

THREE PEOPLE, ONE BUILDING

The AI Mirror

Riches, Ruin and the Indispensability Trap

Three people look at the same data center and see abundance, dispossession and a trap. All three are correct inside their own frame. The disagreement was never about whether the technology works. It is about who ends up owning it.

Contents

The two loudest readings of artificial intelligence are not technological disagreements. They are distributional ones. Both assume the technology works, both assume enormous productivity gains, and they differ entirely on who ends up owning it. This edition adds a third reading drawn from the author's forthcoming book, and grounds all three in a source-locked evidence base: thirteen programmatic figures, every value traceable to a named primary document.

I	Plate I: Three people, one building	03
01	Three People, One Building	04
02	Reflection One: The Case for AI Utopia, Taken Seriously	06
03	Reflection Two: AI Serfdom, and the Question Sanders Keeps Asking	10
II	Plate II: The scarce layer	14
04	Reflection Three: The Indispensability Trap	15
05	The Reflections Are Already Colliding	20
III	Plate III: Light divided	23
06	Who Owns the Abundance?	24
07	AI Jobs and the Missing Timetable	26
08	Geopolitics: When Indispensability Becomes National Policy	28
IV	Plate IV: The same surface	30
09	The Mirror, One Last Time	31
·	Sources and Notes	33
·	Colophon and about the author	35

FOUR VERIFIED NUMBERS THIS REPORT RESTS ON

\$0.011/kWh

VIRGINIA DATA-CENTER ELECTRICITY TAX

First U.S. tax on the electricity data centers consume. July 1, 2026 to July 1, 2028, with no more than \$600 million deposited to the general fund in a fiscal year and excess returned pro rata. Statute, not an estimate.

19%

ENTRY-POINT EMPLOYMENT GAP

Employment among 22 to 25 year olds in highly AI-exposed occupations, against where it would have been had it kept pace with similarly aged workers in less-exposed roles. August 2026 revision. The authors call these early, descriptive indicators rather than causal estimates, report no widespread economy-wide displacement, and note the patterns attenuate when controlling for education.

71%

AMERICANS OPPOSING A LOCAL AI DATA CENTER

Including 48% strongly opposed. Telephone interviews conducted March 2 to 18, 2026 with a random sample of 1,000 adults, margin of error plus or minus 4 points.

39%

ENTERPRISES RUNNING MORE THAN TEN AGENTS

Of executives at global enterprises that already had generative-AI deployments, 39% said their company had launched more than ten AI agents. A vendor-commissioned survey of 3,466 senior leaders across 24 countries: a vanguard sample rather than the economy, and Google does not publish field dates, question wording, sampling method, weighting, margin of error or mode.

Thirteen figures and twelve original art plates were produced for this edition. Every figure is built programmatically from the sources listed in Sources and Notes, and each one is labeled either source-locked evidence or analytical construction. Where a claim could not be verified against a primary document, the figure was reframed rather than filled in.



PLATE I · THREE PEOPLE, ONE BUILDING

One surface, three reflections

The same beam of light returns gold from one face, coral from the next and cold cyan from the third. Nothing about the surface changes. What changes is where the observer is standing.

ORIGINAL ART GENERATED FOR THIS REPORT

01 THE OPENING SCENE

Three People, One Building

Three people stand in front of one building in Northern Virginia. All three are looking at the same concrete, all three are correct inside their own frame, and this report is about why.

In June, Virginia Governor Abigail Spanberger signed a budget containing something no state had ever enacted: a tax on the electricity data centers consume. Eleven-tenths of a cent per kilowatt-hour, levied from July 1, 2026 until July 1, 2028, with no more than \$600 million deposited to the general fund in a fiscal year and excess collections returned pro rata. The first collection occurs in September 2026. It was the compromise that ended months of trench warfare in Richmond, where Senate Finance Chair L. Louise Lucas had spent the session refusing to send the governor any budget preserving the sales-tax exemption the Department of Taxation reported at \$1.941 billion in fiscal 2025, based on company reports to VEDP that were not independently validated.

Virginia spent two decades building Data Center Alley—209 completed data centers and 43 under construction in Loudoun County as of January 1, 2026, according to county property statistics—a meaningful share of the world's internet traffic. And after all that cultivation, it became the first state in America to tell the industry it had courted that the power was no longer free.

Now picture three people standing in front of one of those buildings.

The first sees the physical substrate of the most important scientific instrument humanity has ever constructed—a machine that will compress decades of biomedical research into years, tutor every child on earth, and eventually free most people from the requirement to sell their labor in order to eat.

The second sees a windowless warehouse that will employ a few dozen people permanently, consume the electricity of a small city, raise the utility bills of every household on the same grid, and generate returns that flow almost entirely to shareholders who live somewhere else.

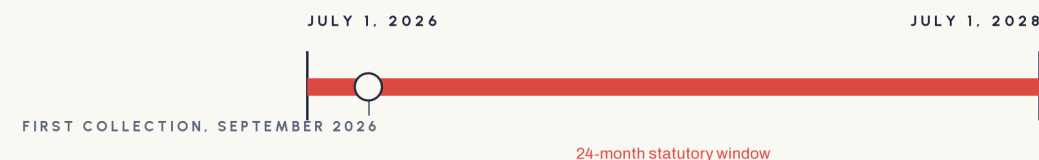
The third—and this is where I sit most days—sees something more unsettling than either: a building that is genuinely indispensable right now, whose owners are betting tens of billions of dollars that it will remain indispensable in exactly its present form. And she has read enough technology history to know how that particular bet usually turns out.

All three are looking at the same concrete. All three are, in their own frame, correct. This report is about why.

FIGURE 01 · SOURCE-LOCKED EVIDENCE

Virginia writes a temporary, capped bill for the buildout

\$0.011 per kilowatt-hour
ELEVEN-TENTHS OF A CENT, LEVIED ON CONSUMPTION



\$600M

ANNUAL GENERAL-FUND CEILING

No more than \$600,000,000 of revenues may be deposited to the general fund in a fiscal year. A cap written into law, not money collected.

PRO RATA

EXCESS IS RETURNED

Collections above the ceiling go to a special nonreverting fund administered by the State Corporation Commission and are refunded as each operator's pro rata share.

SUNSET

IT EXPIRES BY DEFAULT

The tax expires July 1, 2028 unless re-enacted. The equipment sales-tax exemption was preserved; only the power was taxed.

Source: Virginia HB 30, Item 3-5.24 (2026-2028 biennial budget); Office of the Governor of Virginia, June 30, 2026.

Statutory mechanics only. The rate, the window, the cap and the pro rata refund are written into HB 30; the cap is a ceiling on general-fund deposits and is not revenue received.

I want to be direct about the premise. I do not think AI is hype. I think it is the most consequential general-purpose technology since electrification, and I think the people who

dismiss it as autocomplete are going to look as foolish in ten years as the people who dismissed the commercial internet in 1995. That is the starting point, not the argument.

The argument is about what happens next—and specifically about the fact that the two loudest visions of the AI future, the utopian and the dystopian, are not technological disagreements at all. They are distributional ones. Both assume the technology works. They differ entirely on who ends up owning it.

EDITOR'S NOTE

The Indispensability Trap is also the title and central thesis of Alan Shimel's forthcoming book, *The Indispensability Trap: How Becoming Essential Caps the Fortune, From the Railroads to AI*, published September 2026.

And both, I will argue, are less likely to arrive intact than a third outcome, one that comes from the framework in my forthcoming book. I call it the Indispensability Trap, and it explains why the chokepoints everyone is currently pricing as permanent have a strong historical tendency to stop being chokepoints right around the time everyone agrees they never will.

THE FIRST OBSERVER Abundance The physical substrate of the most important scientific instrument humanity has ever built, compressing decades of research into years.	THE SECOND OBSERVER Dispossession A windowless warehouse employing a few dozen people, consuming the electricity of a small city, returning value to shareholders who live somewhere else.	THE THIRD OBSERVER The Indispensability Trap A building that is genuinely indispensable right now, whose owners are betting tens of billions that it stays indispensable in exactly its present form.
---	---	--

HOW TO READ THE EVIDENCE IN THIS EDITION

Source-locked evidence • Every value on the graphic is traceable to a named primary document. • Fiscal periods, sample sizes and field dates are printed on the figure itself. • Ten of the thirteen figures carry this label.	Analytical construction • A diagram of the argument, carrying no quantities at all. • Ordered by capability, role or category rather than by date. • Figures 02, 07 and 11 carry this label.	What is never charted • Any number this report could not verify against a primary source. • Any aggregate assembled from periods that are not coterminous. • Where a claim failed verification the figure was reframed, not filled in.
---	---	---

02 REFLECTION ONE

Reflection One: The Case for AI Utopia, Taken Seriously

Give the optimists their strongest version rather than a caricature. The serious case for abundance is a specific, falsifiable claim about rate, and much of it does not require artificial general intelligence.

Start with the optimists, and give them their strongest version rather than a caricature. Dario Amodei's "Machines of Loving Grace" remains the most fully developed articulation of the case, and its central claim is not vague enthusiasm. It is a specific, falsifiable prediction about rate. Amodei argues that AI-enabled biology could deliver in five to ten years the progress human biologists would otherwise have achieved over fifty to a hundred—what he calls the “compressed 21st century.” From that compression he derives the rest: reliable prevention of most infectious disease, elimination of most cancer, effective treatment for genetic illness, prevention of Alzheimer's, and a plausible doubling of human lifespan.

What makes the argument serious rather than credulous is that Amodei explicitly rejects the more extreme version of his own thesis. He does not think you can get a century of progress in a single year: experiments have latency, hardware has to be built, and some things cannot be deduced logically—they have to be run. He is arguing for massive parallelism on top of an irreducible delay, not the abolition of physics.

Jensen Huang makes the infrastructure builder's version of the same case, and he has been consistent about it. AI, he wrote in March, is not a clever app or a single model; it is “essential infrastructure. Every company will use it. Every nation will build it.” He compares it to electricity and the internet as a claim about category, not as a flourish. Infrastructure implies capex, skilled trades, land, power, and long-term national intent—something you build rather than something you launch.

versal basic income, a technology dividend, broadened equity ownership, a shorter workweek, or some braid of all four could plausibly eliminate material poverty. People would still work—build companies, practice medicine, write software, make art—but because they chose to, not because the alternative was destitution.

That is AI Utopia. It is not a fantasy; it is an extrapolation. The question worth asking is which parts of it require breakthroughs we do not have.

The capability ladder, and where each rung actually gets you

It helps to be precise about what we are talking about, because the utopian case quietly borrows capability from rungs we have not reached.

5 to 10

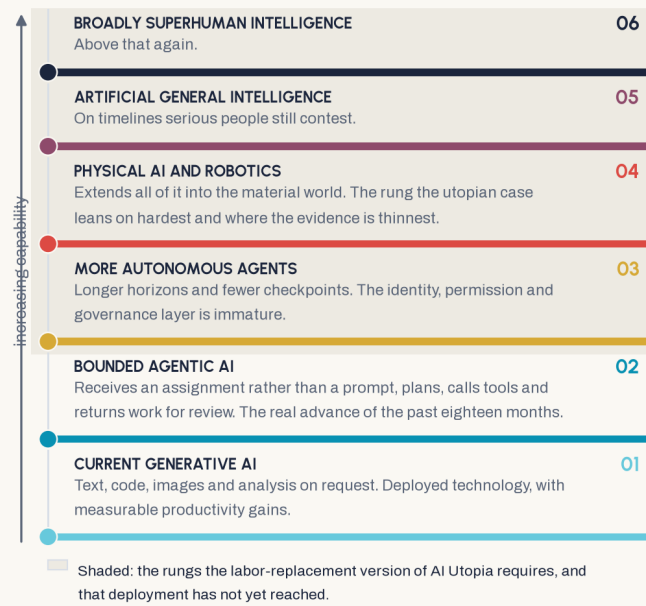
YEARS. AGAINST FIFTY TO A HUNDRED

The compressed 21st century: Amodei's claim is that AI-enabled biology could deliver in five to ten years the progress human biologists would otherwise have achieved over fifty to a hundred. A specific, falsifiable prediction about rate, and the load-bearing assumption of the whole utopian case.

If both men are directionally right, the economic implication is enormous. Productivity gains of that magnitude, sustained over a decade or two, would generate enough surplus to loosen the connection between labor and survival. Uni-

FIGURE 02 · ANALYTICAL CONSTRUCTION

Six rungs, and where the utopian case borrows capability



Ordered by capability, not by date. An analytical construction from the report text; it carries no quantities and no timeline. A surprising amount of AI Utopia does not require the shaded rungs. The labor-income link dissolving does.

Ordered by capability rather than by date. The shaded rungs are the ones the labor-replacement version of AI Utopia depends on and that the deployment record does not yet supply.

Current generative AI produces text, code, images, and analysis on request, and is already delivering measurable productivity gains across software development, support, marketing, legal review, and research synthesis. Deployed technology, not a promise.

Today's bounded agentic AI is the real advance of the past eighteen months. An agent receives an assignment rather than a prompt, plans a sequence, calls tools, manipulates files, changes code, reconciles records, investigates incidents, and returns work for review. The boundaries are genuine—permissions, scope, time horizon, human checkpoints—but the shift from "produce content" to "complete a task" is a change in kind.

More autonomous agents would extend the horizon and reduce the checkpoints. This is where most enterprise anxiety lives, reasonably: the identity, permission, evaluation, and governance infrastructure needed to safely delegate multi-day work to a machine is immature.

Physical AI and robotics extend all of it into the material world. This is the rung the utopian case leans on hardest and where the evidence is thinnest.

Artificial general intelligence and **broadly superhuman intelligence** sit above that, on timelines serious people still contest.

Here is the useful part of the exercise. A surprising amount of AI Utopia does not require AGI.

AI Utopia is not a fantasy. It is an extrapolation.

Universal access to expert-level medical triage, legal explanation, and tutoring is largely a deployment and cost problem, not a capability problem. Dramatic acceleration of drug discovery, materials science, and protein engineering is already underway with specialized models that are nowhere near general. Substantial productivity gains across knowledge work are happening now with bounded agents. None of that needs a machine that can do everything a human can do.

What *does* require capability we do not have is the part of the vision that matters most for the distributional question: replacing enough human labor across enough of the economy that the labor-income link genuinely dissolves. That requires physical AI at scale, and physical AI at scale is not close.

WHAT THE LADDER SETTLES

A great deal of AI Utopia does not need AGI

- **Already a deployment problem, not a capability problem.** Universal access to expert-level medical triage, legal explanation and tutoring.
- **Already underway with narrow models.** Drug discovery, materials science and protein engineering.
- **Already happening with bounded agents.** Substantial productivity gains across knowledge work.
- **Still missing.** Replacing enough human labor across enough of the economy that the labor-income link dissolves. That needs physical AI at scale.

The robot reality check

The humanoid robotics sector has spent two years generating spectacular video and modest deployment, and the verified mid-2026 picture looks nothing like the announced one. Agility Robotics has a handful of Digit units in commercial operation at Toyota after a year-long pilot. Figure AI has real deployments at BMW: after the Figure 02 pilot, BMW said Figure 03 would start at Spartanburg for a sequencing

use case in logistics; Figure’s BotQ facility is targeting capacity of up to roughly 12,000 units a year, not achieved output. Unitree’s first-quarter **adjusted** profit fell more than 52% as R&D and marketing spending rose. Tesla’s Optimus—the program most often cited as evidence of imminent labor replacement—had not begun Fremont production as of mid-July; production was expected later in the year, but no verified unit-output count had been disclosed.

FIGURE 03 · SOURCE-LOCKED EVIDENCE

The humanoid deployment record, as disclosed

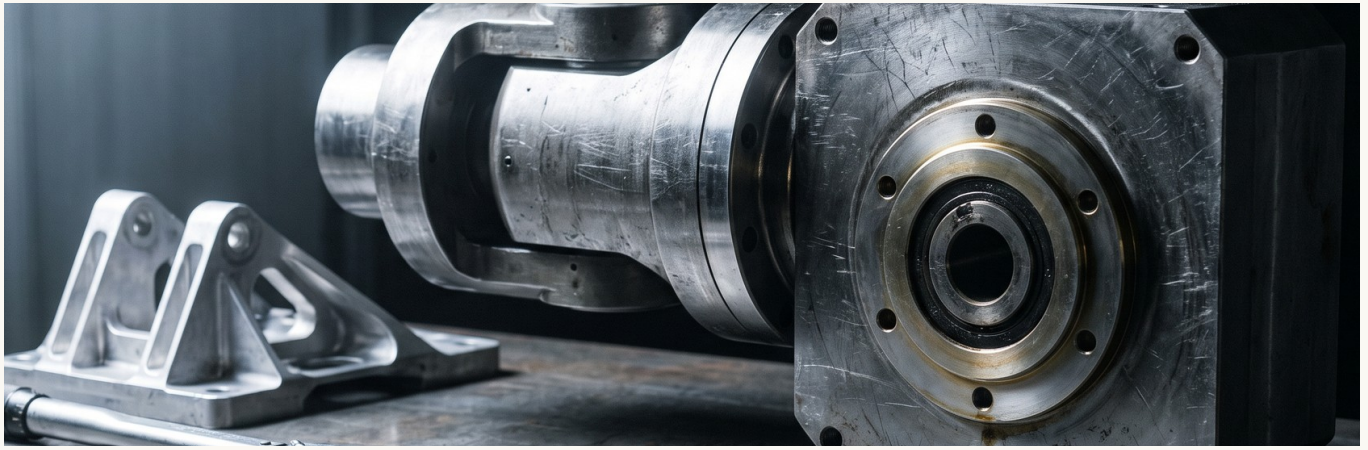
PROGRAMME	WHAT THE PRIMARY SOURCE STATES	VERIFIED STATUS
 FIGURE AI AT BMW	Figure 02 supported production of more than 30,000 BMW X3 vehicles over ten months. Figure 03 now starts at Spartanburg for a sequencing use case in logistics. Neither company discloses a robot count.	Deployed, unit count n.a.
 FIGURE AI BOTQ	BotQ is targeting capacity of up to roughly 12,000 units a year. That is a target for the facility, not achieved output.	Target, not output
 AGILITY ROBOTICS	A robots-as-a-service agreement with Toyota Motor Manufacturing Canada followed a successful pilot. Digit has moved over 100,000 totes at GXO’s Flowery Branch facility.	Commercial, small fleet
 UNITREE	First-quarter adjusted profit fell more than 52% as spending on research and development and on marketing rose.	Scaling, margin pressure
 TESLA OPTIMUS	Fremont production had not begun as of mid-July; production was expected later in the year. No verified unit-output count has been disclosed.	Pre-production
Independent context: NIST proposes the first standardized performance benchmark for humanoid robots, noting that the DARPA Robotics Challenge was the last time humanoid robot performance was rigorously measured between robots. The benchmark is a proposal, not a published standard.		

Source: BMW Group PressClub, June 25, 2026; Figure AI, BotQ high-volume manufacturing facility; Agility Robotics, Toyota Motor Manufacturing Canada agreement, February 19, 2026; CNBC, August 14, 2026; NIST, humanoid robot baseline performance benchmark (proposal).

Company and agency primaries only. Targets are labeled as targets; no unit output count is disclosed for Figure at BMW or for Optimus, and NIST’s benchmark is a proposal rather than a published standard.

NVIDIA’s GTC 2026 keynote was candid about why this is hard: “the real world is massively diverse, unpredictable, full of edge cases.” That is not a marketing problem. A humanoid robot is a ten-thousand-component electromechanical system that must be reliable, safe, certifiable, serviceable, and cheap all at once—and Tesla’s own Model 3 ramp took eighteen months of what Musk called production hell,

for a product category with a century of manufacturing precedent behind it. None of this means physical AI fails. It means the labor-replacement timeline AI Utopia depends on is measured in decades for most manual work, not years. That gap matters enormously, because the dystopian case does not require physical AI at all.



A different clock. Robots need sensors, actuators, power, dexterity, servicing, manufacturing scale, safety certification and acceptable liability. The real world is massively diverse, unpredictable and full of edge cases. Original art for this report.

03 REFLECTION TWO

Reflection Two: AI Serfdom, and the Question Sanders Keeps Asking

Turn the mirror. In this reflection the technology works exactly as the optimists predict and the social outcome is a catastrophe anyway, because the gains flow to the owners of the capital stock.



The rung that manufactures seniors. Junior work is not merely cheap production. It is training. Remove the first rungs without building a new apprenticeship and the eventual shortage is experienced people. Original art for this report.

Turn the mirror. In this reflection the technology succeeds exactly as the optimists predict. The models get better. The agents get more capable. The productivity gains are real and large. And the social outcome is a catastrophe anyway, because the gains flow almost entirely to the owners of the capital stock while the costs land on people who own nothing.

Bernie Sanders has reduced this to a single question that belongs at the center of any honest discussion: who bene-

fits? When he introduced the American AI Sovereign Wealth Fund Act in June, his justification was not primarily about jobs but about provenance. “AI was not created out of thin air,” he said—it rests on the accumulated knowledge and creative work of millions of people, much of it publicly funded. If a public resource generates private wealth, the public has a claim. You can think the 50% stock levy unworkable—I will get to why serious people do—and still recognize the question as the correct one.

Who benefits?

THE QUESTION AT THE CENTRE OF ANY HONEST DISCUSSION

What is actually established

Let me separate what we know from what we are guessing, because this is where most AI labor commentary falls apart.

Established: Employers are citing AI as a layoff reason at rapidly increasing rates. Challenger, Gray & Christmas recorded 54,836 U.S. job cuts attributed to AI in all of 2025; through June 2026 the figure was 101,743—nearly double the prior year’s total in half the time—with AI leading all stated reasons for several consecutive months. Technology accounts for nearly a third of all announced cuts this year—not a third of the AI-attributed cuts.

With a caveat that matters: Challenger measures what companies say, not what is verifiably true. A company claim-

ing AI as the reason is not the same as AI being the reason. The extent to which AI is to blame for layoffs remains contested, and critics have accused firms of “AI washing.” Some of this is genuine substitution; some is cover for ordinary cost discipline, offshoring, or over-hiring correction, dressed in a narrative investors currently reward. Anyone citing these numbers without that caveat is selling something.

Beginning to appear, with the strongest evidence base: The Stanford Digital Economy Lab’s “Canaries in the Coal Mine” work uses ADP payroll data. ADP provides payroll services to firms employing more than 26 million U.S. workers; the paper’s balanced monthly analysis sample contains between 3.5 million and 5 million employees. In the August 2026 revision, employment among workers aged

22–25 in highly AI-exposed occupations sat about 19% below where it would be had it kept pace with similarly aged workers in less-exposed roles. The authors call these early, descriptive indicators rather than causal estimates, and say the patterns attenuate when controlling for education. In levels, that cohort's employment in the two most exposed quintiles fell roughly 11% between November 2022 and June 2026, while the same age group in the three least-exposed quintiles grew about 10%.

The authors are scrupulous about what this does and does not show. Their own headline says it plainly: no widespread, economy-wide displacement. They tested the obvi-

ous alternative explanation—interest rates—and found AI-exposed occupations are, if anything, *less* rate-sensitive than average. The authors also caution that some differential trends are visible before the widespread use of generative AI and that the gaps are more pronounced in the ADP analysis sample than in national-survey benchmarks.

That is what careful evidence looks like. Not "AI is destroying jobs," but something narrower and harder to dismiss: something is happening to young people at the entry point of exposed occupations, it has persisted for nearly four years, and the obvious alternatives do not account for it.

FIGURE 04 · SOURCE-LOCKED EVIDENCE

Three tiers of displacement evidence, graded

ESTABLISHED	BEGINNING TO APPEAR	EXPOSURE, NOT DESTINY
What employers say	Strongest evidence base	Scenario analysis
54,836 U.S. job cuts attributed to AI in all of 2025	19% shortfall for 22 to 25 year olds in highly AI-exposed occupations against similarly aged peers in less-exposed roles, August 2026 revision	40% of global employment exposed to AI
101,743 attributed to AI through June 2026, nearly double the prior year's total in half the time	-11% / +10% levels for the exposed cohort in the two most exposed quintiles against the same age group in the three least exposed, November 2022 to June 2026	60% in advanced economies
one third roughly the share of all announced cuts this year coming from technology, which is not a third of the AI-attributed cuts	3.5–5M employees in the balanced monthly ADP analysis sample	a choice exposure can mean augmentation or displacement, and which one it becomes depends mostly on how employers choose to deploy
Challenger measures the reason a company states, not a verified cause. Critics have accused firms of AI washing.	The authors call these early, descriptive indicators rather than causal estimates, report no widespread economy-wide displacement, and note the patterns attenuate when controlling for education.	An IMF staff discussion note of January 2024, restated in January 2026. Exposure measures overlap. It is not a count of jobs lost.

Source: Challenger, Gray & Christmas, 2025 year-end report; Challenger, Gray & Christmas, June 2026 report; Stanford Digital Economy Lab, August 12, 2026 summary; IMF Staff Discussion Note SDN/2024/001, January 2024.

The tiers are the manuscript's own grading. Challenger measures stated reasons, not verified causes. The Stanford figures are the August 2026 revision and are described by its authors as early, descriptive indicators rather than causal estimates. IMF exposure is not displacement.

Why the entry point is the dangerous part

The career ladder is the mechanism by which inexperienced people become experienced professionals. Junior analysts build models badly, get corrected, and become senior analysts. Junior developers write mediocre code, get reviewed, and become architects. First-year associates grind through document review and learn what actually matters in a case.

Agentic AI is disproportionately good at precisely that tier of work: structured, reviewable, well-specified, high-volume. An organization that replaces its junior tier with agents has no problem this year. It has a problem in eight years, when it discovers it has no senior people, because it stopped manufacturing them.

That is a slower and less visible failure than mass unemployment, and in some ways worse, because no single

decision-maker is responsible for it and no quarterly metric captures it. The IMF's January 2024 staff discussion note estimates that almost 40% of global employment is exposed to AI, rising to about 60% in advanced economies—but exposure is not destiny. It can mean augmentation or displacement, and which one it becomes depends mostly on how employers choose to deploy. That is a choice, not a forecast.

WHAT THE ENTRY-POINT SQUEEZE LOOKS LIKE

A failure nobody is accountable for

- Agentic AI is disproportionately good at structured, reviewable, well-specified, high-volume work.
- That is exactly the tier where inexperienced people become experienced professionals.
- An organization that replaces its junior tier has no problem this year and a serious problem in eight.
- No single decision-maker owns the choice, and no quarterly metric captures it.

The ownership problem underneath the jobs problem

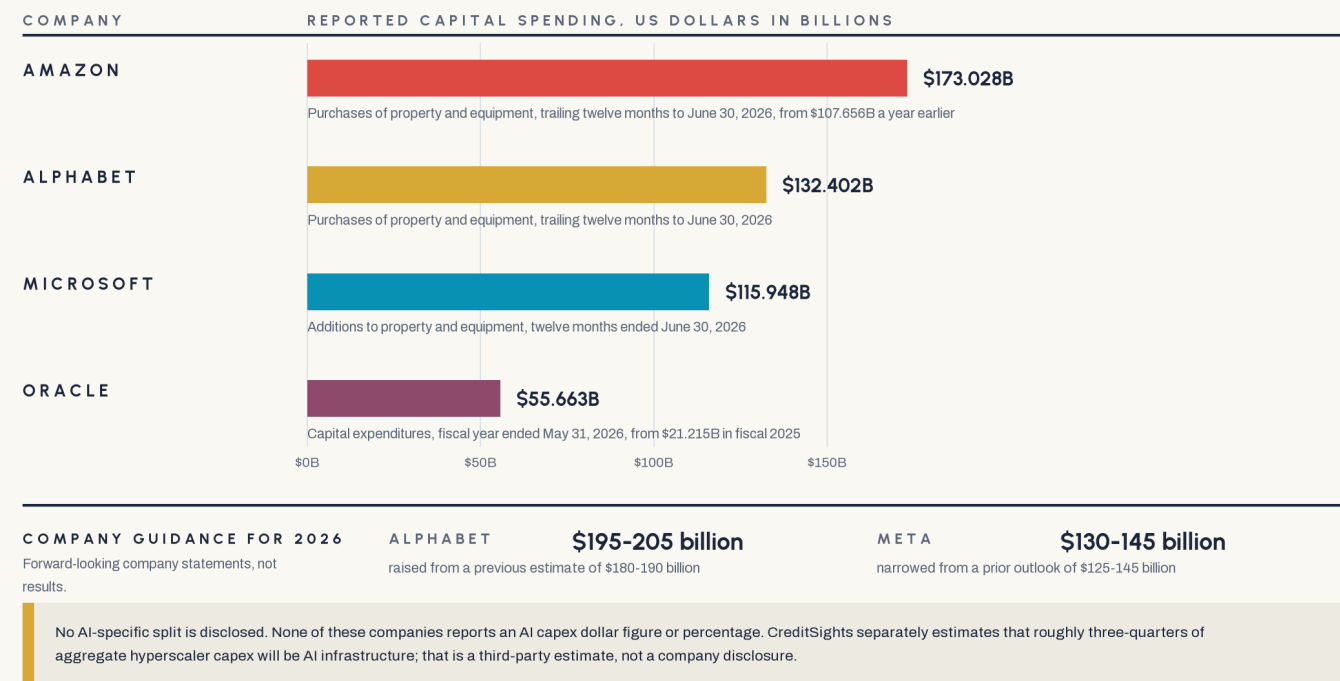
Even if the aggregate job numbers stay fine—and they might—the distributional question does not go away.

Larry Fink put it about as clearly as anyone in his 2026 annual chairman's letter, warning that "Now AI threatens to

repeat that pattern at an even larger scale—concentrating wealth among the companies and investors positioned to capture it." That is not Bernie Sanders. That is the CEO of the world's largest asset manager, and he has no ideological reason to say it.

The mechanism is straightforward. Recent disclosures from Amazon, Alphabet, Meta, Microsoft, and Oracle show an extraordinary buildout, but none of the companies discloses an AI-specific capex percentage. Four-company guidance cited in financial reporting exceeds \$600 billion for 2026; CreditSights separately estimates that roughly three-quarters of aggregate hyperscaler capex will be AI infrastructure. That money flows to a small number of chip designers, foundries, equipment makers, power developers and construction firms. The productivity gains it enables show up as margin at the deploying companies. The returns accrue to equity holders. And U.S. equity ownership is not broadly distributed.

FIGURE 05 · SOURCE-LOCKED EVIDENCE

Where the capital is going, on each company's own clock

Fiscal periods differ and are printed on each row; these values are not a comparable aggregate. Source: Microsoft FY2026 Q4 results; Alphabet Q2 2026 exhibit 99.1 filed with the SEC; Alphabet Q2 2026 earnings call, July 22, 2026; Amazon Q2 2026 results; Oracle Q4 and FY2026 results; Meta Q2 2026 results.

Filing-grade figures with their exact fiscal periods printed. These are not a comparable aggregate: Microsoft's fiscal year ends June 30 and Oracle's ends May 31. None of the five companies discloses an AI-specific capex split.

Sanders' co-sponsor Alexandria Ocasio-Cortez put the political stakes more bluntly still, closing her House version of their data center moratorium bill with "We must choose humanity over profit." You can find the rhetoric excessive. The arithmetic underneath it is not.

Why UBI is not a sufficient answer

Here is where I part ways with a lot of the optimistic policy commentary. Universal basic income may well become necessary—if displacement is broad enough, some floor

under everyone is the humane and probably the politically inevitable response. But income is not ownership, and a check is not a stake.

Betsey Stevenson has articulated the reason better than anyone I have read. Writing for the NBER volume on transformative AI, her argument is that the real fears in a post-AGI world are not about output: the material gains may arrive in a pattern that leaves many behind, and even where incomes rise, the institutions and norms supplying purpose, status, and community may fail to keep pace. That is the bridging argument that carries *me* from diagnosis to policy: AI's benefits should not be treated as an inheritance reserved for those who already own the productive assets.

That framing is the hinge of this entire report, because it converts redistribution from charity into inheritance. It is also, notably, where Sanders lands from the left and where Vinod Khosla—a venture capitalist, not a socialist—lands when he proposes the government take a 10% stake in every public company and pool it for the public. Khosla, speaking at TechCrunch Disrupt in October 2025, proposed that government “take 10% of every corporation and put it in [a] national pool for the people,” and said that “sharing the wealth of AI is a really, really big need.” When Bernie Sanders and a Sand Hill Road billionaire arrive at structur-

ally similar proposals from opposite poles, something real is being detected.

TWO PROPOSALS. OPPOSITE POLES

Where the left and Sand Hill Road converge

Sanders, introduced June 18, 2026. A one-time 50% tax paid in stock by companies with more than \$200 million in annual AI revenue, deposited into a Treasury fund the sponsor estimates at roughly \$7 trillion, governed by a seven-member commission and distributing about 5% of value annually.

Khosla, October 2025. Government takes 10% of every corporation and pools it for the people. Sharing the wealth of AI, in his words, is a really, really big need.

A basic income without ownership produces subsistence without independence. It makes people clients of an institution rather than participants in an economy. Serfdom is, if anything, too generous an analogy—the medieval lord at least needed his peasants' labor, and the distinguishing feature of this scenario is that the people at the top may not need the people below them for anything at all. That has never been a stable arrangement. Extreme concentration of wealth and political power, combined with a large population whose economic function has evaporated, is the standing precondition for populism, expropriation, and worse.

Income is not ownership, and a check is not a stake.

I do not think mass unemployment is inevitable. I think the entry-level squeeze is real and measurable, the aggregate displacement case remains scenario analysis rather than

observed fact, and the ownership concentration is happening right now regardless of how the employment numbers resolve.



PLATE II · THE SCARCE LAYER

What looks indispensable becomes one layer among many

Quantization, distillation, sparsity, better serving stacks and specialized silicon move a growing share of everyday work outward. The enabling layer does not disappear. It becomes less singular.

ORIGINAL ART GENERATED FOR THIS REPORT

04 REFLECTION THREE

Reflection Three: The Indispensability Trap

The framework from the author's forthcoming book. Once a provider becomes essential, indispensability stops being purely a source of strength and becomes a systemic risk, eventually including for the provider.

The Indispensability Trap is the framework I develop in my forthcoming book, and it describes a pattern I have watched play out repeatedly across four decades in this industry. The core observation is this: once a technology provider becomes essential to the operation of an industry, a

government, or an economy, its indispensability stops being purely a source of strength. It becomes a systemic risk—for the customers, for the market, and eventually for the provider itself.

The pattern runs in a predictable sequence.

THE PATTERN, IN FULL

The Indispensability Trap runs in a predictable sequence

The seven steps below are the author's own wording, printed once. They are the spine of the framework developed in the forthcoming book, and every argument in this section refers back to them.

STEP ONE

A company, network, platform, or infrastructure layer becomes essential during a period of scarcity.

STEP FOUR

Engineers do what engineers do. They improve efficiency, reduce the resource intensity of the workload, and invent substitutes. The scarcity that created the premium becomes the target of the industry's best problem-solvers.

STEP SEVEN

Value migrates toward the applications, businesses, and activities the infrastructure enables, rather than remaining with the infrastructure itself.

STEP TWO

Demand exceeds supply. Customers reorganize their operations around securing access. Investors begin treating the chokepoint's pricing power as a permanent structural feature rather than a temporary condition.

STEP FIVE

Standards emerge and reduce dependence. Capability migrates into cheaper, smaller, or more distributed forms.

STEP THREE

That success attracts capital, competitors, and political attention. All three arrive together, and all three are consequences of the same thing.

STEP SIX

Here is the counterintuitive part: the *technology* becomes more indispensable while some of its early *providers* become less singular. Ubiquity and pricing power move in opposite directions.

The precedents behind the pattern are railroads, electrification, the telephone network, the late-1990s fiber build and cloud infrastructure. In every case the technology mattered more over time; in most cases the companies that dominated the scarce period captured a shrinking share of the value.

The precedents are not decoration; they are the evidence base. Railroads were genuinely indispensable to nineteenth-century commerce and produced some of the most spectacular capital destruction in financial history. Electricity became so essential we regulated it as a utility and capped its providers' returns—while the value electrification created accrued overwhelmingly to the factories and industries it enabled. Telephone networks were indispensable enough to justify a legal monopoly, then indispensable enough to justify dismantling it. The late-1990s fiber build was right about

demand and catastrophically wrong about its timing and price; the capacity got used, mostly by companies that bought it out of bankruptcy for pennies. Cloud infrastructure remains indispensable, and cloud margins have been under structural pressure for a decade as the workloads commoditized.

In every case the technology mattered more over time. In most cases the companies that dominated the scarce period captured a shrinking share of the value.



Catalogued chokepoints. Rail, drawn fiber and the ceramic insulator: three artefacts of past scarcity, each genuinely indispensable in its moment, each eventually a commodity. Original art for this report.

Applying the Trap to AI

Now run the pattern against the current AI stack.

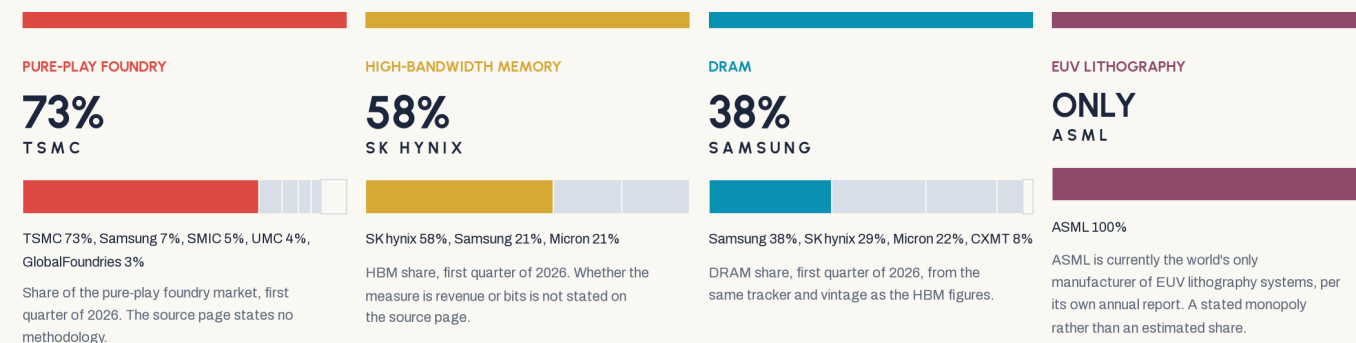
NVIDIA's position is the strongest chokepoint in the market, and it is not static. Public Futurum material says NVIDIA retained more than 90%—and elsewhere 95%—of the GPU market in its Q1 2025 Data Center Semiconductor Spot Check. That is not the same measure as the broader accelerator market, which includes custom silicon. The distinction between those markets is the entire story, and it is reason enough not to turn a single share figure into a permanent-tollbooth thesis.

Because the decomposition matters more than the headline. In *training*, CUDA remains a genuine moat:

FlashAttention, PagedAttention, continuous batching, speculative decoding, TensorRT-LLM, NCCL—two decades of accumulated optimization that does not run anywhere else. In *inference*, the relevant question is not a single share estimate but where large buyers can substitute. Google says TPUs serve as the backbone for AI across nearly all of its products, alongside custom CPUs and NVIDIA GPUs. AWS says Trainium2 already serves more than half of the tokens through Bedrock. Broadcom reported \$20 billion in AI revenue in fiscal 2025 and more than \$73 billion in AI orders on hand. TrendForce projects ASIC-based AI servers will represent 27.8% of the AI-server market in 2026.

FIGURE 06 · SOURCE-LOCKED EVIDENCE

The chokepoint is not one market. It is four, each defined differently



FOR SCALE. NOT FOR SHARE

\$193.7B

NVIDIA data-center revenue for fiscal 2026, on total revenue of \$215.9 billion. Public Futurum material puts NVIDIA above 90% of the GPU market in its first-quarter 2025 spot check, which is a different measure from the broader accelerator market, and that market includes custom silicon.

No accelerator-share series is charted here. No public source supports one, so the single-bar treatment used in the previous edition has been replaced by four separately defined markets, each printed with its own definition and vintage.

Source: Counterpoint Research, global semiconductor foundry market share, June 4, 2026; Counterpoint Research, global DRAM and HBM market share, June 8, 2026; ASML 2025 annual report filed with the SEC; NVIDIA fourth quarter and fiscal 2026 results, February 25, 2026; Futurum Intelligence public material, Q1 2025 Data Center Semiconductor Spot Check.

Four separately defined markets, each printed with its own definition and vintage. No accelerator-share series appears here: no public source supports one. NVIDIA's data-center revenue is shown as scale, not as share.

That is steps four and five of the pattern, happening in real time. Note *why*: Deloitte's 2026 TMT Predictions, as reported, put inference at roughly two-thirds of AI *compute* in 2026—not necessarily two-thirds of spending. It rewards cost-per-token over ecosystem breadth. That is the economic pressure that makes custom silicon and better serving stacks consequential. The economics of specialization are doing what they always do.

Efficiency is the other half of the vise. Quantization, distillation, sparsity, better serving stacks, and small models tuned for narrow tasks have driven the cost of a given unit of capability down by orders of magnitude in three years. DeepSeek-R1 demonstrated that frontier-adjacent capability could arrive from an unexpected direction at an unexpected price—and, the part people forget, it *increased* total compute demand even as it demolished assumptions about what compute a given capability requires. Huang's own five-layer essay concedes the point: open models activate demand across the whole stack. Meanwhile intelligence keeps migrating toward the user, into phones, PCs, vehicles, cameras, and industrial controllers. That migration has a completely different competitive structure from the data center; it is part of the hierarchy I am describing, not a claim that the cloud disappears.

The architecture is a hierarchy, not a war

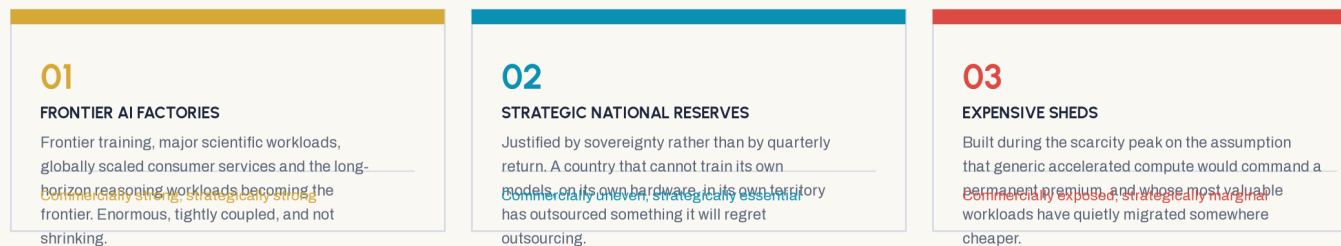
I want to be careful here, because the Trap is regularly misread as an edge-versus-cloud prediction. It is not. Frontier training will continue to require enormous, tightly coupled clusters, as will major scientific workloads, globally scaled consumer services and the long-horizon reasoning workloads becoming the frontier. Those facilities are not going away and their scale is not shrinking.

What changes is everything *below* the frontier. Routine, private, latency-sensitive, regulated and specialized inference has strong technical and economic reasons to move closer to the user. That migration does not eliminate data centers. It restratifies them.

The likely outcome is three tiers. Some facilities remain genuine frontier AI factories, doing work nothing else can do. Some become strategic national reserves, justified by sovereignty rather than by quarterly return. And some—the ones built during the scarcity peak, on the assumption that generic accelerated compute would command a permanent premium—become expensive sheds whose most valuable workloads have quietly migrated somewhere cheaper.

FIGURE 07 · ANALYTICAL CONSTRUCTION

Three tiers, not a war between cloud and edge



An analytical construction from the report text. The Trap is not an edge-versus-cloud prediction: what changes is everything below the frontier. The tiers carry no quantities and no forecast of how many facilities land in each. Source: Jensen Huang, AI Is a 5-Layer Cake, NVIDIA blog, March 10, 2026; Alan Shimel, The Indispensability Trap, forthcoming September 2026.

An analytical construction from the report text. It is not a forecast of how many facilities land in each tier, and one campus can occupy more than one.

The financial exposure is not hypothetical

This is where the Trap stops being a historical argument and starts being a balance-sheet one.

The balance-sheet point is visible without stitching together analyst estimates. Microsoft reported \$115.948 billion of additions to property and equipment and \$38.534 billion of depreciation, amortization and other expense for the twelve months ended June 30, 2026. Alphabet reported \$44.924 billion of property-and-equipment purchases in Q2 and \$132.402 billion over the trailing twelve months, against \$7.104 billion of Q2 depreciation. Amazon reported \$173.028 billion of property-and-equipment purchases over the trailing twelve months to June 30. These are not a single comparable aggregate; their fiscal periods and line items

differ. They do show why the build and the accounting charge move on different clocks.

The useful-life assumption is contested. Accelerators refresh on three-to-four-year cycles; depreciating them over five to six is defensible if they retain value as older-generation inference capacity and indefensible if they do not. Critics including Michael Burry put the resulting understatement at roughly \$176 billion across 2026–2028. Reasonable people dispute the number. Nobody disputes that the direction of the error favors reported earnings.

The financing has shifted too. Meta reported \$24.910 billion of long-term debt issued in Q2 2026. Oracle said it raised \$43 billion in debt financing and \$5 billion in equity financing in the fiscal year ended May 31, 2026; Alphabet

disclosed \$49.6 billion of June financing for general corporate purposes, including capital expenditures to scale AI infrastructure and global compute. Those issuer disclosures are more useful than a vague sector total. Separately, a Moody's analysis reported in February put signed-but-not-commenced data-center lease commitments at \$662 billion—future commitments, not incurred debt. The \$1.5 trillion figures should remain distinct: JPMorgan's is an estimate of investment-grade bonds over five years; Morgan Stanley's is a data-center financing-gap estimate through 2028.

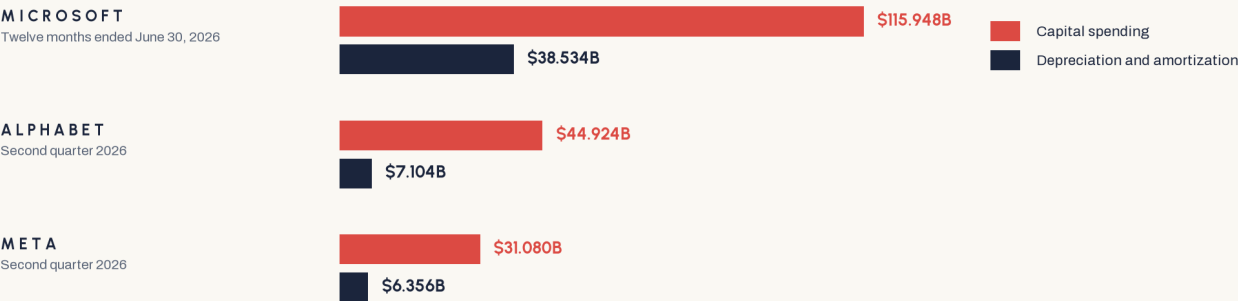
Then add power. Futurum describes a power-generation gap, and its surveyed industry professionals expect 33% of data centers to operate on 100% off-grid power by 2030. That is a practitioner expectation, not an engineering forecast. Berkeley Lab's 2025 Update, published in 2026, projects data centers at 9.5% to 15.3% of U.S. electricity use by 2030, with an 11.8% central estimate; Berkeley Lab's prior report placed the 2023 share at roughly 4.4%. Futurum's own 2030 semiconductor forecast carries a bull-to-bear spread wide enough to be an honest admission that nobody knows.

FIGURE 08 · SOURCE-LOCKED EVIDENCE

Financial exposure runs on three different clocks

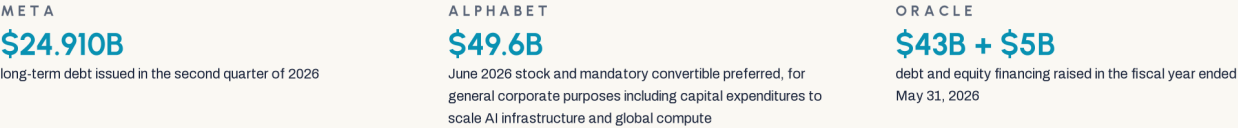
ONE THE BUILD AND THE ACCOUNTING CHARGE

Each pair is drawn from the same reported period for that company. Periods differ between companies and are printed on each row.

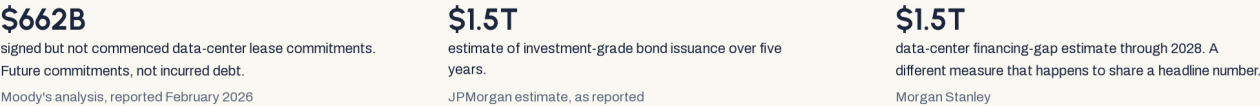


Accelerators refresh on three-to-four-year cycles. Depreciating them over five to six is defensible if they retain value as older-generation inference capacity and indefensible if they do not. Critics including Michael Burry put the resulting understatement at roughly \$176 billion across 2026 to 2028. Reasonable people dispute the number; nobody disputes that the direction of the error favors reported earnings.

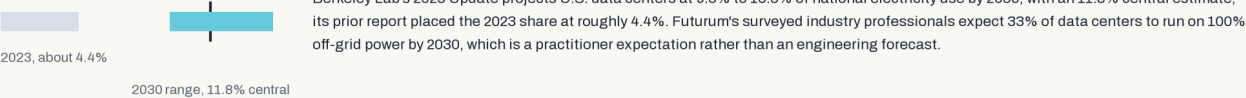
TWO THE FINANCING, FROM THE ISSUERS THEMSELVES



THREE COMMITMENTS AND ESTIMATES, WHICH ARE NOT ADDED TOGETHER



AND THEN POWER



Source: Microsoft FY2026 Q4 results; Alphabet Q2 2026 exhibit 99.1 filed with the SEC; Meta Q2 2026 results; Oracle Q4 and FY2026 results; Moody's analysis of signed but not commenced lease commitments, as reported February 2026; JPMorgan estimate of investment-grade bond issuance over five years, as reported; Morgan Stanley, Bridging the Data Center Gap; Lawrence Berkeley National Laboratory, 2025 U.S. Data Center Energy Usage Update; Berkeley Lab, January 15, 2025 news release on the 2024 report; Futurum Intelligence, survey expectation on off-grid data centers by 2030.

The build, the accounting charge and the financing are shown separately because their periods differ. Capex-versus-depreciation pairs are drawn from the same period for each company. Commitments and estimates are labeled by source tier and are not added together.

I am not predicting collapse. Demand is real, revenue is growing fast, and a company that has already sold the bonds is unlikely to cancel the servers those bonds were

earmarked for. I am saying that a capital structure this leveraged, depreciating this fast, dependent on power that does not yet exist, deployed into a workload mix already

migrating toward cheaper silicon, carries a great deal of embedded assumption.

Commercial versus strategic indispensability

One distinction the policy conversation persistently blurs, and I will return to it later.

A facility, a chip program, or a national compute reserve can be commercially mediocre and strategically essential at

the same time. Those are different tests. A country that cannot train its own models, on its own hardware, in its own territory, has outsourced something it will regret outsourcing—and that argument justifies redundancy no discounted cash flow analysis would ever approve. Conflating the two tests is how a national-security case quietly becomes a subsidy case.

Commercial mediocrity and strategic necessity are different tests. Conflating them is how a national-security case quietly becomes a subsidy case.

What the Trap actually predicts

Not that AI fails—the opposite. The Trap predicts that AI succeeds so comprehensively it stops being scarce, and that some of the companies currently priced as permanent tollbooths discover indispensability and pricing power are not the same asset.

That is not a comfortable middle, and muddling through is not painless. Every wave that ended this way destroyed real investments, occupations, companies and communities on its way to producing greater aggregate wealth. Ask a telecom equipment engineer from 2001 how the fiber build worked out for her. The economy adapted. She did not necessarily.

WHAT THE TRAP DOES AND DOES NOT PREDICT

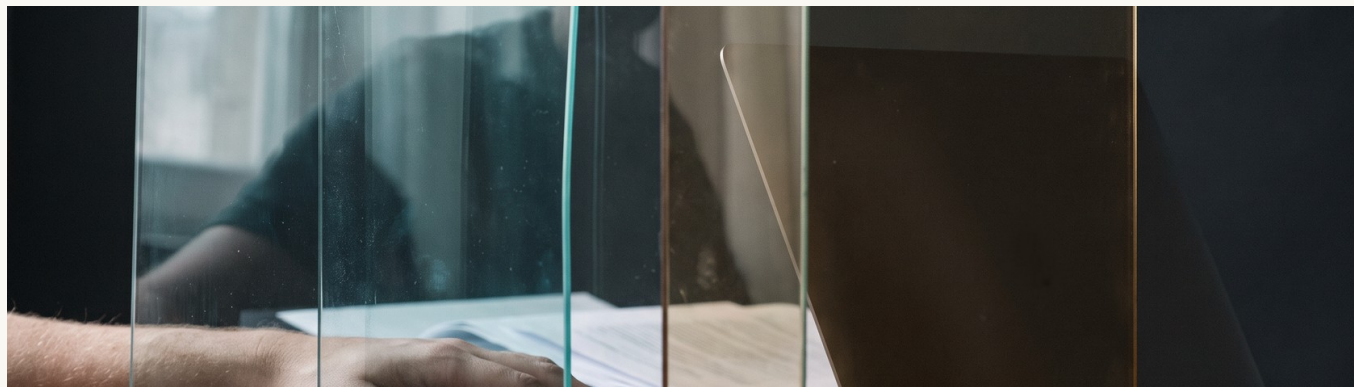
Read it as a claim about pricing power, not about capability

- **It does not predict that AI fails.** It predicts the opposite: that AI succeeds so comprehensively it stops being scarce.
- **It is not an edge-versus-cloud forecast.** Frontier training keeps needing enormous, tightly coupled clusters.
- **It separates two assets.** Indispensability and pricing power are not the same thing, and they move apart.
- **It is not painless.** Every wave that ended this way destroyed real investments, occupations, companies and communities.

05 THE PRESENT TENSE

The Reflections Are Already Colliding

None of this is futurism. All three reflections are visible now, and they collide in two specific places: inside companies deploying agents, and outside buildings drawing power.



Delegated assignments. Organizations are no longer asking models for content. They are delegating longer, harder, multi-step work with real system access and real consequences. Original art for this report.

Everything above sounds like futurism. It is not. All three reflections are visible in the present tense, and the collision is happening in two specific places: inside companies deploying agents, and outside buildings drawing power.

Agentic AI is today's AI

I want to dispense with the framing that treats agents as a speculative precursor to something real. They are the something real.

Google Cloud's second-annual ROI of AI study found that 52% of executives reported their organizations actively using AI agents, with 39% having launched more than ten. Read that number carefully, because the methodology matters: Google Cloud commissioned National Research Group to survey 3,466 senior leaders and executives across 24 countries, all at global enterprises *that already had generative AI deployments*. That is a sample of leading adopters, not the American economy. Google does not publish field dates, question wording, sampling method, weighting, margin of error, or mode. Fifty-two percent of the vanguard is a very different claim from fifty-two percent of employers.

FIGURE 09 · SOURCE-LOCKED EVIDENCE

Agents inside the vanguard, with the methodology attached

52 of 100 executives said their organization was actively using AI agents



METHODOLOGY. AS PUBLISHED

3,466 senior leaders and executives across 24 countries, all at global enterprises that already had generative-AI deployments. Commissioned by Google Cloud and conducted by National Research Group, September 2025.

Not published: field dates, question wording, sampling method, weighting, margin of error, mode. Fifty-two percent of the vanguard is a very different claim from fifty-two percent of employers.

A vendor-commissioned survey of leading adopters, not of the American economy. Field dates, question wording, sampling method, weighting, margin of error and mode are not published.

But within that vanguard the shift is qualitative. Organizations are no longer asking models for content; they are delegating assignments—longer, harder, multi-step work with real system access and real consequences. And that delegation forces a choice no vendor survey captures. Every organization deploying agents at scale is deciding, usually implicitly and usually without putting it in a memo, whether the productivity gain shows up as *more output from the same people* or *the same output from fewer people*.

That single decision, made independently thousands of times over the next few years, determines which reflection we end up living in. It is not a technology question. Nobody

at NVIDIA or OpenAI or Anthropic decides it. A few hundred thousand managers decide it, one budget cycle at a time.

A few hundred thousand managers decide which reflection we live in, one budget cycle at a time.

The evidence so far is mixed in exactly the way you would expect. Workers gain real capability. Employers appear to capture most of the resulting surplus. Entry-level and routine cognitive work is affected well before any whole occupation disappears.

The backlash is the same argument, in concrete

The data center fight is not a separate story about NIMBY-ism. It is the physical manifestation of the ownership question.

The public opinion numbers are extraordinary. Gallup found that 71% of Americans opposed AI data-center construction in their local area, including 48% who were strongly opposed, in telephone interviews conducted March 2–18, 2026 with a random sample of 1,000 adults (± 4 points). Spotlight PA's reporting of Gallup data put total opposition at

75% among Democrats and 63% among Republicans. Spotlight PA also reported that conservative Republicans opposed local data centers at higher rates than moderate Republicans, and quoted Yale's Anthony Leiserowitz on the unusual alignment.

The organizing has followed. Data Center Watch counted at least 75 projects worth roughly \$130 billion blocked or delayed in the first quarter of 2026 alone—"roughly matching the scale of all of 2025 in just three months," in the tracker's wording. The tracker says opposition groups more than doubled across 49 states; secondary reporting counted 833 groups. On July 14, New York's governor issued an executive order pausing discretionary state permits for new hyperscale facilities of 50 MW and above for up to one year; Monterey Park, California's Measure NDC made its city ban durable by ballot.

FIGURE 10 · SOURCE-LOCKED EVIDENCE

The backlash, in numbers that come from different places

POLLING

71%

of Americans opposed AI data-center construction in their local area, including 48% strongly opposed.



Strongly opposed 48%, somewhat opposed 23%

Telephone interviews conducted March 2 to 18, 2026 with a random sample of 1,000 adults, margin of error plus or minus 4 points. The 75% Democratic and 63% Republican splits are Spotlight PA's reading of Gallup data, not Gallup's own published figures.

ORGANIZING AND STATUTE

AT LEAST 75

projects blocked or delayed in the first quarter of 2026, worth roughly \$130 billion, roughly matching the scale of all of 2025 in just three months, in the tracker's own wording

49 STATES

opposition groups more than doubled across 49 states; secondary reporting counted 833 groups

MORE THAN 300 BILLS

state data-center bills filed in the first six weeks of 2026, with statewide moratorium proposals in 14 states

50 MW AND ABOVE

New York's July 14 executive order paused discretionary state permits for new hyperscale facilities at or above 50 MW for up to one year; Monterey Park, California made its city ban durable by ballot

PRICES

7.1%

observed rise in residential electricity prices in 2025, more than double the inflation rate, with some states above 20%. Consumer Reports, using federal data.

5.6%

forecast rise in the 2026 average, to 18.27 cents per kilowatt-hour. A federal forecast, not an observed price.

On causation: PJM's independent market monitor says data-center load growth is the primary reason for recent and expected capacity-market conditions and high prices, while warning in the same section that the accuracy of the relevant load forecasts is highly questionable. Both halves are published here.

Source: Gallup, telephone interviews March 2–18, 2026, random sample of 1,000 adults, plus or minus 4 points; Spotlight PA reporting of Gallup data, June 2026; Data Center Watch (10a Labs), Q1 2026 report; ExtremeTech reporting of the 833-group count; Consumer Reports analysis of federal residential electricity price data; U.S. Energy Information Administration, Short-Term Energy Outlook, August 2026; Monitoring Analytics, 2026 Q1 State of the Market Report for PJM.

Polling, project tracking, statute and price data are shown in separate registers because they are separate kinds of evidence. The forecast price is labeled as a forecast.

UCLA's Megan Mullin identified the common thread precisely: the opposition is rooted in "our deep affinity for where

we live." That is not an ideology. It is the thing ideologies are built on top of.

The grievances are concrete and mostly legitimate. Residential electricity prices have been rising faster than inflation: *Consumer Reports*, using federal data, put the 2025 increase at 7.1%—more than double inflation—with some states above 20%. The EIA’s August 2026 outlook forecasts a further 5.6% rise in the 2026 average, to 18.27 cents per kilowatt-hour; that is a forecast, not an observed price. In PJM, the independent market monitor says data-center load growth is the primary reason for recent and expected capacity-market conditions and high prices, while warning that the relevant load forecasts are highly questionable. Water, noise, land use, and diesel backup generation are real local costs. And the employment math is genuinely unfavorable relative to the investment: Virginia’s Department of Taxation reported that in fiscal 2024 data-center operators told VEDP they had added 1,197 net new jobs against about \$32 billion of investment—figures the operators self-reported and that VEDP and the Department did not independently validate. The same report estimates that the exemption generated \$2.1 billion more in cumulative tax revenue than its value over two overlapping five-year periods, an average annual return on incentive of 11% and a revenue-to-cost ratio of 1.7. Construction employment is substantial and real. Permanent employment is modest.

Which brings us back to Louise Lucas, who has been making a straightforward argument that the industry has struggled to answer: that companies operating in Virginia should be paying their fair share, and that the revenue foregone through the exemption is revenue not available for teachers, health coverage, transit, or childcare. Her colleague Danica Roem put the principle even more plainly—“If you want to operate here, you’ve got to play by the same rules that everybody else does.”

The industry’s counter is not trivial. A Northern Virginia Technology Council report found that data centers generated nearly \$40 billion in statewide economic impact in 2025, supported more than 112,000 Virginia jobs, and contributed more than \$1.5 billion in annual state tax revenue. It is an industry-sponsored impact measure, not a neutral measure of distribution. Those are real people with real stakes, and any analysis that treats them as astroturf is not serious.

The question was never whether data centers provide benefit. It is the distribution of benefit and cost, and the allocation of risk if projected demand or the underlying technology changes. Virginia’s answer—preserve the equipment exemption, tax the power, cap it, sunset it in two years—satisfies nobody and is probably about right for a situation this uncertain.



The argument in concrete. Water, noise, land use, diesel backup and transmission are real local costs. Virginia preserved the equipment exemption, taxed the power, capped it and sunset it in two years. Original art for this report.

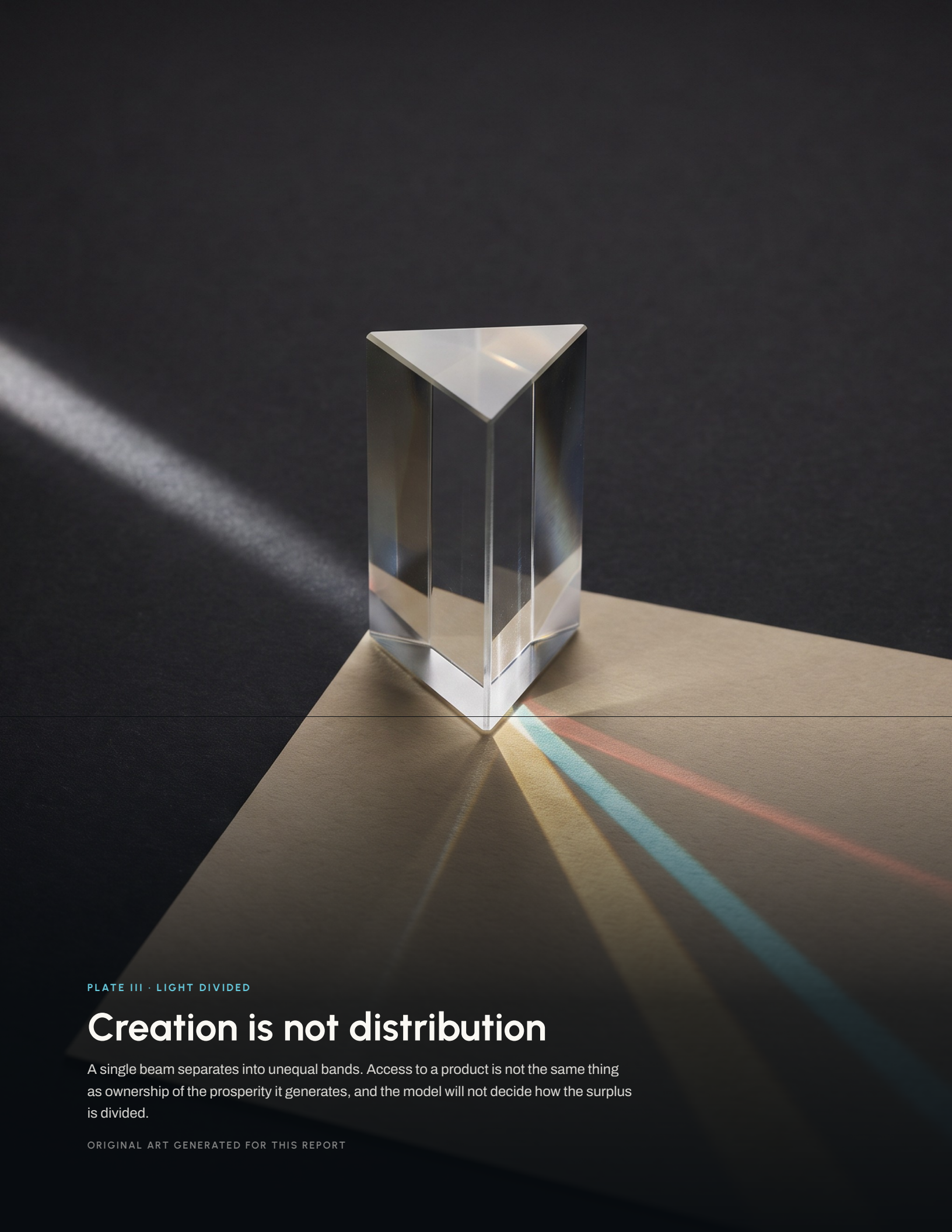


PLATE III · LIGHT DIVIDED

Creation is not distribution

A single beam separates into unequal bands. Access to a product is not the same thing as ownership of the prosperity it generates, and the model will not decide how the surplus is divided.

ORIGINAL ART GENERATED FOR THIS REPORT

06 OWNERSHIP

Who Owns the Abundance?

If the difference between Utopia and Serfdom is ownership rather than capability, the policy question is not how to make AI work. It is who holds the assets when it does.

If the difference between Utopia and Serfdom is ownership rather than capability, then the policy question is not "how do we make AI work." It is "who holds the assets when it does."

I want to separate the debate into five categories, because collapsing them into a single argument about taxation is how the conversation keeps failing.

FIGURE 11 · ANALYTICAL CONSTRUCTION

Five categories, five different instruments

1	2	3	4	5
INFRASTRUCTURE COST ALLOCATION	TAXATION	BROADENED OWNERSHIP	COMPETITION	WORKER TRANSITION
- Large-load tariffs that assign grid upgrade costs to the loads requiring them - Direct financing structures for transmission - Long-term take-or-pay commitments against stranded infrastructure - Community benefit agreements - Decommissioning obligations	- Corporate income taxes and capital gains rates - Potential excess-profit mechanisms - Actual enforcement against profit shifting	- Sovereign wealth funds and public equity stakes - Employee ownership and profit sharing - Expanded retirement-fund participation in private AI capital - Smaller stakes taken as compensation for specific public inputs	- Antitrust enforcement across the whole stack - Interoperability and data portability - Open standards and capable open-weight models - Public and academic compute allocations	- Wage insurance and portable benefits - Meaningful unemployment support - Apprenticeships and advance notice requirements - Employer-connected retraining tied to real openings
Most immediately actionable, least ideologically fraught	Conventional and well understood	Carries the state-dependence problem the report examines at length	Works with the Indispensability Trap rather than against it	Least attention, and probably matters most to actual people

The numbering appears here so the running subheads do not double it. None of the five categories solves the problem alone; anyone selling a single mechanism as the answer is not describing a policy so much as an identity. Source: Sanders summary of the American A.I. Sovereign Wealth Fund Act; Vinod Khosla at TechCrunch Disrupt, October 28, 2025.

A navigation summary of the five categories the report then develops in order. The numbering appears here so the running subheads do not double it.

Infrastructure cost allocation

The most immediately actionable, and the least ideologically fraught. Large-load tariffs that assign grid upgrade costs to the loads requiring them. Direct financing structures for transmission. Long-term take-or-pay commitments so rate-payers are not left holding stranded infrastructure if a project cancels. Community benefit agreements. Decommissioning obligations.

Taxation

Conventional and well understood: corporate income taxes, capital gains rates, potential excess-profit mechanisms, and—probably most consequentially—actual enforcement against profit shifting. A great deal of the concentration problem is downstream of the fact that the most profitable

companies in history have been extraordinarily effective at booking income where it is not taxed.

Broadened ownership

Sovereign wealth funds, public equity stakes, employee ownership, profit sharing, and expanded retirement-fund participation in private AI capital.

Sanders' **introduced** proposal is the most aggressive version: a one-time 50% tax paid in stock by companies with more than \$200 million in annual AI revenue, deposited into a Treasury fund the sponsor estimates at roughly \$7 trillion, governed by a seven-member commission, distributing about 5% of value annually.

Let me give it a fair hearing, because it deserves one and mostly does not get one.

A FAIR HEARING FOR THE SOVEREIGN FUND

The strongest case for it

- **Provenance.** The models were trained on the collective output of human civilization, much of it publicly funded and none of it compensated.
- **Precedent.** Norway and Alaska show that sovereign funds can be governed competently over decades.
- **It is not fringe.** Vinod Khosla, a venture capitalist rather than a socialist, proposed that government take 10% of every corporation and pool it for the people, and called sharing the wealth of AI a really, really big need.

The strongest argument for it is the provenance argument, and it is a good argument. These models were trained on the collective output of human civilization, much of it produced with public funding, none of it compensated. There is a real sense in which the public is an unpaid contributor of capital to this enterprise. Norway and Alaska demonstrate that sovereign funds can be governed competently over decades. And Sanders is correct that the idea is not fringe—Sam Altman has publicly floated public wealth vehicles, and Khosla's 10%-of-everything proposal is arguably more radical in scope.

The arguments against are also serious. A 50% one-time levy is large enough to materially affect capital formation and would be litigated for years. Distributing 5% of fund value annually before the underlying assets have proven durable is a mechanism for paying out during a bubble and defaulting after it. And seven presidential appointees exercising voting control over the boards of the country's most strategically significant companies has obvious failure modes in both directions.

The deepest problem applies to every version of state equity ownership, including the ones I find more attractive: it makes government financially dependent on the companies it must regulate. A state that owns half an industry has a fiscal interest in that industry's profits. That is not a hypothetical conflict. It is the entire reason we have historically been careful here.

Which is why I am more persuaded by mechanisms that broaden ownership without concentrating it in the state: employee equity participation, mandated profit-sharing, public pension access to private AI capital at institutional

The serious objections

- **Scale.** A one-time 50% levy is large enough to affect capital formation materially, and would be litigated for years.
- **Timing.** Distributing 5% of fund value annually before the assets have proven durable pays out during a bubble and defaults after it.
- **Governance.** Seven presidential appointees voting the boards of the most strategically significant companies in the country has obvious failure modes in both directions.
- **The deepest problem.** State equity makes government financially dependent on the industry it must regulate.

rather than retail terms, and smaller sovereign stakes taken as compensation for specific public inputs—compute, spectrum, land, energy infrastructure, publicly funded research—rather than as a general levy on success.

Competition

Antitrust enforcement, interoperability requirements, data portability, open standards, and the availability of capable open-weight models are the most direct mechanisms for preventing durable concentration—and they have the advantage of working *with* the Indispensability Trap rather than against it. Every force in the Trap that erodes a choke-point (substitution, standardization, efficiency, distribution) is a force competition policy can accelerate. Public and academic compute allocations do more to keep frontier research from becoming a private club than any tax does.

Worker transition

The category that gets the least attention and probably matters most to actual people: wage insurance, portable benefits, meaningful unemployment support, apprenticeships, advance notice requirements, and employer-connected retraining tied to real openings rather than certificate mills. It matters more than aggregate policy because of the mismatch problem, which is next.

None of these five categories solves the problem alone. Anyone selling a single mechanism as the answer—UBI, antitrust, a sovereign fund, retraining—is not describing a policy so much as an identity.

Anyone selling a single mechanism as the answer is not describing a policy so much as an identity.

07 WORK

AI Jobs and the Missing Timetable

New jobs appearing in the aggregate is true and useless to individuals. The gap between those two facts is where most of the human damage lives.

I am not going to reassure you that new jobs will appear. That claim is true in the aggregate and useless to individuals, and the gap between those two facts is where most of the human damage lives.

Here is what is visible now, in real hiring. Data center construction and the skilled trades—electricians, pipefitters, millwrights—in genuine shortage. Power engineering, substation work, transmission planning, and industrial cooling. Semiconductor fabrication and the equipment supply chain, including advanced packaging where TSMC's CoWoS capacity is itself a chokepoint. AI implementation and systems integration. Agent security, evaluation, red-teaming, and governance. Process redesign and human oversight roles—the people who define what an agent is allowed to do and who is accountable when it does something else.

What may expand later, on timelines I would not bet on precisely: robotics deployment and field service, physical-AI safety engineering, autonomous fleet operations, AI-assisted science and manufacturing.

WHERE THE WORK ACTUALLY IS

Three different timetables in one labor market

Visible now, in real hiring. Data-center construction and the skilled trades in genuine shortage; power engineering, substation work, transmission planning and industrial cooling; semiconductor fabrication and advanced packaging; implementation and systems integration; agent security, evaluation, red-teaming and governance; process redesign and human oversight.

Possibly later. Robotics deployment and field service, physical-AI safety engineering, autonomous fleet operations, AI-assisted science and manufacturing.

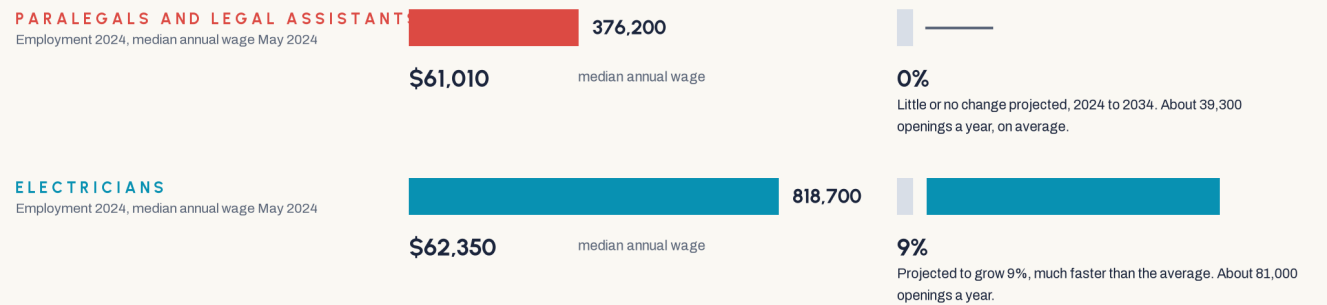
Mostly invisible in job titles. Existing professions acquiring accountability for work a person did not personally perform.

And here is the point most workforce commentary misses. Most AI jobs will not have "AI" in the title. Existing professions are acquiring new responsibilities—defining objectives precisely enough for a machine to pursue them, scoping machine permissions, judging outputs, handling escalated exceptions, and accepting accountability for work a person did not personally perform. That last one is a genuine professional shift, and most organizations have not thought about it at all.

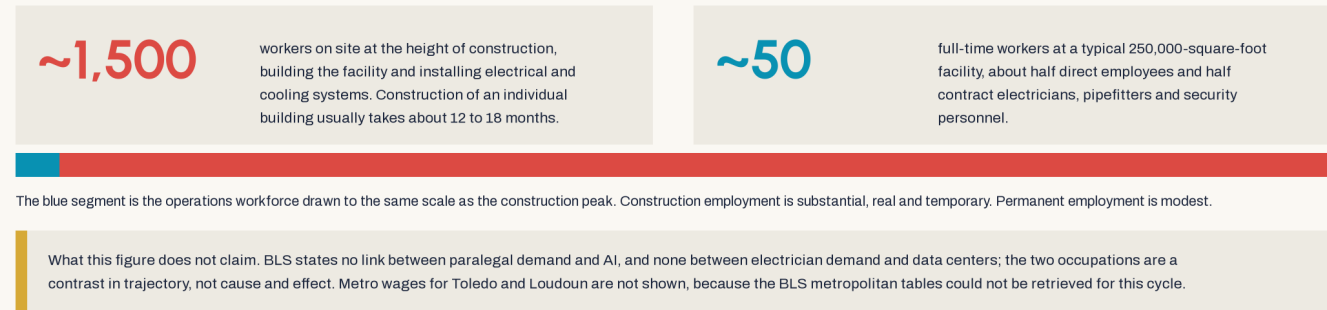
FIGURE 12 · SOURCE-LOCKED EVIDENCE

The mismatch is national, and it is not a retraining problem

ONE TWO NATIONAL OCCUPATIONS, OPPOSITE TRAJECTORIES



TWO ONE VIRGINIA DATA CENTER, TWO VERY DIFFERENT WORKFORCES



Source: U.S. Bureau of Labor Statistics, Occupational Outlook Handbook, Paralegals and Legal Assistants; U.S. Bureau of Labor Statistics, Occupational Outlook Handbook, Electricians; Virginia JLARC, Data Centers in Virginia, Report 598, December 2024.

National BLS occupational data and Virginia's own legislative audit. BLS states no link between paralegal demand and AI, and none between electrician demand and data centers; the two occupations are shown as a contrast in trajectory, not as cause and effect. Metro-level wages are not published here because the BLS metro tables could not be retrieved.

Now the mismatch, stated as plainly as I can. Jobs may grow in aggregate while specific people suffer permanent loss. A 47-year-old paralegal in Toledo whose firm just deployed a document-review agent does not benefit from a shortage of substation electricians in Loudoun County. The jobs are in the wrong place, require different skills, and arrive

on a different schedule than the loss. Aggregate employment statistics are true, and they are not a livelihood. Every serious policy response has to be built around the individual timeline rather than the economy-wide one, and almost none of them are.

Aggregate employment statistics are true, and they are not a livelihood.

08 GEOPOLITICS

Geopolitics: When Indispensability Becomes National Policy

Compute, semiconductors, models, energy and robotics are now instruments of state power, and eighteen months of policy churn has a cost of its own.

Al compute, semiconductors, models, energy, and robotics are now instruments of state power, and the policy environment reflects it.

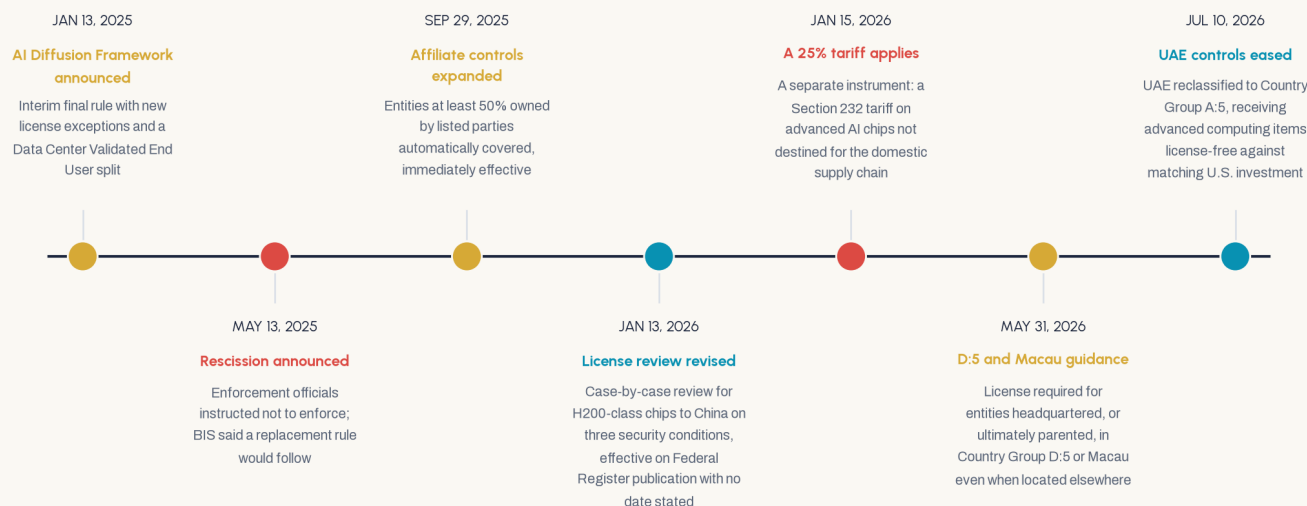
U.S. export controls have run through several regimes in eighteen months—a tiered diffusion framework, its rescission, and then a January 2026 pivot to case-by-case licens-

ing for H200-class chips to China under specified security conditions, paired with a 25% tariff on advanced AI chips not destined for the domestic supply chain. Whatever one thinks of the substance, the churn itself has a cost: allied governments and multinationals now maintain compliance architectures that assume the rules will change again.

FIGURE 13 · SOURCE-LOCKED EVIDENCE

Eighteen months of export-control churn

FOUR REGIMES IN EIGHTEEN MONTHS



The churn itself has a cost: allied governments and multinationals now maintain compliance architectures that assume the rules will change again. Source: BIS, Framework for Artificial Intelligence Diffusion, announced January 13, 2025; BIS, rescission of the AI Diffusion Rule, announced May 13, 2025; BIS, expansion of end-user controls to cover affiliates, September 29, 2025; BIS, Revision to License Review Policy for Advanced Computing Commodities, January 13, 2026; White House fact sheet on certain advanced computing chips, January 14, 2026; BIS guidance on advanced computing items, May 31, 2026.

Dated primary instruments. The January 2026 license-review revision is effective on Federal Register publication and carries no stated date; January 15, 2026 is the tariff's effective date, a separate instrument.

That unpredictability is among the strongest drivers of sovereign AI. Chatham House has made the argument directly—states that cannot rely on the stability of an AI stack they do not control have a core strategic interest in diversification. Every national compute program, every domestic model effort, every Gulf-state AI factory is partly a hedge

against access turning out to be a political variable rather than a commercial one. The result is parallel ecosystems, deliberate redundancy, and domestic chip subsidies in a dozen countries. Some of that is rational. Some will prove to be economically indefensible capacity built for defensible reasons.

Does the security interest require this company's continued dominance, or merely the capability to exist under domestic control?

Which returns me to the distinction I flagged earlier, the one most likely to be abused. A national-security justification is a legitimate reason to fund what a market would not, and it is also the most effective device available for converting private commercial interest into public subsidy without accountability. The test I would apply is simple: does the security interest require this specific company's continued

dominance, or merely the capability to exist under domestic control? If the latter—and it usually is—the security argument supports *more* domestic capacity and *more* competition, not the protection of an incumbent. Strategic and commercial indispensability point in different policy directions more often than their advocates admit.



Deliberate redundancy. The same capability, duplicated across a border, lit differently on each side. Some of that is rational hedging. Some will prove to be economically indefensible capacity built for defensible reasons. Original art for this report.



PLATE IV · THE SAME SURFACE

The reflection still depends on where you stand

A scientist sees faster discovery. An electrical contractor sees years of work. A recent graduate sees the first rung of a career disappearing. A homeowner near a campus sees transmission lines and a rising bill. All of them are looking at the same revolution.

ORIGINAL ART GENERATED FOR THIS REPORT

09 CONCLUSION

The Mirror, One Last Time

Utopia and Serfdom are not opposing technological forecasts. They differ on one variable: who owns the productive assets when the gains arrive.

Al Utopia and AI Serfdom are not opposing technological forecasts. That is the thing I most want readers to take away. Both assume the technology works, both assume enormous productivity gains, and both assume machines take on a large share of what humans currently do. They differ on one variable: who owns the productive assets when the gains arrive.

That is not a question engineers answer, and it will not be settled by scaling laws, benchmark scores or the next model release. It is a question about tax policy, antitrust enforcement, corporate governance, labor law, energy regulation and the political capacity of a democracy to make decisions about concentrated private power. Those are choices, and they are being made right now—mostly by default, mostly without anyone framing them as choices.

The Indispensability Trap is why I think neither extreme arrives intact.

AI will likely create more wealth and capability than the skeptics expect—the deployment evidence is already too strong to argue otherwise. It will likely disrupt employment and concentrate its early gains more severely than the optimists admit—the entry-level data is already too consistent to wave away. But its very success will attract competition, technical substitution, regulation, and political resistance, exactly as every prior wave of indispensable technology has.

THE ONE VARIABLE

What the two loudest forecasts actually disagree about

- Both assume the technology works.
- Both assume enormous productivity gains.
- Both assume machines take on a large share of what humans currently do.
- **They differ on who owns the productive assets when the gains arrive.** That is a question about tax policy, antitrust, corporate governance, labor law and energy regulation, and it is being answered right now, mostly by default.

Some of today's chokepoints will survive. NVIDIA's training position looks durable for years. Others will weaken; the inference market is already fragmenting. Value will migrate from the layer that supplies intelligence toward the layers that use it, because that is what has happened every previous time.

The third reflection is not a comfortable one. It is neither nirvana nor proletarian hell but the historically familiar outcome in which a society argues, overbuilds, regulates, competes, redistributes, adapts, and muddles its way into a changed economic order—destroying a great deal on the way to creating more. Muddling through is not a happy ending. It is just the most likely one, and it is only tolerable if the muddling includes deliberate decisions about ownership rather than deferring them until the concentration is complete and irreversible.

Muddling through is not a happy ending. It is just the most likely one.

This summer, the Stanford Digital Economy Lab organized “We Must Act Now: A Statement on AI’s Transformation of the Economy,” signed by more than 200 economists and AI researchers, including 16 Nobel laureates and signatories such as Yoshua Bengio. The public descriptions available

here do not support assigning the statement’s more specific language to Brynjolfsson or Bengio. The point still stands in my own words: the distributional question will not answer itself, and the window for answering it deliberately is not indefinite.

IN CLOSING

AI will influence what becomes possible. It will not decide who inherits the possibility.

That choice remains ours.

• REFERENCE

Sources and Notes

Twenty-five entries in the eight groups the argument uses. Every published source links to the original document. Where a figure or a sentence rests on an estimate, a survey or a forecast rather than a reported value, the entry says so.

The initiating conflict and wealth-sharing debate

- 01 Virginia HB 30 (2026–2028 biennial budget), Item 3-5.24, data-center electricity-consumption tax; signed June 30, 2026. Rate of \$0.011/kWh beginning July 1, 2026, expiring July 1, 2028; no more than \$600 million is deposited to the general fund in a fiscal year, with excess collections returned pro rata. First collection: September 2026.
budget.lis.virginia.gov/item/2026/2/HB30/Chapter/3/3-5.24 www.governor.virginia.gov/newsroom/news-releases/2026/june-releases/name-1120477-en.html
- 02 Virginia Department of Taxation, *Virginia Tax Exemptions for Data Centers*, Report Document RD40, January 2, 2026. FY2025 reported tax benefit: \$1.941 billion. The figures were company-reported to VEDP and not independently validated by VEDP or the Department.
rga.lis.virginia.gov/Published/2026/RD40/PDF
- 03 Virginia Department of Taxation, *Virginia Tax Exemptions for Data Centers*, Report Document RD40, January 2, 2026: VEDP-collected, company-reported FY2024 figures of 1,197 net new jobs and approximately \$32.0 billion investment. Virginia Joint Legislative Audit and Review Commission, *Data Centers in Virginia*, Report 598, December 2024, for construction and operations employment.
rga.lis.virginia.gov/Published/2026/RD40/PDF jlarc.virginia.gov/pdfs/reports/Rpt598-2.pdf
- 04 Statements by Sen. L. Louise Lucas and Sen. Danica Roem, as reported by *Virginia Business* and *The American Prospect*.
virginiabusiness.com/virginia-budget-with-data-center-tax-moves-to-spanbergers-desk-pr ospect.org/2026/07/06/data-centers-win-lose-in-virginia

AI Utopia

- 05 Dario Amodè, “Machines of Loving Grace.”
darioamodei.com/essay/machines-of-loving-grace
- 06 Jensen Huang, “AI Is a 5-Layer Cake,” NVIDIA Blog, March 10, 2026; NVIDIA GTC 2026 announcement, March 3, 2026.
blogs.nvidia.com/blog/ai-5-layer-cake investor.nvidia.com/news/press-release-details/2026/NVIDIA-CEO-Jensen-Huang-and-Global-Technology-Leaders-to-Showcase-Age-of-AI-at-GTC-2026/default.aspx
- 07 Humanoid robotics deployment and benchmark context: BMW Group, “BMW Group Advances the Use of Physical AI in Production with Figure 03 Project in Spartanburg,” June 25, 2026; Figure AI, “F.02 Contributed to the Production of 30,000 Cars at BMW”; Figure AI, “BotQ: A High-Volume Manufacturing Facility for Humanoid Robots”; Agility Robotics, Toyota agreement and Digit throughput; CNBC reporting on Unitree and Optimus; and NIST’s humanoid-robot baseline performance benchmark.
www.press.bmwgroup.com/global/article/detail/T0458778EN/bmw-group-advances-the-use-of-physical-ai-in-production-with-figure-03-project-in-spartanburg?language=en www.figure.ai/news/production-at-bmw www.figure.ai/news/botq www.agilityrobotics.com/content/agility-robotics-announces-commercial-agreement-with-toyota-motor-manufacturing-canada www.agilityrobotics.com/content/digit-moves-over-100k-totes www.cnn.com/2026/08/14/china-humanoid-robots-unitree-ipo-tesla-optimus.html www.nist.gov/el/intelligent-systems-division-73500/humanoid-robot-baseline-performance-benchmark

AI Serfdom and displacement

- 08 Erik Brynjolfsson, Bharat Chandar, and Ruyu Chen, “Canaries in the Coal Mine? Six Facts about the Recent Employment Effects of Artificial Intelligence,” Stanford Digital Economy Lab, revised August 2026; and “Canaries, Interest Rates, and Timing,” February 2026.
digitaleconomy.stanford.edu/publication/canaries-in-the-coal-mine-six-facts-about-the-recent-employment-effects-of-artificial-intelligence digitaleconomy.stanford.edu/news/canariesaug26
- 09 Challenger, Gray & Christmas reports, January–June 2026; June 2026 report, 2025 year-end report, and May 2026 report.
www.challengergray.com/wp-content/uploads/2026/07/Challenger-Report-June2600986996.pdf www.challengergray.com/blog/2025-year-end-challenger-report-highest-q4-layoffs-since-2008-lowest-ytd-hiring-since-2010 www.challengergray.com/wp-content/uploads/2026/06/Challenger-Report-May-2026.pdf
- 10 International Monetary Fund, *Gen-AI: Artificial Intelligence and the Future of Work*, Staff Discussion Note SDN/2024/001, January 2024; current restatement, SDN/2026/001.
www.elibrary.imf.org/view/journals/006/2024/001/article-A001-en.xml www.imf.org/en/publications/staff-discussion-notes/issues/2024/01/14/gen-ai-artificial-intelligence-and-the-future-of-work-542379 www.imf.org/-/media/files/publications/sdn/2026/english/sdnea2026001.pdf
- 11 Larry Fink, 2026 BlackRock annual chairman’s letter.
www.blackrock.com/corporate/investor-relations/larry-fink-annual-chairmans-letter
- 12 Betsey Stevenson, “What Is There to Fear in a Post-AGI World?,” in *The Economics of Transformative AI*; author description of the chapter.
www.nber.org/books-and-chapters/economics-transformative-ai/what-there-fear-post-agi-world betseystevenson.com/
- 13 Sen. Bernie Sanders, American A.I. Sovereign Wealth Fund Act, S.4825, introduced June 18, 2026; Sanders and Ocasio-Cortez, Artificial Intelligence Data Center Moratorium Act, S.4214; House companion H.R.9442, introduced June 24, 2026; and Rep. Ocasio-Cortez’s House release.
www.congress.gov/bills/119th/congress/senate-bill/4825/all-info www.sanders.senate.gov/wp-content/uploads/AmericanAISovereignWealthFundActSummary.pdf www.congress.gov/bills/119th/congress/senate-bill/4214/ocasio-cortez.house.gov/media/press-releases/ocasio-cortez-introduces-house-version-ai-data-center-moratorium-act www.congress.gov/bills/119th/congress/house-bill/9442
- 14 Vinod Khosla, remarks at TechCrunch Disrupt 2025.
techcrunch.com/2025/10/28/vc-vinod-khosla-says-the-us-government-could-take

Agentic AI adoption

- 15 Google Cloud and National Research Group, “Second-annual ROI of AI Study,” September 4, 2025: 3,466 senior leaders and executives across 24 countries at global enterprises with generative-AI deployments; 52% actively using AI agents and 39% having launched more than ten. Google does not publish field dates, question wording, sampling method, weighting, margin of error, or mode.
www.googlecloudpresscorner.com/2025-09-04-Google-Cloud-Study-Reveals-52-of-Executives-Say-Their-Organizations-Have-Deployed-AI-Agents,-Unlocking-a-New-Wave-of-Business-Value_1

Infrastructure, energy, and the backlash

- 16 Lawrence Berkeley National Laboratory, 2025 U.S. Data Center Energy Usage Update: 9.5%–15.3% of U.S. electricity by 2030, central estimate 11.8%; 2024 report for the 4.4% 2023 share.
datacenters.lbl.gov/modeling-forecasting www.energy.gov/documents/united-states-data-center-energy-usage-report-2025-update newscenter.lbl.gov/2025/01/15/berkeley-lab-report-evaluates-increase-in-electricity-demand-from-data-centers
- 17 Gallup, March 2026 polling on local AI-data-center construction; Spotlight PA reporting of Gallup data for the partisan and ideological splits and Anthony Leiserowitz quotation.
news.gallup.com/poll/709772/americans-oppose-data-centers-area.aspx www.spotlightpa.org/news/2026/06/data-center-opposition-bipartisan-backlash-ai-tech-environment
- 18 Data Center Watch (10a Labs), Q1 2026 report and 2025 baseline; ExtremeTech reporting for the 833-group count.
www.datacenterwatch.org/q1-2026 www.datacenterwatch.org/report www.extremetech.com/energy/unprecedented-surge-in-local-backlash-puts-75-new-us-data-centers-on-hold
- 19 Hyperscaler capex, depreciation, lease, and financing figures: Microsoft FY2026 Q4; Alphabet Q2 2026 filing and earnings-call guidance; Amazon Q2 2026; Meta Q2 2026; Oracle FY2026; plus secondary reporting of Moody's, JPMorgan, and Morgan Stanley estimates where identified in the text.
www.microsoft.com/en-us/investor/earnings/fy-2026-q4/press-release-webcast www.sec.gov/Archives/edgar/data/1652044/000165204426000066/googexhibit991q22026.htm abc.xyz/investor/events/event-details/2026/2026-Q2-Earnings-Call-2026-GgTAq7Is0z/default.aspx ir.aboutamazon.com/news-release/news-release-details/2026/Amazon-com-Announces-Second-Quarter-Results www.prnewswire.com/news-releases/meta-reports-second-quarter-2026-results-302838214.html investor.oracle.com/investor-news/news-details/2026/Oracle-Announces-Record-Q4-and-FY-2026-Results-Driven-by-Cloud-Infrastructure--Cloud-Applications/default.aspx finance.yahoo.com/news/moody-flags-662-billion-risk-081000616.html finance.yahoo.com/news/ai-5-trillion-data-center-165808446.html www.morganstanley.com/content/dam/msdotcom/en/assets/pdfs/Research_Bridging-Data-Center-Gap.pdf
- 20 Consumer Reports analysis of residential electricity prices; U.S. Energy Information Administration, *Short-Term Energy Outlook*, August 2026; Monitoring Analytics, Q1 2026 State of the Market Report for PJM.
www.consumerreports.org/data-centers/ai-data-centers-impact-on-electric-bills-water-and-more-a1040338678 www.eia.gov/outlooks/steo/archives/aug26.pdf www.monitoringanalytics.com/reports/PJM_State_of_the_Market/2026/2026q1-som-pjm-sec1.pdf

Semiconductors and market structure

- 21 Futurum Intelligence public material on data-center GPU market share and on-site power. Public Q1 2025 wording is “more than 90%”/“95%” of the GPU market; the off-grid result is a survey expectation.
futurumgroup.com/press-release/data-center-semiconductor-market-is-set-to-exceed-500-billion-in-2029 info.futurumgroup.com/tfg-data-ctr-semi-spot-check futurumgroup.com/press-release/ai-grid-constraints-will-push-over-33-of-data-centers-off-grid-by-2030
- 22 No public source supports the previously cited 75–88% accelerator band, 92% 2023 comparison, training-above-90% series, or 60–75% inference series; those figures are not used. For distinct regional context only: Reuters reporting on IDC's China AI-accelerator-server market denominator.
www.reuters.com/world/china/chinese-chipmakers-claim-nearly-half-of-local-market-nvidias-lead-shrinks-idc-2026-04-01

Geopolitics

- 23 U.S. Bureau of Industry and Security, *Revision to License Review Policy for Advanced Computing Commodities*, announced January 13, 2026 and effective on Federal Register publication; Framework for Artificial Intelligence Diffusion announced January 13, 2025; rescission announced May 13, 2025; affiliate controls expanded September 29, 2025; and D:5/Macau guidance issued May 31, 2026. Section 232 advanced-computing-chip tariff: White House fact sheet dated January 14, 2026, with the 25% tariff applying from January 15, 2026.
www.bis.gov/press-release/department-commerce-revises-license-review-policy-semiconductors-exported-china www.bis.gov/press-release/biden-harris-administration-announces-regulatory-framework-responsible-diffusion-advanced-artificial www.bis.gov/press-release/department-commerce-announces-rescission-biden-era-artificial-intelligence-diffusion-rule-strengthens www.bis.gov/press-release/department-commerce-expands-entity-list-cover-affiliates-listed-entities www.bis.gov/media/documents/bis-guidance-may-31-2026.pdf www.whitehouse.gov/fact-sheets/2026/01/fact-sheet-president-donald-j-trump-takes-action-on-certain-advanced-computing-chips-to-protect-americas-economic-and-national-security
- 24 Rowan Wilkinson, Chatham House, “AI export controls are not the best bargaining chip,” April 29, 2026; and Francisco Javier Varela Sandoval, Isabella Wilkinson, Alex Krasodonski and Rowan Wilkinson, Chatham House, *How middle powers can weather US and Chinese AI dominance: The case for ‘sovereign AI’ strategies*, February 2026.
www.chathamhouse.org/2026/04/ai-export-controls-are-not-best-bargaining-chip www.chathamhouse.org/sites/default/files/2026-02/2026-02-13-sovereign-ai-strategies-sandoval-et-al.pdf

Governing framework

- 25 Alan Shimmel, *The Indispensability Trap: How Becoming Essential Caps the Fortune—From the Railroads to AI* (forthcoming September 2026; author-confirmed).
digitalcxo.com/article/the-big-short-meets-the-indispensability-trap

• ABOUT THIS REPORT

Colophon

Publication, authorship, evidence standard, art direction and production details for Edition Two.

AUTHOR

Alan Shimmel

Alan Shimmel is founder and chief executive of Techstrong Group, the media network behind DevOps.com, Security Boulevard, Cloud Native Now, Techstrong.ai and Techstrong TV. His book *The Indispensability Trap: How Becoming Essential Caps the Fortune, From the Railroads to AI* is published September 2026 and supplies the framework in Reflection Three.

PUBLICATION

The AI Mirror: Riches, Ruin and the Indispensability Trap

Techstrong Special Report, Edition Two
September 2026

EVIDENCE STANDARD

Every figure in this edition is drawn programmatically from a source-locked ledger of twenty-five entries. Figures are labeled either source-locked evidence, in which case every value on the graphic is traceable to a named primary document, or analytical construction, in which case the graphic carries no quantity at all. Company figures are printed with their exact fiscal periods, because those periods are not coterminous.

WHAT WAS REFRAMED

Where a number could not be verified against a primary source it was removed rather than estimated. The single accelerator-share bar used in the first edition was replaced by four separately defined chokepoint markets, because no public source supports an accelerator-share series. Metropolitan wage detail was dropped from the mismatch figure because the federal metropolitan tables could not be retrieved. Nothing was interpolated to fill a gap.

ART DIRECTION

Original cover and plate artwork produced on a single art direction: fractured mirror surfaces, near-black mineral grounds, warm paper, and three semantic accents for abundance, strain and distributed intelligence. Three new compositions were made for this edition, for the opening scene, the catalogue of past chokepoints and the sovereign fence line. All charts and data graphics are programmatic.

TYPOGRAPHY

Urbanist for the cover, display headlines, section labels, pull quotes and numerals.
Archivo for long-form text, captions, labels and charts.
Both families embedded.

CONTACT

techstronggroup.com
Editorial and sponsorship inquiries welcome.

FIGURE INDEX

- [01](#) Virginia writes a temporary, capped bill for the buildout (*evidence*)
- [02](#) Six rungs, and where the utopian case borrows capability (*construction*)
- [03](#) The humanoid deployment record, as disclosed (*evidence*)
- [04](#) Three tiers of displacement evidence, graded (*evidence*)
- [05](#) Where the capital is going, on each company's own clock (*evidence*)
- [06](#) The chokepoint is not one market. It is four, each defined differently (*evidence*)
- [07](#) Three tiers, not a war between cloud and edge (*construction*)
- [08](#) Financial exposure runs on three different clocks (*evidence*)
- [09](#) Agents inside the vanguard, with the methodology attached (*evidence*)
- [10](#) The backlash, in numbers that come from different places (*evidence*)
- [11](#) Five categories, five different instruments (*construction*)
- [12](#) The mismatch is national, and it is not a retraining problem (*evidence*)
- [13](#) Eighteen months of export-control churn (*evidence*)

RIGHTS

© 2026 Techstrong Group. All rights reserved. This report may be shared in its complete, unaltered form with attribution. Excerpts require written permission.

AI will influence what becomes possible. It will not decide who inherits the possibility.

THE AI MIRROR · TECHSTRONG SPECIAL REPORT · EDITION TWO · SEPTEMBER 2026

Reflection one: Abundance

A general purpose capability that compresses research, tutors every child and loosens the link between labor and survival.

Reflection two: Dispossession

The same technology, working exactly as promised, with the gains flowing to the owners of the capital stock and the costs landing on people who own nothing.

Reflection three: The Trap

Success so comprehensive that AI stops being scarce, and some of the companies priced as permanent tollbooths discover that indispensability and pricing power are not the same asset.



Techstrong Group is the media network behind DevOps.com, Security Boulevard, Cloud Native Now, Techstrong.ai and Techstrong TV. This edition is based on Alan Shimmel's research and on a source-locked evidence ledger of twenty-five primary and secondary references.

techstronggroup.com