUCT, ENGINEERING AND FINANCE OWNERS

**Replace assumptions with measurements before treating energy estimates as a capacity case.**

Evidence to require: Accepted workload, attempts, measured IT energy and shared activity

## FACILITIES AND GRID COUNTERPART

**Distinguish deliverable peak power from average energy; price delay exposure.**

Evidence to require: Coincident-load assessment, connection evidence and commissioning schedule

## INVESTMENT COMMITTEE AND PROCUREMENT

**Commit where workload and delivery evidence support it; price options for uncertainty.**

Evidence to require: Committed demand, flexibility and lower-demand/delayed-power cases

mayden connects senior judgement, research and engineering to turn strategic choices into working capabilities.

## EVIDENCE YOU CAN INSPECT

# Research notes

## Method and boundaries

Comparison of eight opened primary energy, statistics, regulatory and company sources; cutoff 10 October 2026. Forecast vintages and observed values remain explicit; the investment framework is original mayden synthesis.

IEA for global demand, supply and grid context; DOE-hosted research for US estimates; CSO for metered national data; Commission for reporting context; Google for a clearly labelled company disclosure example.

No site engineering study, energy-price forecast, project valuation or Gulf-country capacity assessment. Data-centre demand is not AI-only demand. Company disclosures are not independent causal evidence; proposed investment gates require project-specific due diligence. The workload-to-energy exhibit uses explicitly invented inputs to demonstrate dimensional accounting; it estimates neither actual model efficiency nor peak capacity.

## Primary sources and author research

### [1] Key Questions on Energy and AI: Executive summary

International Energy Agency / 2026 edition; accessed 2026-10-10. Updated data-centre demand outlook; 485 TWh 2025 and 950 TWh 2030.

[Read the source](#)

### [2] United States Data Center Energy Usage Report: 2025 Update

Lawrence Berkeley National Laboratory; hosted by US Department of Energy / 2026-06-18 (report title: 2025 Update). PDF page 10 / printed page 8, Total Energy Use.

[Read the source](#)

### [3] Electricity 2026: Grids

International Energy Agency / 2026. Infrastructure planning and construction times.

[Read the source](#)

### [4] Energy and AI: Energy supply for AI

International Energy Agency / 2025-04-10. Electricity supply mix methodology.

[Read the source](#)

### [5] 2026 Environmental Report

Google / 2026; 2025 reporting period. Water stewardship; 78% replenishment statement.

[Read the source](#)

### [6] Energy performance of data centres

European Commission DG Energy / Current page accessed 2026-10-10. Reporting scheme and Delegated Regulation EU 2024/1364.

[Read the source](#)

### [7] Data Centres Metered Electricity Consumption 2025

Ireland Central Statistics Office / 2026-07-07. Release infographic transcript: 5% in 2015, 23% in 2025; publication date 7 July 2026.

[Read the source](#)

### [8] Modernising Grids in the Age of Electricity

International Energy Agency / 2026-09-21. Report overview; digital grid toolkit; chapter structure. General opportunities do not establish project headroom, deliverable connection date or a universal AI effect.

[Read the source](#)