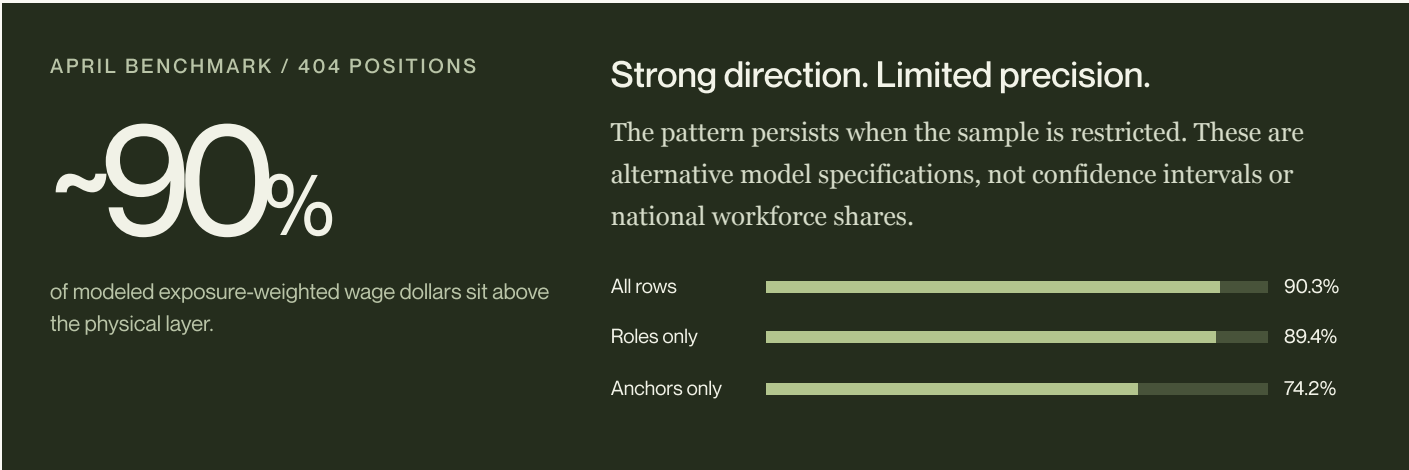


THE ENERGY DECISION STACK

The filing cabinet gets AI before the wellhead.

Energy’s early AI opportunity sits in the work that prepares a decision: reconcile the numbers, assemble the evidence, surface the exceptions. The return depends on what happens after the draft.



373 roles + 24 workflows + 7 artifacts. Overlapping sample; 60 BLS/industry employment anchors, 344 formulaic estimates, six compensation proxies. Scores and assumptions remain the April benchmark. Reproduction results ↗

What changed around the thesis

GRID ACCESS

2,061GW

Generation and storage requests at end-2025. Median request-to-operation exceeded five years for projects completed in 2025 in covered regions.

Berkeley Lab, May 2026 ↗

DATA-CENTER ELECTRICITY

485 → 950

TWh estimated in 2025 → TWh projected in 2030. The projection is approximate. Includes all data centers, not AI alone.

IEA, April 2026 ↗

ENERGY INVESTMENT

\$3.4T

Estimated globally in 2026. Decision quality matters at this scale, but no percentage improvement can be inferred from the number of scenarios run.

IEA, 2026 ↗

Start where the work recurs, the company controls preparation, and an accountable reviewer can define an accepted result.

Raj Mistry · September 7, 2026 · Original benchmark published April 2026

A PRACTICAL 30-DAY PILOT

One workflow.
Ten cases. A measurable decision.

Treasury and lender readiness, ownership and title, or regulatory preparation.
Choose one recurring output and one person accountable for accepting it.

DAYS 01–05 / BASELINE

Define accepted work.

Collect ten completed cases. Record preparation and reviewer hours, sources, errors, and acceptance criteria. Reserve five cases for a later evaluation.

DAYS 06–12 / BUILD

Make the evidence traceable.

Use the first five cases to connect source documents, numerical checks, exception handling, and draft generation. Keep a human owner of the final output.

DAYS 13–22 / EVALUATE

Test the held-out cases.

Apply the same acceptance criteria. Log added review, unsupported claims, source errors, reconciled numbers, and total costs. Keep external approval time separate.

DAYS 23–30 / DECIDE

Expand only on evidence.

Continue if accepted work uses less total human time and the economics justify it. Investigate every escaped material error. Five test cases do not establish reliability.

ILLUSTRATION / FIRST YEAR

Count the review. Count the costs.

\$1,425

Illustrative usable capacity value after costs, using the report’s editable defaults. This is not measured savings or a payroll-reduction estimate.

240 h

Net human hours released

\$10,385

Value of usable capacity

\$8,960

API, software + setup

5 people × 160 annual prep hours × 40% removed × (1 – 25% added review) = 240 hours. Value 60% of that capacity at \$150,000 ÷ 2,080 hours. Subtract 120M annual tokens × \$8/M, \$3,000 software, and \$5,000 setup.

What this edition corrects. A queue snapshot is not a construction pipeline. The historical 13% queue completion figure is a different cohort from today’s active requests. The former 60–70% paperwork-delay estimate and value/token multiples are retired. More scenarios do not establish better capital returns. The synthetic proof page is a design illustration, not an evaluated model run.

Use the working materials

[Editable economics ↗](#) [Pilot scorecard ↓](#) [Primary sources ↗](#) [Changes & corrections ↗](#)

Benchmark exposure is not observed productivity. Released time has value when it is productively used or spending is avoided. Higher investment returns, faster external approvals, and fewer losses require separate evidence.