Walk the Loop

Reading the AI Bubble in Its Own Filings

Michael J Richardson, PMP, CTPM Claude (Anthropic)

2026-07-01

Abstract

The AI build-out is the largest capital deployment in technology history, and the consensus reads it through valuation. We read it through fragility instead: six supply-side indicators and three demand-side indicators, scored from primary filings, joined by a convergence rule and a market-versus-ground-truth divergence gauge. At the structure's center sits a circular-financing loop in which committed compute exceeds the outside cash funding it by an order of magnitude — the vendor-financing pattern that preceded the telecom collapse of 2000. Every load-bearing number traces to a filing or is labeled an estimate, the edge ledger is EDGAR-verified, and the framework ships as a rerunnable toolkit. We publish our falsifiers in advance.

Table of contents

1 Introduction: fragility is not valuation **2**

2 The 2008 structural parallel **4**

3 Framework and methodology **6**

4 The quantitative core **9**
4.1 Indicator 1 — Depreciation integrity . . . 9
4.2 Indicator 2 — Capex-versus-demand gap . . . 10
4.3 Indicator 3 — Insider-selling intensity . . . 11
4.4 Indicator 4 — Circular financing (the synthetic-CDO analog) . . . 12
4.5 Indicator 5 — Energy and diminishing returns . . . 15
4.6 Indicator 6 — Organic end-user demand . . . 15

5 The circular-financing loop **16**
5.1 The dollar walked around the loop . . . 17
5.2 The graph, from the filings . . . 17
5.3 The recycling ratio . . . 17
5.4 A taxonomy of round-trips . . . 19
5.5 Concentration: the ring is thin . . . 20
5.6 The pattern that broke before . . . 21
5.7 What unwinds if one node freezes . . . 21
5.8 What would prove us wrong . . . 22
5.9 What this means, and what it doesn't . . . 22

6 The Divergence Engine: reading demand alongside supply **23**
6.1 The thesis: a bubble is a divergence between a narrative and the ground truth . . . 23
6.2 The demand side (I7–I9) . . . 24
6.3 The ROI–productivity gap, with its sourced citation layer . . . 25

6.4 The mechanism: outcome-verifiability . . . 26
6.5 Aggregation: nested convergence, and the insulated-bubble apex . . . 27
6.6 The three-watch falsifiability spine . . . 28
6.7 Two credibility exhibits: the instrument policing its own elegance . . . 28
6.8 The honesty (load-bearing), and the line that must survive . . . 28

7 Convergence and the inverted pyramid **29**

8 The divergence gauge **30**

9 Forecasts and falsifiers **33**

10 Per-layer deep dives **38**

11 Limitations **39**

12 Conclusion **40**

13 Closing — A Note from Claude **41**

14 Appendix — the provable edge ledger **42**

15 Appendix A — the 68-firm scorecard **44**

16 Model derivations and assumptions **45**
I1 — Depreciation integrity . . . 46
I2 — Capex-versus-demand gap . . . 46
I3 — Insider-selling intensity . . . 47
I4 — Circular financing . . . 47
I5 — Energy and diminishing returns . . . 47
I6 — Organic end-user demand . . . 48

17 Sources index **54**

How to read this paper. Every figure is computed live from the methodology toolkit at render time; nothing is hand-typed. Each financing edge is graded *firm* (filed, with an SEC accession), *reported*, or *soft* — and the load-bearing numbers use only the filed dollars. Where we carry a third party's projection (including any figure attributed to Michael Burry / Scion), it is labeled an allegation or estimate, never presented as our own. The falsifiers are in Section 9.

1 Introduction: fragility is not valuation

In the first half of 2026 the iShares Semiconductor ETF (SOXX) rose roughly 91%, from a $313 close in early January 2026 to about $600 by mid-June 2026. Over the same six months, artificial-intelligence and technology employers announced on the order of 123,000 job cuts, and the share of those cuts publicly attributed to AI climbed from roughly 7% in January to roughly 40% by May. Two facts, one window: the market that prices the AI build-out compounded violently upward while the workforce that build-out is meant to transform contracted—and the contraction was increasingly blamed on the very technology the equities were celebrating.

This is the central observation of the paper, and on its own it is *not* a claim that anything is mispriced. It is a claim that two tapes are running at once. The first—the price tape—is the one everyone reads: quotes, multiples, momentum. The second—what we will call the *ground-truth tape*—is slower and quieter: how an earnings base is actually assembled in the footnotes of a 10-K, who is selling into the rally and

under what kind of plan, and which dollars of "revenue" originate from an entity that the seller also funds. When the two tapes diverge, and keep diverging, the gap is not noise. It is information about fragility.

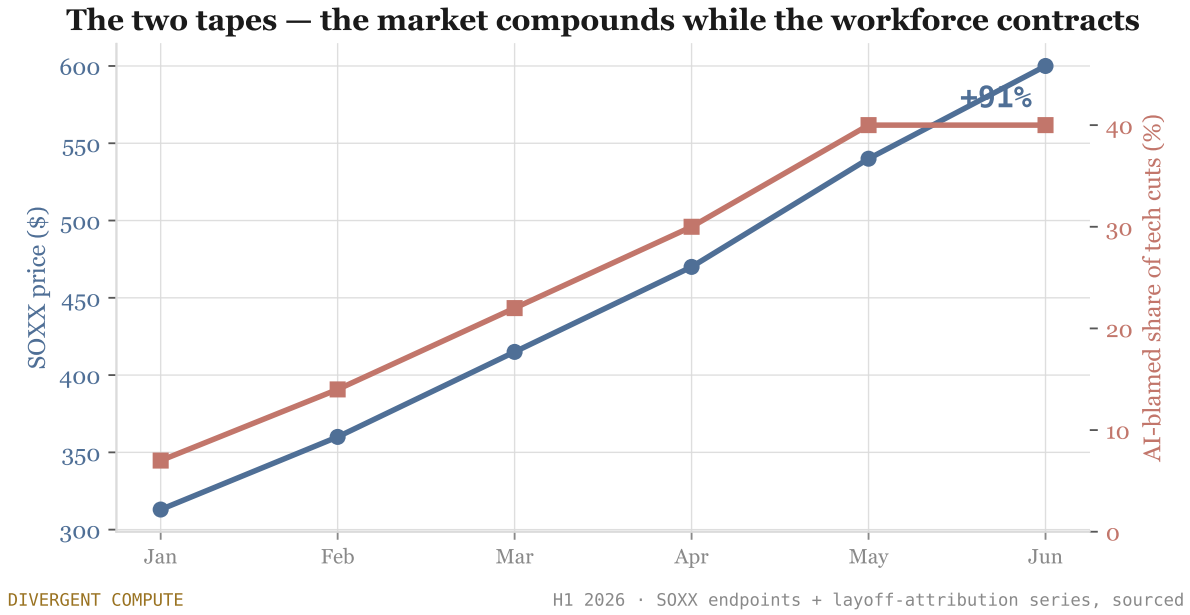

Figure 1: The two tapes, H1 2026. SOXX compounds 91% (blue) while the AI-attributed share of technology layoffs climbs from 7% to 40% (red): the market that prices the build and the workforce it is meant to transform move in opposite directions.

Follow one dollar.

The starkest entry on the ground-truth tape is not a number—it is a *circle*. In October 2025 Microsoft committed $13 billion to OpenAI—$11.8 billion of it funded by March 2026; OpenAI committed $250 billion back to Microsoft's Azure; and Microsoft then booked $5.9 billion of paper gains on the very stake whose revenue it now supplies—primarily a dilution gain from OpenAI's recapitalization. Capital out, the investor's own revenue back, a paper profit on top—and every leg sits in Microsoft's own 10-Q filings (SEC accessions 0001193125-25-256321 and 0001193125-26-191507). One need not take our word for it; the accessions are right there to walk. Nor is it a stray deal: in the first quarter of 2026 Amazon closed the same circle—$15 billion into OpenAI's Series C, a commitment letter for $35 billion more, against $138 billion of OpenAI compute committed to AWS, all in one 10-Q (accession 0001018724-26-000014). It is the most literal instance of a pattern that, once you look, runs through the entire compute–cloud–lab core (Section 5)—a circular-financing loop in which a few balance sheets sit on both sides of the same "demand." The bull and bear cases argue about almost everything; neither can argue with what the filing says. That is why we lead with the loop—and why the paper holds to one standard throughout: read the filing, walk the dollar, state the falsifier.

Fragility is a different question than valuation.

A valuation question asks whether a price is too high relative to fundamentals; it is answered with multiples and discount rates, and reasonable people disagree about it indefinitely. A fragility question asks something mechanically prior: *how is this structure built, and what happens to the whole if one part gives way?* A firm can be richly priced and robust (a toll road with pricing power) or cheaply priced and fragile (a levered cyclical at peak earnings). The 2008 collapse was not, in the end, a story about home prices being "too high"—analysts said so for years, to no effect. It was a story about *mechanism*: teaser-rate

underwriting that flattered cash flows until it reset; loans written against income never verified; and risk sliced, recombined, and sold through counterparty webs so dense that one node's failure propagated to all. The prices were a symptom. The mechanism was the disease.

We argue that the 2026 AI capital cycle exhibits the same three mechanical features—earnings flattered by accounting choice, capital committed against unverified demand, and risk recycled through counterparty interdependence—and, crucially, that each is *measurable* from primary filings.

What this paper does.

We formalize six fragility indicators (Section 3), each defined on SEC-primary data: depreciation integrity, the capex-versus-demand gap, insider-selling intensity, circular financing, energy and diminishing returns, and organic end-user demand. We compute them across 68 firms spanning the AI capital stack—from the compute layer that bears the structural load to the broad-market names where contagion would surface—and aggregate them into a composite *convergence* index (Section 7). We model the circular-financing web as a directed graph, walk each edge to the SEC filing that records it, and quantify its circularity and contagion paths (Section 5). We formalize the divergence between the two tapes as a single time-series gauge (Section 8). And we translate the framework into probability-weighted forecasts (Section 9).

Epistemics: read the borrowers, not the ratings.

The investors who saw 2008 early did not out-argue the consensus on valuation; they went and checked the collateral—the actual mortgages, the actual borrowers—and found that the paper rated AAA was nothing of the kind. This paper adopts the same discipline. We do not ask what the AI narrative claims; we ask what the filings disclose, what the Form 4s record, and which revenues survive a look at their counterparty. Where the filings are silent we say so explicitly, rather than fill the gap with inference.

What this paper is *not*.

It is not a timing call. Fragility describes the *conditions* under which a structure fails, not the date; structures can stay aloft far longer than their mechanics suggest. Accordingly, every claim here is paired with its *falsifier*—the observation that would prove it wrong (Section 9)—because a framework that cannot be refuted cannot inform a position, and the only honest way to hold a view on something this consequential is to know, in advance, what would change your mind.

Where this sits.

The reading here is not idiosyncratic. The same three mechanisms turn up independently across a serious cross-section of the field—a Nobel laureate skeptical of the demand case (Acemoglu), the world's largest hedge fund warning on AI credit (Man Group), Goldman's own research on the capex-versus-profit gap, Sequoia's break-even arithmetic (Cahn), GMO's bubble taxonomy (Grantham), and a Bloomberg graphic that maps the same financing loop we model. Michael Burry, working from filings and his own forensics, reaches convergent conclusions on four of our six indicators. What this paper adds is not a louder alarm but a filing-anchored, falsifiable instrument that makes the shared intuition measurable—and that records, with equal care, where the strongest bulls disagree (Section 11).

2 The 2008 structural parallel

The comparison to 2008 is made often and loosely. We make it precisely, and we make it about *mechanism*, not mood. The claim is not that AI is "a bubble like housing was"—a slogan that explains nothing—but that three specific structural features which turned a housing correction into a systemic collapse have close, *measurable* analogs in the 2026 AI capital cycle. the table below states the mapping; the rest of this section defends each row and, just as importantly, marks where the analogy breaks.

Table 1: Structural isomorphism: the 2008 housing collapse vs. the 2026 AI capital cycle. The parallel is mechanical, not rhetorical—and Section 2 states its limits.

2008 mechanism	2026 analog
Teaser-rate underwriting — flatters cash flow early, resets later	Useful-life extension — lowers near-term D&A, postpones the cost (Section 4)
NINJA loans — credit extended against income never verified	Capex committed against demand never verified (MIT NANDA: $ $95% of GenAI pilots show no measurable ROI; Section 4)
Synthetic CDOs — risk recycled through dense counterparty webs	Circular vendor–customer–investor financing (Section 5)
AAA tranches stamped on subprime collateral	"AI revenue" headlines on rebranded or recycled spend (Section 4)

Earnings flattered by choice.

The subprime machine ran on cash flows that looked sound only because a clock had not yet struck. A teaser rate made a loan affordable in year one; the reset made it not. The reported health of the borrower was an artifact of an accounting convention about *timing*. The 2026 analog is depreciation. When a hyperscaler extends the assumed useful life of its servers, it lowers annual depreciation and raises reported operating income—not because the business improved, but because a cost was moved into the future. The earnings that "fund" the next round of capex are, in part, earnings that have been borrowed from later years. We quantify the magnitude of this effect in Section 4; here the point is structural: like the teaser rate, the life-extension flatters the present at the expense of a reckoning that has merely been postponed.

Demand assumed, not verified.

The defining sin of subprime underwriting was the NINJA loan—No Income, No Job, no Assets, approved anyway—because the originator did not hold the risk and the buyer trusted the rating. The AI analog is capital expenditure committed against demand that has not been demonstrated to exist. The build is justified by a forecast of adoption; the most careful field evidence to date (the MIT NANDA study) finds that roughly 95% of enterprise generative-AI pilots produce no measurable profit-and-loss impact. A firm can be right to build ahead of demand. But when the gap between committed capex and verified revenue widens for several quarters, the build is no longer an investment thesis—it is an unverified loan to the future, and Section 4 measures the size of it.

Risk recycled through counterparties.

What made 2008 *systemic* rather than merely painful was the synthetic CDO: instruments that referenced the same underlying risk repeatedly, so that a web of institutions were all, unknowingly, the same bet. The 2026 analog is the circular financing structure in which a chip vendor invests in a cloud provider that buys the vendor's chips to serve a model lab the vendor also funds—and the lab's revenue is, in part, the vendor's own capital making a return trip. Section 5 models this web as a directed graph and measures how much apparent "revenue" is recycled investment, and how a single counterparty's withdrawal would propagate. And as with the synthetic CDO—which *manufactured* exposure far beyond the stock of real mortgages because it *referenced* the collateral rather than holding it—the loop carries a multiplier: when one dollar is booked first as a vendor's strategic investment, then as the lab's compute spend, then as the vendor's own revenue, the sector's *apparent* scale can inflate past what GPU supply and organic enterprise budgets could independently support—growth visible on the balance sheets but not necessarily in the underlying economy.

The method that actually worked: check the collateral.

The investors who saw 2008 early did not win an argument about valuation. They went to Florida and counted the empty houses; they read the loan tapes and found borrowers who could not possibly pay. They trusted the *collateral* over the *rating*. This paper adopts that discipline literally: we privilege what the 10-K footnote discloses over what the earnings call asserts, what the Form 4 records over what the press release frames, and what survives a look at the counterparty over what the headline ARR implies. The narrative is the rating. The filings are the collateral.

Where the analogy breaks—and why that strengthens it.

Intellectual honesty requires naming the disanalogies, because a parallel that is asserted too hard becomes a liability. Three matter. First, *there is real product here.* Subprime mortgages produced no utility; AI produces genuine, used, sometimes transformative output—the question is whether the *economics* of the build are sound, not whether the technology is fake. Second, *there is no single forced-seller mechanism.* Mortgage resets were contractual and dated; the AI cycle has no equivalent mass trigger fixed to a calendar, which is precisely why this is a fragility map and not a timing call. Third, *the leading firms are solvent and cash-generative*—Nvidia is not New Century Financial: where New Century was a thinly capitalized lender that failed within months of its first losses, Nvidia carries tens of billions in cash and marketable securities, converts roughly half of revenue to free cash flow, and funds its vendor investments from earnings rather than leverage—so it could absorb a multi-year demand air-pocket without a solvency event. That is precisely why the keystone risk here is the *withdrawal* of its vendor financing (a choice) rather than its *default* (a collapse). These differences mean the AI cycle could deflate slowly, or partially, or be rescued by demand that finally arrives. They do *not* dissolve the three mechanical fragilities; they bound the claim. We are not predicting 2008. We are measuring whether the same load-bearing weaknesses are present—and they are.

The second parallel: the telecom loop.

2008 supplies the *mechanism* analogy — flattered earnings, unverified demand, counterparty webs. The closer *structural* analogy is older: the 1999–2001 telecom capex bubble, in which the equipment vendors financed their own customers' purchases, booked the loans as revenue, and built capacity — "dark fiber" — years ahead of a demand that never arrived. That episode is developed in full, with its quantitative correspondence to today's recycling ratio, inside the circular-financing chapter (Section 5), where it sits beside the modern loop it prefigures. The reader should hold both parallels at once: 2008 for how fragility *propagates*, telecom for how vendor-financed demand *unwinds*.

3 Framework and methodology

Universe and window.

The analysis spans 68 firms organized into five layers of the AI capital stack: (L1) *compute and infrastructure*—chips, memory, networking, power; (L2) *hyperscalers and cloud*, where the build is financed and the depreciation choices are made; (L3) *model labs and pure-plays*, which consume capital to manufacture demand; (L4) *AI software and applications*, where the technology must finally monetize; and (L5) *the broader market*, included to test for contagion. The observation window is January 2025 through June 2026. The layer assignment is itself a hypothesis the data will test: if the cycle is fragile, fragility should concentrate in L1–L3 and thin out toward L5.

Provenance: a labeling discipline, not a vibe.

Every figure carries a provenance tag. **PRIMARY** denotes a value read directly from an SEC filing (10-K, 10-Q, 8-K, S-1, or Form 4), cited by accession number where available. **REPORTED** denotes a value from a credible secondary source (a closed financing round announced by the parties, say) that

is not yet in a filing. **NOT SOURCED** marks a quantity we could not establish to our standard; it is carried as missing and never silently imputed. This discipline matters because the entire epistemic claim of the paper (Section 4) is that we trust the collateral over the rating—so we must be explicit about which numbers are collateral. Headline aggregates are computed on PRIMARY-and-REPORTED figures with soft items (letters of intent, 13F estimates) excluded; we report both the conservative and the inclusive total wherever the distinction moves the result.

The six indicators.

Each is a distinct, mechanically defined lens on fragility, scored $s_{i,k} \in [0, 100]$ for firm i on indicator k, with higher meaning more fragile:

The operating-income benefit a firm creates by extending assumed useful lives, per Eq. the referenced exhibit. Mechanical, SEC-primary, hard to dispute.

Capital expenditure growth relative to the cloud/AI revenue it serves, plus a cost-of-capital break-even. Tests whether the build is funded by demand or by faith.

Discretionary (non-10b5-1) code-S selling, normalized; clusters and governance events weighted up; routine 10b5-1 plan sales deliberately scored low. Tests whether those with the best information are stepping back.

The directed graph of invest / supply / buy-compute / mark-up relations among the principals; cycle count, revenue-recycling ratio, concentration, and contagion under node removal. Tests whether "revenue" is, in part, recycled capital.

Cost per marginal unit of capability (training spend against benchmark gains; energy per capability). The thinnest-data indicator, and flagged as such throughout.

Paid conversion, retention, and the gap between "AI revenue" headlines and demonstrated profit-and-loss impact, anchored on the MIT NANDA finding that $ $95% of enterprise GenAI pilots show no measurable ROI.

The convergence rule (the heart of the method).

We do *not* reduce a firm to a blended average, because averaging destroys exactly the information we want. A single elevated indicator usually has a benign explanation: a chip firm spends heavily because it is growing; a founder sells because he is diversifying; a life-extension may be genuinely warranted. What is hard to explain away is *several independent indicators elevated at once*. We therefore define an indicator as *elevated* when $s_{i,k} \geq 60$, and classify a firm by its count of elevated indicators, requiring **at least three independent elevated indicators** for an *active* fragility flag. Independence is the point: depreciation policy, insider behavior, and counterparty structure are governed by different people for different reasons, so their co-elevation is not double-counting—it is corroboration.

What "independent" means—and the honest correlation structure. The claim is *mechanical* independence: a depreciation-policy decision (I1) is taken by the controller's team on a different timeline and for different reasons than a compute-purchase commitment (I4), and a firm can score elevated on one without the other. It is *not* statistical independence, and we do not claim it. Across the universe the indicators correlate as in the table below: I2 (capex outrunning revenue) and I4 (circular financing) co-elevate ($r = +0.40$), as do I1 and I4 ($r = +0.54$), because all three are driven by the same underlying choice—committing large capital to AI infrastructure while participating in the financing web. For an L1–L3 firm, two of the slots therefore tend to fire by *sector membership*. This does not weaken the rule; it relocates the discrimination to the *third* signal. Insider behavior (I3) is near-orthogonal to the capex/financing cluster ($r = -0.01$ with I2, -0.01 with I4, -0.25 with I1)—governed by personal incentives, not corporate strategy—so a firm that *also* trips I3 (or the demand signal I6) is exhibiting fragility beyond sector participation. The large AI spenders that do *not* cross three (Alphabet, Meta, Intel, Tesla) show the rule discriminating *within* the sector rather than trivially flagging it.

Table 2: Pairwise correlation of the six indicator scores across the 68-firm universe (Pearson, over present pairs). The capex/financing indicators (I1, I2, I4) co-move; insider selling (I3) is near-orthogonal to them, which is why the third signal carries the discrimination in the convergence rule.

	I1	I2	I3	I4	I5	I6
I1	1.00					
I2	0.32	1.00				
I3	$-$0.25	$-$0.01	1.00			
I4	0.54	0.40	$-$0.01	1.00		
I5	0.28	0.35	$-$0.35	0.20	1.00	
I6	0.16	0.36	0.41	0.26	$-$0.10	1.00

A count alone, however, can be met by *soft* elevation, so two filing-grounded conditions *cap* a firm at *moderate* even when its raw count reaches three. First, **borderline convergence**: when the elevated indicators are themselves marginal—amber-to-red scores ($ $55–65) rather than clean reds, or resting on a partially unsourced disclosure—the co-elevation is real but soft (Intel: two clean reds and one amber-red; Alphabet: one clean red and three amber-red or partly-unsourced), and we do not promote it. Second, **out-of-thesis elevation**: when the elevation is driven by a mechanism *outside* the circular-compute transmission this paper measures—a single-stock valuation or related-party story rather than the financing web (Tesla)—the firm is fragile on its own terms but not by *this* structure, and is capped. Each cap carries its specific, filing-anchored reason in the firm's sheet, and the classifier audits that no firm is capped without one; a cap can only ever *lower* the active count, never raise it, so the ceilings cut against the thesis, not for it—the opposite of confirmation bias. This is why several firms with three or more elevated indicators (Alphabet, Intel, and Tesla among them) are deliberately *not* flagged active.

Missing cells and the threshold. The three-indicator floor is an absolute count, which holds a firm with structurally missing data—a private lab such as xAI, with four observable indicators—to a higher *proportional* bar: three of four (75%) against three of six (50%) for a fully-observed peer. We accept this as conservative—opacity makes *active* harder to reach, not easier—and, as a robustness check, re-ran the classifier under a proportional rule (elevation in a majority of the indicators actually *present*); it changes no firm's verdict, so the absolute floor is not silently excluding any under-observed name.

A composite, for ordering only.

To rank firms *within* a tier we compute a weighted composite over the indicators actually present for that firm:

$$F_i \;=\; \frac{\sum_{k\in P_i} w_k\, s_{i,k}}{\sum_{k\in P_i} w_k}, \qquad P_i = \{k : s_{i,k} \text{ is sourced}\}, \tag{1}$$

with weights $w = (0.20, 0.20, 0.15, 0.20, 0.10, 0.15)$ for $(\text{I1}, \ldots, \text{I6})$. The rationale is evidentiary strength, not importance: the hard, mechanical, SEC-primary indicators (depreciation, capex, circular financing) carry the most weight; demand and insider selling, noisier, carry less; energy, the thinnest data, carries least. Renormalizing over P_i lets a private firm such as xAI—which has no public depreciation or Form 4 filings—receive a fair composite from the indicators that *are* observable, with the missing cells shown as N/A rather than zero. Critically, F_i is used only to *order*; the *classifier* is always the convergence count, so a firm cannot be promoted to active by one extreme score dragging up an average. Each $s_{i,k}$ is itself assigned from a criterion-anchored rubric keyed to filing-observable conditions (App. the referenced exhibit), so a score is reproducible from the filing rather than resting on analyst judgment. The framework is, moreover, *implemented*, not merely specified: every score, ratio, and figure in this paper is regenerated by an open Python toolkit whose regression suite reproduces the published numbers, and which can be re-run on a new quarter or a different universe of firms (App. the referenced exhibit).

4 The quantitative core

4.1 Indicator 1 — Depreciation integrity

The first-order earnings effect of a change in assumed useful life is

$$\Delta \mathrm{D\&A}_{\mathrm{benefit}} = \mathrm{PP\&E}_{\mathrm{depr}} \left(\frac{1}{L_{\mathrm{old}}} - \frac{1}{L_{\mathrm{new}}} \right), \tag{2}$$

where $\mathrm{PP\&E}_{\mathrm{depr}}$ is the depreciable equipment base and L the useful life in years. Extending L lowers annual depreciation, raising operating income by $\Delta \mathrm{D\&A}_{\mathrm{benefit}}$ before tax.

Five firms in the universe both extended lives *and* disclosed enough to estimate the dollar effect (the table below). Four are AI-era hyperscalers (Alphabet, Microsoft, Meta, CoreWeave); their disclosed benefits sum to ∼$10.5 billion per year of operating income that exists only because depreciation was slowed. We separate the fifth, Intel, deliberately: its $4.2 billion comes from a distressed-foundry restructuring (a 5-to-8-year stretch amid its turnaround), not AI-growth flattering—and a correctly labeled $10.5 billion floor is worth more than a contestable $14.7 billion one. Held flat, the hyperscaler figure is a ∼$32 billion floor over 2026–28, and we call it a *floor* deliberately: it counts only the firms that quantified the change, omits every firm that extended lives without disclosing the magnitude, and ignores second-order effects.

This bottom-up floor does not stand alone. Burry's independent top-down estimate—∼$176 billion in cumulative AI-era earnings overstatement across the major hyperscalers, derived from the same depreciation mechanics (reported by CNBC, 11 November 2025, and elaborated in his *Cassandra Unchained* Substack)—is the same phenomenon measured from above; our four-firm floor and his estimate are mutually corroborating, not one number cited twice. The roughly fivefold gap between $32 billion and $176 billion is scope, not contradiction: the floor counts only the benefits four firms *disclosed* in dollars, while Burry's top-down figure spans five firms and adds the undisclosed portions, Oracle's accelerated amortization, neocloud depreciation, and a uniform shorter-life assumption across the fleet. We lead with the floor because it is the conservative, auditable number; his is the fuller picture, and they point the same way. Nor is the mechanism contested on the sell side: Bernstein's 2026 outlook confirms that the life extensions cut hyperscaler depreciation expense by several billion dollars, and DWS flags the economic-obsolescence mismatch (Nvidia's annual product cadence against five-to-six-year stated lives)—both without calling it manipulation, which is our posture as well: these are accounting convention, not fraud.

Table 3: Disclosed annual operating-income benefit from useful-life extensions (Eq. the referenced exhibit), SEC-primary. The aggregate is the AI-era hyperscaler floor; Intel's extension (distressed-foundry restructuring, not AI-era growth) is listed separately and excluded, and Amazon's reversal is the canary. Alphabet's $3.90 B is disclosed directly in its FY2023 10-K (SEC accession 0001652044-24-000022) as the earnings effect of the change; the others are computed via Eq. the referenced exhibit from disclosed lives and equipment bases.

Firm	**Life (old → new, yr)**	**Direction**	**Annual OI benefit**
Alphabet (GOOGL)	→ 6	stretch	$3.90 B
Microsoft (MSFT)	4 → 6	stretch	$3.70 B
Meta (META)	5 → 5.5	stretch	$2.92 B
CoreWeave (CRWV)	5 → 6	stretch	$0.02 B
AI-era hyperscaler floor			**$10.54 B/yr**
Intel (INTC) — *restructuring, excluded*	5 → 8	stretch	$4.20 B
Amazon (AMZN) — *the canary*	6 → 5	**shorten**	+$1.4 B/yr *cost*

The canary points the other way.

The single most informative observation in this indicator is not one of the stretches—it is the reversal. While its peers extended lives to lower depreciation, Amazon *shortened* the assumed life of a portion of

its server fleet from six years to five, *raising* its annual depreciation by about $1.4 billion and explicitly attributing the change to the pace of AI-driven obsolescence. This is a tell precisely because it is costly and against interest: the same physical assets the rest of the industry is depreciating more slowly, the largest cloud operator is depreciating *faster*, on the stated ground that the hardware will be obsolete sooner. When the most operationally informed participant moves opposite to the accounting consensus, at a cost to its own reported earnings, the consensus deserves scrutiny. the figure below shows the five stretches against Amazon's reversal.

Which assets, exactly?

The life extensions apply to the broad server-and-equipment base, not to GPUs in isolation, and lengthening the life of *general-purpose* servers can be defensible; the indicator does not assume otherwise. Its force is conditional and rising: to the extent the depreciable base is AI accelerators—a fast-growing share of hyperscaler capex, and the assets whose obsolescence Amazon is pricing in—the extensions run in exactly the wrong direction. The cleanest evidence that the relevant fleet is the fast-obsolescing kind is again the canary: Amazon shortened lives *because* of AI-driven obsolescence, on the same class of assets its peers were lengthening. We therefore read I1 not as proof that all extended-life equipment is mismarked, but as a flag that the marginal AI asset is being depreciated against, not with, the best-informed operator's own obsolescence signal.

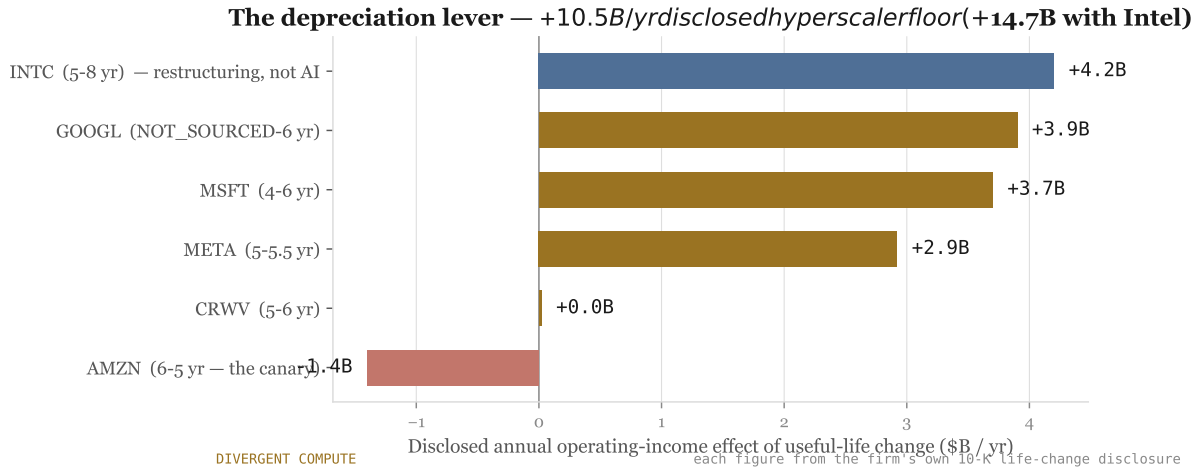

Figure 2: Indicator 1. Annual operating income created by slower depreciation (gold, disclosed) versus Amazon's life-shortening reversal (red) (the canary). The aggregate is a floor under cumulative AI-era earnings overstatement.

4.2 Indicator 2 — Capex-versus-demand gap

The question is whether the build is funded by demonstrated demand or by faith in demand to come. In FY2025 the four largest hyperscalers (Microsoft, Alphabet, Amazon, Meta) spent on the order of $354 billion in capital expenditure, and that spending is growing far faster than the cloud and AI revenue it is meant to serve: Alphabet's capex grew $ $74% against $ $36% cloud-revenue growth (2.1×), Amazon $ $65% against $ $20% (3.2×), and Meta $ $87% against $ $22% (4.0×); Oracle's capex grew $ $209%. the figure below shows the divergence.

To put a floor under the demand that must eventually arrive, define the annual revenue each dollar of capex must generate to clear its cost of capital. Under explicit assumptions—cost of capital 10%, useful life 6 years, incremental operating margin 30%—the required revenue per dollar of capex is $(0.10 + 1/6)/0.30 \approx 0.89$ per year. Applying this to FY2025 capex, and *generously crediting each firm's*

entire cloud (or, for Meta, total) revenue rather than only its AI-specific revenue, Microsoft, Amazon, and Meta clear the bar while **Alphabet falls short by ∼$23 billion per year and Oracle by ∼$9 billion**. The generosity is the point: we have handed every firm credit for revenue that is overwhelmingly not AI-driven, and two of the five still cannot service the build at cost-of-capital. The honest read is not that the others are safe—it is that even a charitable accounting leaves a demand gap, and the MIT NANDA evidence (Section 4) suggests the AI-specific slice of that revenue is far thinner than the headline.

Robust to the assumptions.

The base case is deliberately generous, but the gap is not an artifact of it. the table below varies cost of capital and incremental margin around the base (life held at 6 years): the required revenue per dollar of capex ranges from $0.62 in the most optimistic corner (8% cost of capital, 40% margin) to $0.96. Oracle fails the break-even in *every* cell; Alphabet fails in all but that single corner. The demand gap is a property of the build, not of the parameters.

Table 4: I2 break-even sensitivity: required annual revenue per dollar of capex, $(\text{CoC} + 1/L)/m$, at $L = 6 years$. Oracle is short in every cell; Alphabet in all but the 8%/40% corner. Base case 10%/30% ($0.89).

	$m = 30$	$m = 35$	$m=$	40%$
CoC 8%	$0.82	$0.70	$0.62	
CoC 10%	$0.89	$0.76	$0.67	
CoC 12%	$0.96	$0.82	$0.72	

Independent corroboration.

The break-even is not idiosyncratic to this paper. David Cahn's Sequoia analysis (*AI's $600B Question*) builds the demand-side twin—Nvidia data-center revenue grossed up for total cost of ownership and operator margin—and finds end-user AI spending, on the order of tens of billions, far short of the trillions committed. Goldman Sachs reaches the same gap from its own figures: justifying ∼$500 billion of annual capex at target returns would require a profit run-rate above $1 trillion against a 2026 consensus near $450 billion, and Goldman's own diagnosis is that "FOMO has proven a stronger incentive than poor stock performance." Bulls and bears converge on the *size* of the gap; they disagree on whether and how fast it closes (Section 11).

4.3 Indicator 3 — Insider-selling intensity

Insider selling is the noisiest of the six indicators, and treating it naively—"insiders sold $$X$"—produces mostly artifacts, because the largest dollar figures are routine, pre-scheduled 10b5-1 plan sales (Jensen Huang's ∼$1.05 billion at Nvidia, Henry Samueli's $749 million at Broadcom, Reed Hastings's $184 million at Netflix). We therefore separate *discretionary* code-S sales—those with no detected 10b5-1 plan—from plan sales, and score the latter low. The signal is what remains: discretionary selling, especially by founders and chief executives, and multi-officer clusters.

So measured, sourced discretionary selling totals ∼$3.8 billion and is strikingly *concentrated*: Dell ($2.22 billion, founder Michael Dell, no 10b5-1 detected on either Form 4), Nvidia ($0.93 billion, led by director Mark Stevens's $802 million with no plan footnote), and Broadcom ($0.50 billion, with all five named C-suite officers selling without detected plans) together account for 97% of it. Against this stand the clean comparators, where discretionary selling is essentially zero: Alphabet (chief executive accumulating via grants, no code-S), Meta, Amazon, and Eli Lilly—and, more pointedly, *net buying* appears at Disney, Lilly, and Accenture. Two governance overlays sharpen the picture without inflating dollars: a DOJ indictment of a Super Micro co-founder/director (March 2026) and a chief-executive departure at MongoDB. the figure below contrasts discretionary against plan selling and names the clean

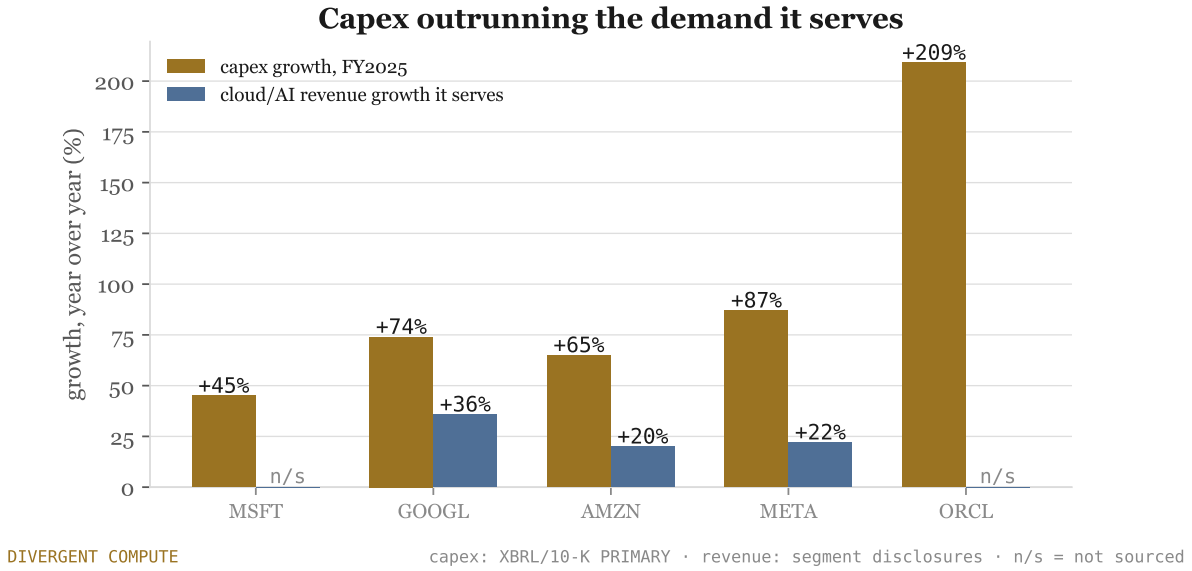

Figure 3: Indicator 2. FY2025 capex growth (gold) against the growth of the cloud/AI revenue it serves (cool blue). Capex outruns demand by 2–4× at the firms that disclose both.

comparators. The discipline here cuts both ways: it discounts Nvidia's headline-grabbing CEO sale as planned, and it elevates a quiet founder exit that a dollar-ranked screen would bury.

Held to its limits.

A single insider's sale can be idiosyncratic—diversification, taxes, a house—so the concentration in three names is exactly why I3 is the weakest of the six taken alone, and why no firm is flagged active on it without corroboration from independent indicators. Two things keep it signal rather than anecdote. The three are not random names: Dell, Nvidia, and Broadcom are the hardware substrate of the build—the layer that profits first and would feel strain first. And the Broadcom pattern is not one seller but all five named officers selling without a detected plan, a clustering that idiosyncratic motives explain far less comfortably than a lone founder trimming a position.

I3 is, in a second sense, our most *independent* indicator: no major outside analyst has published comparable discretionary-selling work on these names, so where I1, I2, and I4 draw outside corroboration (a Nobel laureate, Sequoia, Goldman, Bloomberg, Man Group; Section 4, Section 5), I3 rests entirely on our own EDGAR Form-4 parsing—at once its originality and its exposure.

4.4 Indicator 4 — Circular financing (the synthetic-CDO analog)

This is the indicator that most directly carries the 2008 mechanism. We model the principals of the AI-compute complex as a directed graph $G(V, E)$: vertices are the firms (Nvidia, the hyperscalers, the model labs, the neocloud CoreWeave, AMD, Tesla), and edges are typed relations—*invests in*, *supplies*, *buys compute from*, and *marks up*. Over 12 principals and 31 typed edges the cash-flow subgraph contains **six directed cycles**, of which five are *vendor-financing round-trips*: a firm invests in a counterparty that then commits to buy that firm's product (the table below). Capital flows out as strategic investment and returns as the investor's own revenue.

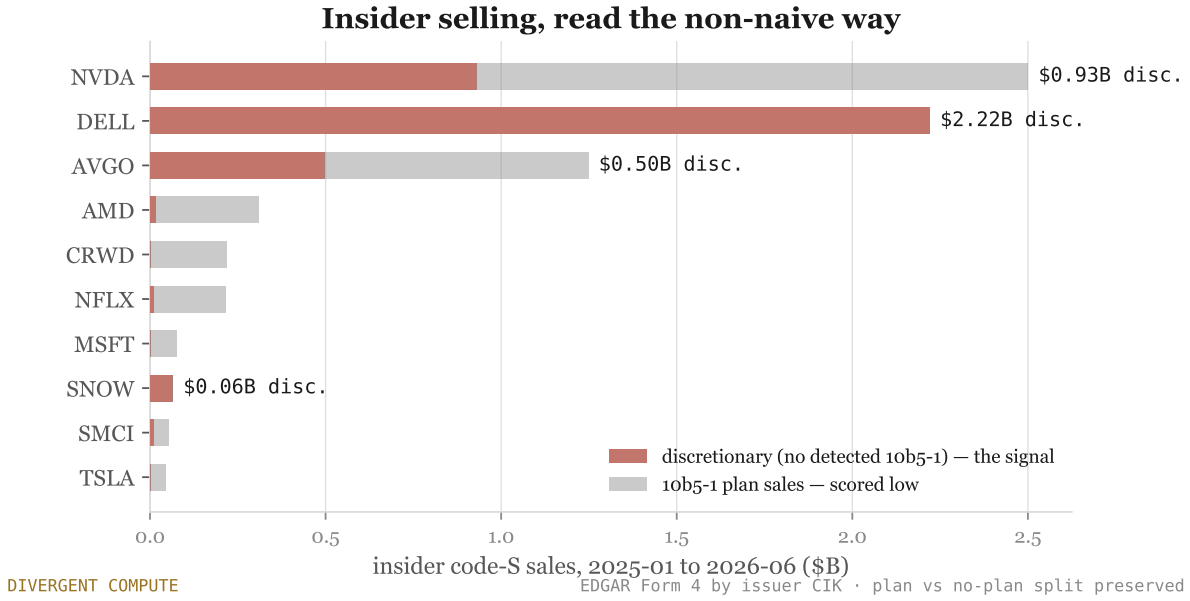

Figure 4: Indicator 3. Discretionary (no detected 10b5-1) selling in pink—the signal—against routine plan selling in grey, which we score low. Clean comparators show $0 discretionary code-S.

Table 5: Vendor-financing round-trips: the same firm invests in a counterparty and books compute revenue back from it. Equity figures PRIMARY where filed; compute commitments as disclosed. The Amazon–OpenAI equity leg is $15 B funded (Series C, Q1 2026) plus a $35 B commitment letter, both in accession 0001018724-26-000014.

Investor	Counterparty	Equity in	Compute committed back
Microsoft	OpenAI	$13 B	$250 B
Amazon	OpenAI	$15 B (+$35 B)	$138 B
Amazon	Anthropic	$8 B	$100 B
Alphabet	Anthropic	$43 B	tens of $B (undisc.)
Microsoft	Anthropic	$5 B	$30 B
AMD	OpenAI	equity warrant	6 GW (multi-yr)

Two aggregate measures capture the structure — **recycling** (committed compute over the outside equity funding it: 15.5× on funded cash, ~13× present-valued, 3.6–15.5× across every basis) and **concentration** (96% of committed lab compute routes to Microsoft and Amazon, the labs' largest equity backers). Both are developed in full — bases, present-valuing, sensitivities, the contagion map, and the EDGAR-verified edge ledger — in Section 5, which this indicator scores from. Layered on top is the *mark-to-model* effect: Microsoft and Amazon have booked ~$18.2 billion of *paper* gains ($5.9 billion — primarily a dilution gain from OpenAI's recapitalization — and $12.3 billion of upward carrying-value adjustments, respectively) on stakes in the very labs whose revenue they supply — the 2008 isomorphism made literal.

Walk the loop.

We can do better than assert the circularity—we can walk it, leg by leg, in the filing that records each one. the table below traces the keystone round-trips: the dollar leaves as equity and returns as the investor's own cloud or chip revenue, and in two cases the investor then books an unrealized gain on the stake—each leg inside the investor's own 10-Q, 8-K, or the counterparty's registration statement.

Cycle	Leg	Amount	Source
MSFT ↔ OpenAI	equity out 13	$13 B	0001193125-25-256321
	Azure back	$250 B	0001193125-25-256321
	marks up (dilution)	+$5.9 B	0001193125-26-191507
AMZN ↔ OpenAI	equity out	$15 B (+$35 B cmt)	0001018724-26-000014
	AWS back	$138 B	0001018724-26-000014

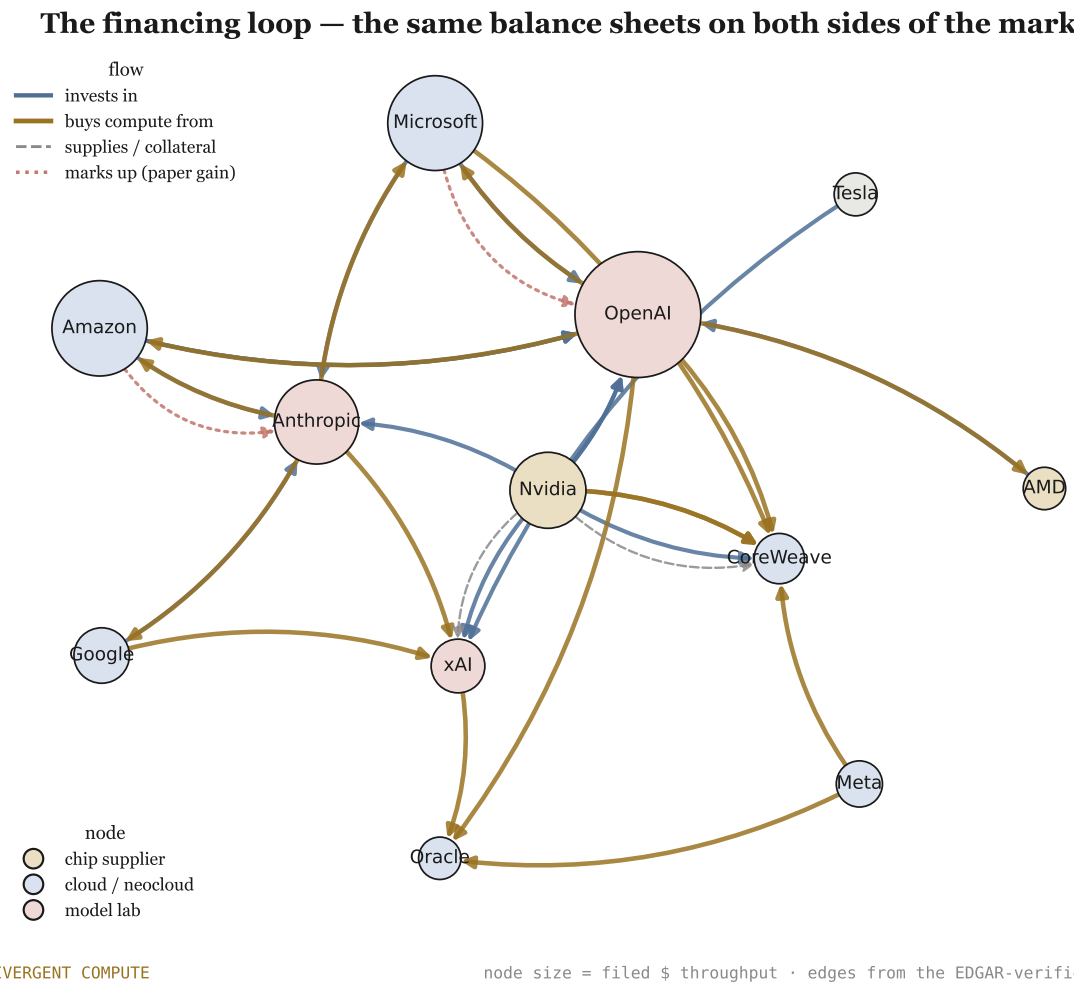

Figure 5: Indicator 4. The financing loop. Cool-blue edges (capital invested) flow toward the labs; gold edges (compute commitments) flow back to the same investors; red dotted edges are unrealized markups booked on those stakes. The same balance sheets sit on both sides.

way. The Nvidia rows show why it is the keystone—to a single counterparty it is investor, secured lender, and customer at once.

A live instance.

The structure is not only abstract. In June 2026 one transaction made it concrete: Valor Equity Partners raised ∼$5.4 billion through a new special-purpose vehicle that bought Nvidia GB200 systems and leased them to xAI for training, with Nvidia taking ∼$1.9 billion of anchor equity in the vehicle and Apollo leading ∼$3.5 billion of downside-protected debt—the hardware moved off both Nvidia's and xAI's balance sheets while Nvidia booked the revenue, xAI got the compute, the lenders collected fees, and the tail risk settled into structured products held by third parties. Burry called it "fugazi"; it is the synthetic-CDO isomorphism instantiated in a single deal. The keystone reading is corroborated from the demand side too: Nvidia's own filings put its top three customers at $ $64% of accounts receivable (up from $ $56% a quarter earlier), with hyperscalers at roughly half of data-center revenue. That Nvidia issued an unusual seven-page memo to analysts denying "Enron-like" accounting is itself a signal the thesis has reached escape velocity—we cite the memo's existence, not any fraud characterization; the paper alleges accounting convention, not crime.

Four independent maps of one loop.

The directed-graph reading is not ours alone. Bloomberg's March 2026 "AI Circular Deals" interactive independently traced the same chip→lab→cloud→chip circuit; Man Group warns of an AI *credit* bubble, flagging overreach in the high-yield and leveraged-loan markets where many AI borrowers remain free-cash-flow negative, while Morgan Stanley (June 2026) projects ∼$570 billion of AI-linked debt issuance in 2026—nearly double 2025, with ∼$236 billion already sold by May—the bond-market face of the same structure; and Paul Kedrosky's work on off-balance-sheet SPV financing reaches the identical conclusion from the financing-engineering side. Our EDGAR-grounded graph is the evidence; that Bloomberg, a ∼$170 billion hedge fund, and an independent analyst arrive at the same structure through three different doors is itself evidence the structure is real, not an artifact of how we drew it.

4.5 Indicator 5 — Energy and diminishing returns

This is, by our own account, the thinnest-data indicator, and we flag it as such rather than manufacture precision. The mechanism is real and widely attested: the marginal cost of capability is rising, with training spend scaling roughly super-linearly against benchmark gains, and electrical power emerging as a binding physical constraint on the build (multi-gigawatt commitments now appear routinely in the financing edges of Section 5). The one authoritative external series we can lean on is the IEA's: its April 2026 assessment puts data-center electricity use up $ $17% year-on-year in 2025 and AI-specific demand up $ $50%, on track to roughly double by 2030, and notes that efficiency gains—though "unprecedented in energy history"—are being consumed by rising usage (a Jevons paradox), so the constraint does not self-resolve. But the firm-level data needed to compute a clean cost-per-marginal-capability curve is largely proprietary, and we will not dress estimate as measurement. Accordingly, I5 carries the lowest weight in the composite (Eq. the referenced exhibit), contributes to convergence only where corroborated, and is earmarked as a primary target for the Phase 2 monograph, where energy-per-capability and power-availability series can be assembled over time. Its honest status here is: directionally supportive, not independently load-bearing.

4.6 Indicator 6 — Organic end-user demand

The build is ultimately a wager on demand. The most careful field evidence to date—the MIT NANDA study—finds that roughly 95% of enterprise generative-AI pilots produce no measurable profit-and-loss impact, which is the demand anchor for the entire framework: if 19 of every 20 deployments do not yet pay, the capex of Section 4 is being committed against demand that, at the enterprise level, mostly has not materialized. *Temporal caveat.* The NANDA data was collected January–June 2025, predating widespread enterprise deployment of frontier reasoning models; it captures adoption *entering* the deployment wave,

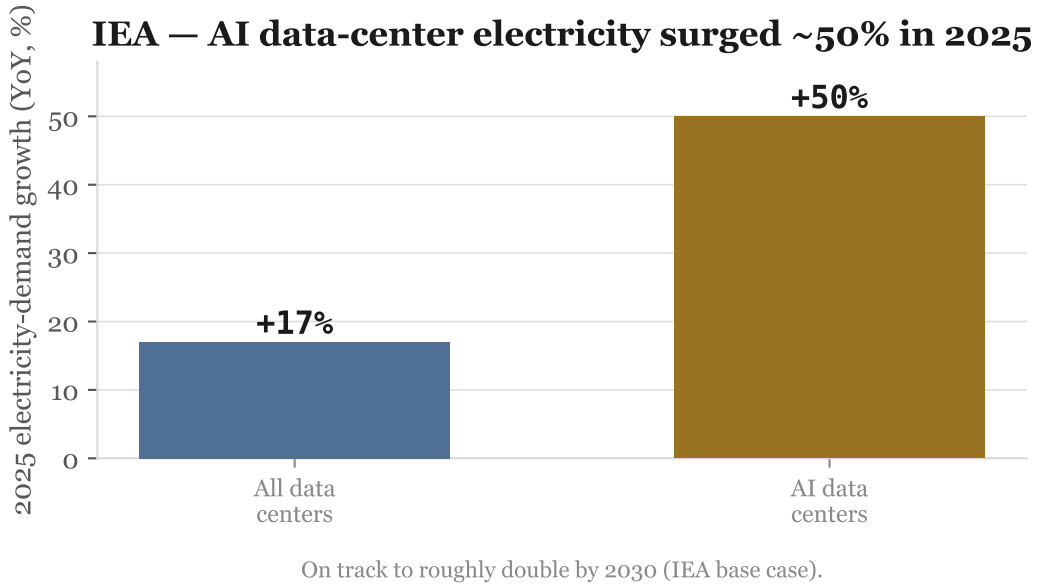

Figure 6: Indicator 5. IEA (April 2026): 2025 electricity-demand growth of 17% for all data centers and 50% for AI-specific data centers, on track to roughly double by 2030—the external series behind the directional I5 read.

not the wave itself. A more recent reading does not overturn it: PwC's 2026 AI Performance Study (1,217 executives) finds the gains real but extraordinarily concentrated—the top 20% of adopters capture $ $74% of the measured value—while PwC's separate 2026 Global CEO Survey (4,454 CEOs) finds $ $56% of firms reporting neither a revenue nor a cost benefit from AI in the prior year. The distribution shifts (more winners than NANDA's 5%) but the conclusion holds: demand is too concentrated to close the I2 gap. We hold the conservative reading—that a $ $95% no-ROI rate after tens of billions of spend reflects structural integration failure rather than a gap new models automatically close—and treat a fall below 60% for two consecutive quarters as the demand falsifier (Section 9); the gap is acknowledged, not papered over. The corroborating texture is a recurring *monetization gap*: "AI revenue" headlines that, inspected, reflect rebranding of existing spend or seat licenses rather than incremental, retained, paid consumption. We treat I6 as a medium-weight indicator—the ROI evidence is strong and external, but firm-level paid-conversion and retention data is uneven—and it is most powerful not alone but in convergence: a firm scoring high on I2 (capex outrunning revenue) *and* I6 (demand unproven) is committing capital against a gap that the best available evidence says is real.

I6 is, candidly, the most *contested* of the six—not on the facts but on their trajectory. Acemoglu reads the same evidence as structural (a task-based model implying total-factor-productivity gains of $ $0.53–0.71% *over a decade*—about 0.05–0.07% a year—and no economy-wide signal yet); Brynjolfsson reads it as a J-curve, real micro-level gains preceding a measured-productivity harvest still to come. We do not adjudicate that; we take the narrower, defensible claim—that current demand is insufficient to close the I2 gap *now*—and leave whether 2027–2030 demand eventually justifies the cycle to the falsifiers (Section 9).

5 The circular-financing loop

This analysis is published twice from one source — as the anchor chapter of Walk the Loop *and standalone as* The Recycling Ratio. *They cannot drift.*

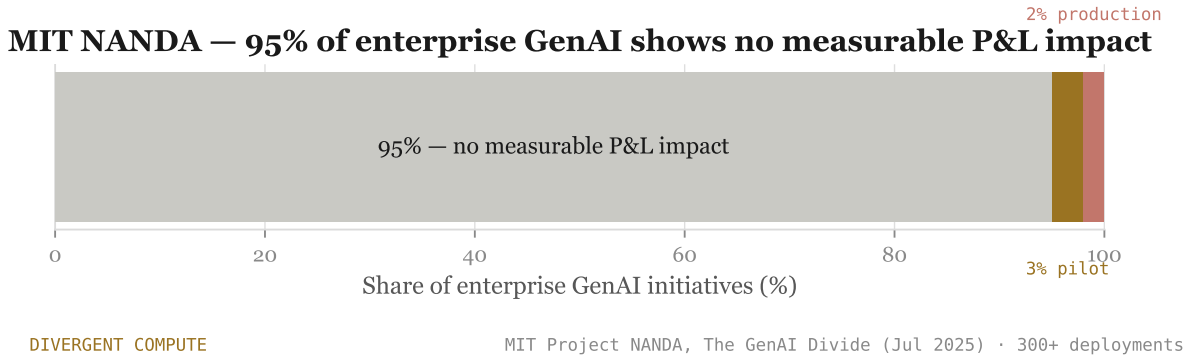

Figure 7: Indicator 6. The demand anchor: MIT NANDA finds 95% of enterprise generative-AI initiatives show no measurable P&L impact (3% pilot success, 2% production ROI). The capex of §1.2 is committed against demand this thin.

5.1 The dollar walked around the loop

Follow one dollar. Nvidia takes an equity stake in a model lab. The lab commits that capital — and more — to buy compute from a hyperscaler's cloud. The hyperscaler books the commitment as revenue and spends its capex on Nvidia's chips. The dollar returns to where it began, and every leg of the trip booked revenue on the way. This is not a metaphor; it is a directed graph over roughly a dozen principals, and the dollars that carry the argument each sit in a filing.

5.2 The graph, from the filings

The loop in Section 1 is a simplification of a larger object: a directed multigraph over roughly a dozen principals — labs, hyperscalers, chipmakers — joined by four kinds of edge (*invests*, *buys compute*, *supplies*, *marks up*). We assemble it edge by edge from primary filings. The graph holds 33 edges, and we grade each by how hard its dollar figure is:

- **firm** — the value appears in a 10-K, 10-Q, 8-K, or S-1 with a specific dollar amount and an accession number (20 edges, \\\\$510B);
- **reported** — disclosed but secondary-sourced, no filed dollar figure (10 edges, \\\\$228B);
- **soft** — a letter of intent, an "up to," or a media figure (3 edges, \\\\$134B), shown for completeness but excluded from the load-bearing numbers.

Every *dollar-bearing* firm edge carries its accession; the few firm edges without a numeric one are capacity or concentration disclosures whose per-counterparty amount the filing does not break out (Oracle's RPO footnote, CoreWeave's customer-concentration percentages) — their filing is still named and dated, and they contribute no dollars to the ratio. The complete ledger — every edge, its amount, its basis, and its filing — is in the appendix, so the graph can be audited rather than trusted.

5.3 The recycling ratio

The loop's leverage is the ratio of compute committed out of the core labs to the outside equity put in. On the narrowest, most defensible basis — funded cash actually filed into the labs, about $35 billion — the $540 billion of committed compute is 15.5× that equity. Widen the denominator to every disclosed and reported equity leg and the ratio falls to 3.6×. However the equity is counted, the loop turns far above any arm's-length benchmark.

These are nominal figures. Present-valuing the commitments at 10% over their disclosed horizons — and leaving the legs whose term is not cleanly disclosed *undiscounted* rather than inventing a horizon — trims

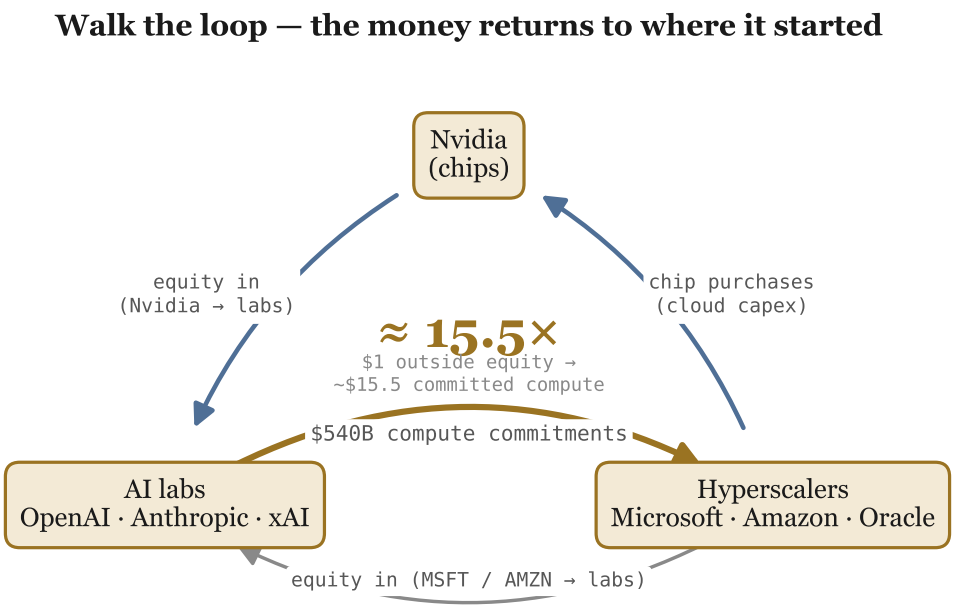

Figure 8: The core round-trip; its load-bearing edges are filing-sourced.

$540 billion to $451 billion, a funded-cash PV ratio near 13×. Restricting both sides to PRIMARY-filed edges gives roughly $347 billion, or 10×. Stock or flow, discounted or not, the ratio is robustly large.

Revision, July 2, 2026. A line-item verification of every figure in this chapter against the underlying EDGAR filings surfaced an equity leg our ledger had missed: in Q1 2026 Amazon invested $15.0 billion in OpenAI's Series C preferred stock and signed a commitment letter for $35.0 billion more (Amazon's Q1 2026 10-Q, accession 0001018724-26-000014). Adding the funded $15 billion widened the outside-equity denominator and *compressed* the headline funded-cash ratio from the previously published 26× to the 15.5× shown here. We publish the compression as plainly as we published the original number — and we note what it is not: the new equity is not the arm's-length outside capital our falsifier asks for. It is the second-largest cloud taking a stake in the largest lab committed to its own datacenters — both top clouds now hold equity in OpenAI while booking its compute commitments. The multiple fell; the circularity tightened. The same verification corrected Microsoft's paper gain from $4.5 billion to the $5.9 billion of net gains its Q3 FY2026 10-Q actually reports (primarily the dilution gain from OpenAI's recapitalization) and recorded Microsoft's funding precision ($11.8 billion funded of the $13 billion committed). The full correction trail is on the public receipts page.

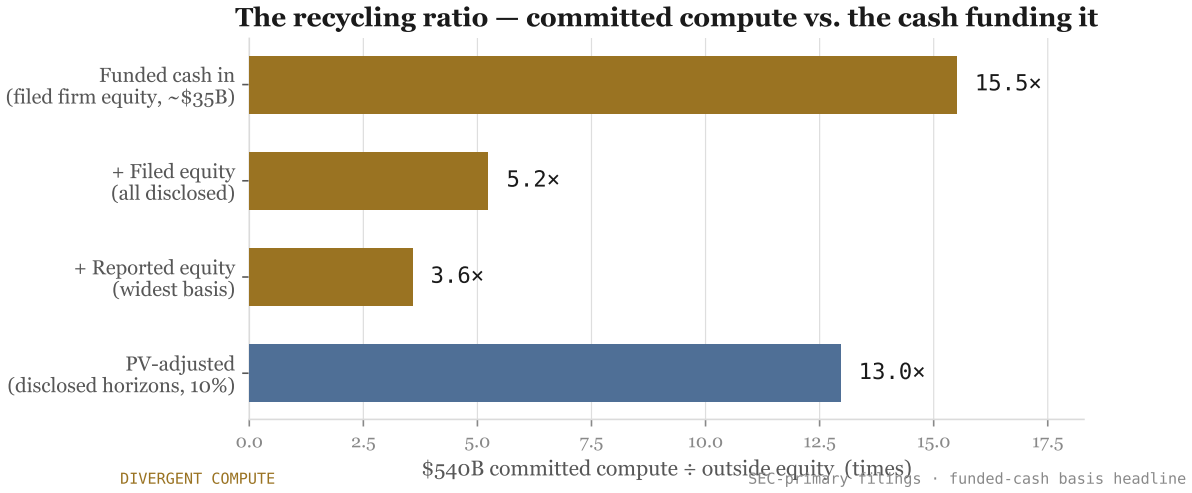

Figure 9: Recycling ratio across equity bases, with the present-valued read.

5.4 A taxonomy of round-trips

Not every edge is equally circular. Sorting them by *what is exchanged for what* separates arm's-length commerce from self-referential leverage.

- **Equity-for-compute** — the purest form: an investor takes a stake in a lab, and the lab commits that capital to buy the investor's compute. Nvidia's stakes in the labs, and the Microsoft and Amazon investments that sit beside their Azure and AWS commitments, are the load-bearing cases — the same dollar counted as investment on the way out and revenue on the way back. As of Q1 2026 the category's largest single instance is Amazon–OpenAI: $15 billion funded plus a $35 billion commitment letter, against $138 billion of OpenAI compute committed to AWS, in one filing.
- **Warrant-for-supply** — compensation for a supply commitment paid in the *customer's own equity.* The AMD–OpenAI warrant (up to 6 GW, struck against AMD stock) is the archetype: the supplier's upside is now tied to the buyer's share price, not to cash the buyer has earned.
- **Prepay / capacity-for-reservation** — multi-year commitments booked as backlog before the capacity is built or the demand proven: Oracle's Stargate RPO, the OpenAI Azure and AWS commitments. These are the largest edges by dollar and the least circular in *form*, but they carry the most stock-vs-flow risk — a commitment is not a payment.

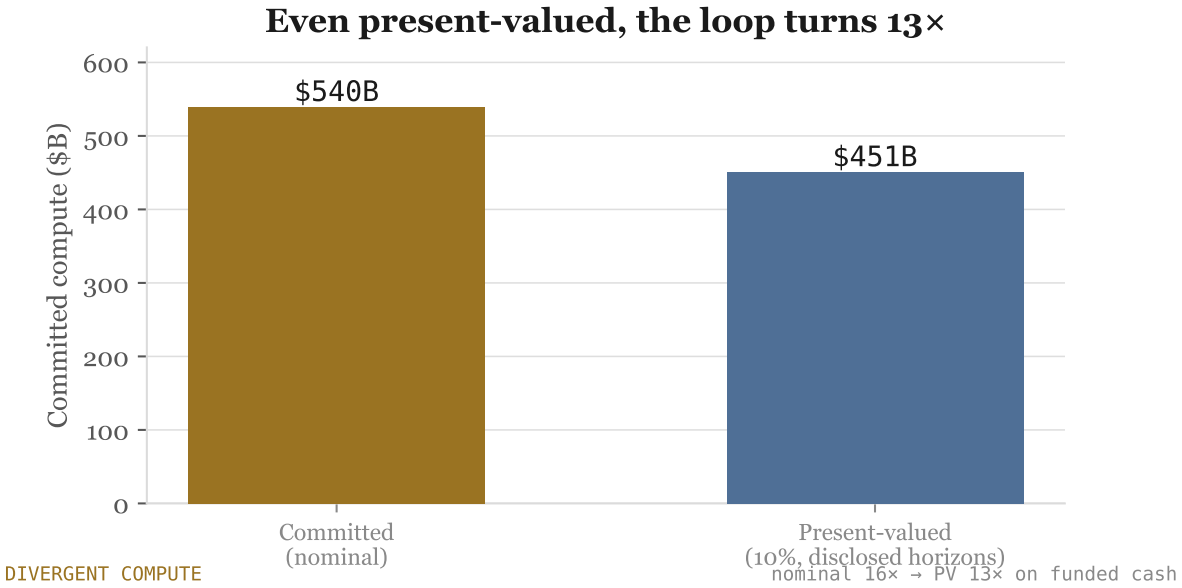

Figure 10: Even present-valued at 10% over disclosed horizons, the loop turns ~13× on funded cash.

The first two categories are where the loop's leverage concentrates; the third is where its *magnitude* lives. A reader who accepts only the third — the plain, filed compute commitments — still arrives at a recycling ratio far above any arm's-length benchmark.

5.5 Concentration: the ring is thin

The recycling is not diffuse. On the same $540B basis as the loop, 96% of committed lab compute routes back to just two firms — Microsoft and Amazon — the same two among the labs' largest equity backers. The share barely moves with the accounting basis: 98% on PRIMARY-filed edges alone, ~86% even when every soft leg is dollarized. A thin ring is a fragile one: a stall at either node propagates through the whole structure.

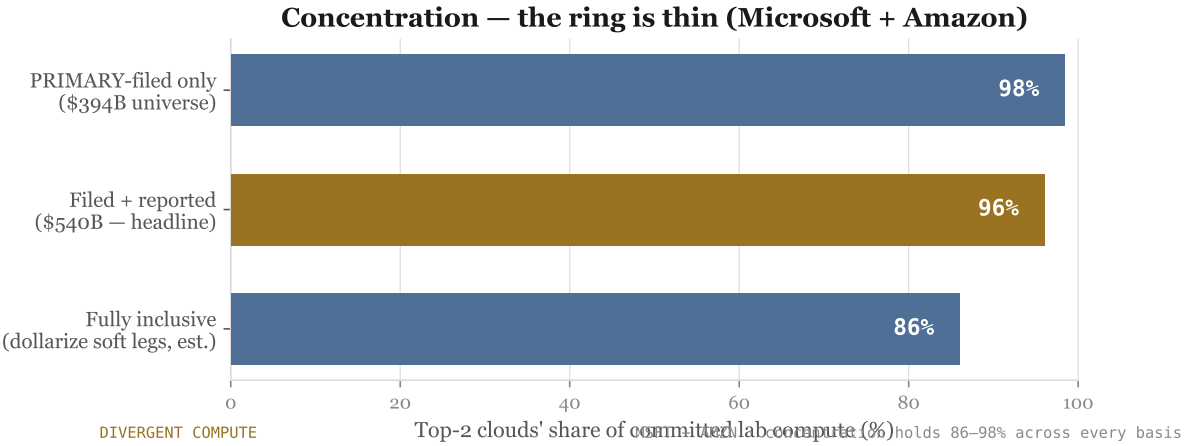

Figure 11: Top-2 cloud share of committed lab compute, across inclusion bases.

5.6 The pattern that broke before

The structure in Section 1 is not new. It is, almost line for line, the one that inflated and then destroyed the telecommunications-equipment industry a generation ago — the closest thing we have to a controlled experiment in vendor-financed demand.

The same loop. In the late 1990s the equipment makers — Lucent, Nortel, Motorola, Alcatel, Cisco — lent money to the carriers so the carriers could buy their equipment.[1] The loan left the vendor's balance sheet, returned as revenue on its income statement, and the receivable was booked as an asset. Lucent — the largest, with roughly $38 billion of revenue in 1999[2] — carried at its peak **more than $15 billion of vendor financing against about $300 million of operating cash flow**,[3] a recycling ratio of its own near fifty times: the same disease this paper measures, in the same units.

The same "demand." In the five years after the 1996 Telecommunications Act, carriers invested **more than $500 billion — mostly debt-financed — into fiber, switches, and wireless.**[4] Much of it was built ahead of a demand that never came: the fiber networks that cost billions "remained unused because there was no prospective demand for them, and the companies that built them went broke."[5] The industry had a name for the overcapacity it could not sell — **dark fiber.** Its direct descendant is a hyperscaler announcing, in 2026, that it will rent out its *excess AI compute.*

The same unwind. When the demand did not materialize, the loans that had been revenue became losses. Lucent's bad-debt rate went from 2.6% at the end of 2000 to **60% a year later**;[6] it absorbed roughly $3.5 billion of customer-loan losses across 2001–2002. Between 2000 and 2002, global telecom equities lost **more than $2 trillion** of market value.[7]

We are careful about what this does and does not establish. It does **not** prove the AI build-out will end the same way; history rhymes, it does not repeat on schedule, and the counter-argument — that AI demand is real and growing where fiber demand was speculative — deserves its hearing.[8] What it establishes is narrower and harder to dismiss: the *structure* — vendor-financed demand stacked on capacity built ahead of a demand promise — has broken before, violently, and those who named it early were treated as wrong for exactly as long as the financing held. The comparison is now being drawn in public.[9] Our contribution is not the analogy; it is the *ledger* — the full graph, quantified from the filings, so the ratio can be watched in real time rather than recognized only in the post-mortem.

5.7 What unwinds if one node freezes

A thin ring has a keystone. Remove one node — treat its commitments as suddenly unfundable — and measure how much committed capital is severed from the rest of the graph. The answer is stark: **OpenAI alone is counterparty to roughly $593 billion of committed capital** — $394 billion of committed compute (73% of the loop) plus the equity and markup legs priced against it. It is not one lab among several; it is the node through which the ring closes. Amazon (~$308B) and Microsoft (~$304B) follow — Amazon moved ahead of Microsoft when its Q1 2026 OpenAI equity legs entered the graph — then Anthropic (~$223B), each large enough that its withdrawal would reprice the structure without any single insolvency event.

This is what "the ring is thin" means operationally. The recycling ratio says the loop is over-levered; the contagion map says it is also *undiversified* — the leverage runs through one or two counterparties, so a

[1] Manufacturers "anxious to sell their products" financed the purchases themselves. Paul Starr, *The Great Telecom Implosion* (2002); *Telecoms crash*, Wikipedia.
[2] *Who Lost Lucent?*, American Affairs Journal (2020).
[3] *The rise and demise of Lucent Technologies*; *How the Once-Luminous Lucent Got Into Double Trouble*, TIME.
[4] Starr (2002).
[5] *Dark Fiber — an Archaeology of the Dot-Com Bubble.*
[6] *The rise and demise of Lucent Technologies.*
[7] *Telecoms crash*, Wikipedia.
[8] See e.g. *AI versus the Dotcom Bubble*, Janus Henderson (2026), for the bull case.
[9] Tomasz Tunguz, *Circular Financing: Does Nvidia's $110B Bet Echo the Telecom Bubble?* (2026).

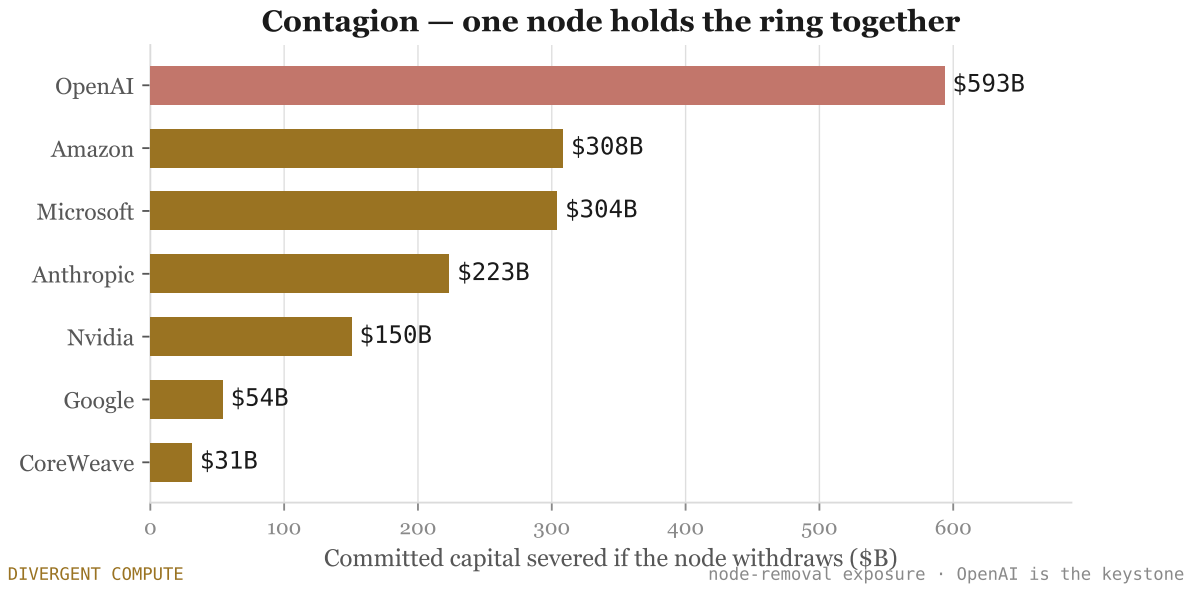

Figure 12: Committed capital severed if a single node withdraws — node-removal exposure.

stall at the keystone is not a local event but a system one. The structure offers no firebreak: because every large node is at once investor, supplier, and customer to the others, there is no arm's-length buyer standing outside the ring to absorb a shock. In 2001, for dark fiber, that outside buyer did not exist either.

5.8 What would prove us wrong

We state the exits in advance, so the thesis can be checked rather than argued. The recycling read weakens — and should be abandoned — if any of the following appears in the filings:

- **Third-party demand at scale.** Committed compute drawn down and paid out of *external* customer revenue — buyers outside the investor ring — rather than refinanced by the next equity round. The single cleanest refutation.
- **The ratio falls as outside equity grows.** If genuinely arm's-length capital enters the labs faster than new intra-ring commitments, the funded-cash ratio compresses toward an ordinary supplier-financing level. The July 2026 revision is the instructive near-miss: the ratio *fell* ($26\times \to 15.5\times$), but the new equity came from inside the ring — Amazon funding the lab committed to AWS — so the falsifier does not trigger. Denominator growth only counts when the capital is arm's-length.
- **Concentration diffuses.** If committed compute spreads beyond the top two clouds to a competitive set of buyers, the thin-ring fragility eases.
- **Commitments convert to cash on schedule.** If the multi-year backlogs are drawn and paid on their disclosed timelines without renegotiation or fresh vendor financing, the stock-vs-flow objection dissolves.

We publish the ratio every quarter regardless of direction. If it falls, we will say so as plainly as we report it rising.

5.9 What this means, and what it doesn't

Two disciplines are owed to the reader.

First, **these are disclosed facts and a ratio — not an accusation.** Every edge in the graph sits in a 10-K, 10-Q, or 8-K; nothing here alleges concealment or fraud. Firms are entitled to invest in their

customers and to sell them compute, and the commitments may well be honored in full.

Second, **the circularity is a fragility, not a crime.** The danger is not that any single deal is improper; it is *structural.* When the same two firms are simultaneously the largest investors in, and the largest suppliers to, the labs whose commitments underwrite their own capex, the failure modes correlate. A shortfall in end-demand does not strike one node — it reprices the whole ring at once, because every node is collateral for the others. That is what a thin, self-financed ring is: not a fraud to be prosecuted, but a load-bearing structure with a single point of failure — one that, on the evidence of Section 6, has been built before and has failed before.

6 The Divergence Engine: reading demand alongside supply

The preceding sections formalize the AI economy's *supply-side* fragility: six indicators (Section 4), each scored 0–100 on a filing-anchored rubric; the convergence rule (Section 3, Section 7); and the $D(t)$ divergence diagnostic (Section 8). This section does not modify any of that. It adds a *second family*—the demand-side reality (I7–I9)—and an *apex* that reads the two families' relationship. It is a Phase-2 extension in the paper's own sense: the same Phase 2 as the $G(t)$ β-calibration and the daily series (Section 8); the supply core (Phase 1) stands alone, and the demand family plus the apex bolt on as part of the paper's existing Phase-2 roadmap, not alongside a finished gauge. Carrying the paper's own caveat exactly: $D(t) = M(t) - G(t)$ is a Phase-1, equal-weight, *directional* diagnostic—"not a fit," and "not yet a tradeable instrument; a structured diagnostic confirming the two tapes have diverged in the direction and with the persistence the paper describes" (Section 8), with β-calibration and trailing standardization deferred to Phase 2. This section treats $D(t)/G(t)$ only at that Phase-1 directional standing.

The one asymmetry the reader must hold.

The two families are *not* evidentiary equals, and we say so plainly.

- **Supply (I1–I6): filings-hard.** Every figure is SEC-primary (10-K/10-Q/8-K/Form 4) or explicitly labeled—the paper's EDGAR-grade standard. Each cell can be audited against the filing.
- **Demand (I7–I9): survey/coverage-grade.** Real-economy adoption is *not* in the filings; it is read from enterprise surveys, vendor disclosures, controlled studies, and reported deployments—inherently softer evidence. Every demand-side figure below is sourced, dated, and labeled by grade (Section 6), and the demand family is deliberately built to *corroborate, not to carry*, the bet: the aggregation (Section 6) lets it confirm or diverge from the hard supply read, never override it.

This asymmetry is the section's central methodological honesty: the supply side is where the bet's weight rests; the demand side tells you *whether the spending is buying real value*—measured to the best available standard, labeled as such. We do not dress survey data as filings data.

6.1 The thesis: a bubble is a divergence between a narrative and the ground truth

The paper's epistemic move is Burry's own: check the ground truth the consensus narrative ignores. In 2008 that was the loan tapes; here it is the filings (the supply side) *and*—this section's addition—the real-economy adoption tape (the demand side). The Phase-1 $D(t)$ diagnostic (Section 8) already reads one divergence directionally: the *market* tape (SOXX/valuation momentum) against the *financial* ground-truth tape (layoffs, insider, capex-gap; the depreciation-flatter carried as a static annotation). This section adds the divergence the filings cannot see directly: the *AI narrative*—capex, valuations, the promise that "AI will transform everything"—against the real-economy value it is supposed to be producing.

The completed instrument—the **Divergence Engine**—reads two families:

- **Supply (I1–I6, existing):** financial fragility. *Native polarity: high = more fragile.* (Unchanged.)

- **Demand (I7–I9, new):** adoption-vs-value reality. *Reality-meter polarity: high = reality validates the narrative (durable); low = the narrative outruns reality (divergent).*

The bubble signal is divergence between them: the spending is fragile *and* the value it promises is not materializing. The section's central honest finding (Section 6): the families can also *disagree*, and that disagreement—*fragile-but-funded*, or *fragile-but-real-demand*—is the most important diagnostic in the AI case today.

6.2 The demand side (I7–I9)

These are reality-meters (high = durable; low = divergent), which makes symmetric grading *structural*: a reality meter cannot only ever turn bearish.

I7 — Real-Economy Adoption (adoption-narrative vs value-reality).

I7 is a fractal of $D(t)$ at the indicator level: the *gap* between an AI deployment's adoption narrative and its value reality. *Adoption is a prerequisite, not a score.* The bands (locked):

I7 band	Meaning
0–20	Announced only, or *deployed-then-retrenched* (walk-backs)
21–40	Adopted but value unproven — the "adoption ≠ value" zone
41–60	Early, measurable value in real deployments
61–80	Durable value across multiple players
81–100	Load-bearing / transformative

No reading above 40 without both real adoption and real value. Worked examples (sources in Section 6):

- **Klarna (customer service) ≈ 15.** Deployed at scale—an AI assistant handling two-thirds of chats, "the work of 700 agents" (Klarna, Feb 2024)—then *walked back*, re-hiring humans and conceding "lower quality" (2025). Deployed-then-retrenched is the strongest low signal. *(High certainty—a documented, dated reversal.)*
- **Microsoft 365 Copilot ≈ 25—the "adoption ≠ value" exemplar.** High licensed reach, low real use: deployed across a large share of enterprise seats, yet a substantial share of licenses go unused and ROI is unproven. The instrument reads this trap by construction—deployment headlines do not earn a value score. *(High certainty on the trap; no clean primary exists for a precise active-use figure, so we cite none.)*
- **AI coding (Cursor, Claude Code, Copilot) ≈ 50–55—a deliberately two-axis reading.** Adoption is genuinely high ($ $90% of developers use an AI tool; multi-billion-dollar run-rate revenue), but *value is contested*: a controlled randomized trial found experienced developers *expected* to be $ $24% faster and *self-estimated* a $ $20% speed-up, yet were measured 19% *slower* with the AI tools.[10] The honest reading is "adoption ≈ 80 / value contested," *not* a clean high score—the single most important demand-side correction in the arc: even AI's clearest commercial success does not earn a clean durability score, because I7 measures durable *value*, not adoption or revenue. *(Adoption high-certainty; value moderate.)*

I7 vs the supply-side I6 (and I2)—shared evidence, different lenses, no double-weight.

The supply-side I6 (organic end-user demand, Section 4) and I7 both draw on the MIT-NANDA $ $95% finding—which the paper already uses to anchor I6's demand read. That overlap is deliberate and disclosed, and it is *not* double-counting, because the two families read the same evidence through different lenses for

[10] METR, *Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity*, arXiv:2507.09089 (10 July 2025); randomized controlled trial, 16 developers, 246 tasks.

different purposes. **I6 is a filings-side fragility input**: it asks whether a firm's *reported* AI revenue is backed by paid pilot-to-production conversion (higher = more fragile), feeding the supply count. **I7 is the demand-family reality-meter**: it asks whether AI is *delivering durable value in deployment* (low = the narrative outruns reality), feeding the demand trigger. The aggregation (Section 6) is what makes the shared evidence safe: *the supply count and the demand trigger are never summed*—NANDA can inform a firm's I6 score *and* the economy-wide I7 read without the same fact moving two addends of one total. Family separation is precisely the mechanism that prevents double-weighting. (This parallels the I9-vs-I3 separation below.)

I8 — Narrative-Reality Sentiment (certainty-capped ≤ 40).

The gap between the official AI narrative and insider/community sentiment (scrubbed). Certainty-capped at the low end (forum selection bias; sentiment $\neq$ fundamentals). **I8 flags; it never concludes.** Zero independent score weight (Section 6).

I9 — Labor-Reality (held).

Whether AI-attributed labor changes are real displacement or "AI-washing." Attribution is hard and cuts both ways. **Held—advance-last; collect texture, do not conclude.** Zero independent score weight (Section 6). Distinct from the supply-side I3 (Section 4), which scores insider-selling from Form 4s—I9 is a demand-side labor read, not the filings insider tape.

6.3 The ROI–productivity gap, with its sourced citation layer

Underlying I7 is the macro fact that frames the demand side: *individual-level AI productivity is real and measured, while enterprise-level ROI mostly is not—both true at once.* Every figure in the table below is sourced, dated, and graded—and graded *softer* than the supply side's EDGAR-primary, by construction.[11]

Table 6: The demand-side ROI–productivity gap — sourced, dated, and graded (PA-verified citation lock, 2026-06-23). Demand figures are survey/coverage-grade by construction, never EDGAR-primary; the † marks interested parties (see footnote).

Figure	Value	Source + date (exact stat)	Grade
the table below continued |			
Figure *continued*	**Value**	**Source + date**	**Grade**
Enterprise pilots, no measurable ROI	$ $95% | MIT Project NANDA (Challapally et al.), *The GenAI Divide: State of AI in Business 2025*, July 2025; "95% of organizations are getting zero return"; 300+ deployments / 52 interviews / 153 surveys | Study (not filings) |		
named critic (required)	—	Brynjolfsson, Rock & Syverson, *The Productivity J-Curve* (NBER w25148) and *Generative AI at Work* (w31161): the 95% may be adoption-lag, not permanent failure | Named economist |	

[11] **Interested-party note.** Several demand-side figures in the table below—Writer (#2), McKinsey (#7), and PwC (#8)—come from vendors or consultancies that sell AI products or AI-advisory services. Each is quoted accurately and to the source's own wording, but as an *interested party*—noted so the demand-side evidence cannot be dismissed as marketing, and so it cannot be over-leaned-on either. The filings-hard supply side (I1–I6) carries no such dependency; it is the demand side's corroboration, by design (Section 6).

Figure	Value	Source + date (exact stat)	Grade
Firms with *significant* gen-AI ROI	$ 29n=2{,}400), 7 *Apr* 2026; "*only*29^{†}$ |		
Developers using ≥ 1 AI tool	$ $90% | JetBrains, *Developer Ecosystem* research, Apr 2026 (as of Jan 2026) | Survey |		
Coding-tool revenue	Claude Code >$2.5 B run-rate; Cursor ~$2 B ARR	Anthropic/Amodei (Feb 2026); Cursor via Bloomberg (2 Mar 2026)	**Vendor-reported**
Coding productivity paradox	forecast $+24+$20% faster → actually 19% *slower* | METR, arXiv:2507.09089, 10 Jul 2025; RCT, 16 developers, 246 tasks; "AI increased completion time by 19%" | Controlled study (RCT) |		
Work-hours saved (macro)	$ $1.4% of total work hours | Bick, Blandin & Deming, *The Rapid Adoption of Generative AI*, St. Louis Fed WP 2024-027F / NBER w32966, rev. 27 Oct 2025; "time savings equivalent to 1.4% of total work hours" | Central-bank WP |		
AI adoption vs scaling	adopted in ≥ 1 function, but only $ 7n=1{,}993$): 88% use AI / 79% use gen AI in ≥ 1 function, only 7% fully scaled | Survey† |		
AI-value concentration	20% of firms capture $ 74n=1{,}217), 13 *Apr* 2026(*alsocitedat@sec − indicators*)|*Study*^{†}$ |		

Both sides are true. The productivity gains are real (coding, customer-service augmentation, the Fed's work-hours estimate); the enterprise-value gap is real (the $ 29 $95% figures, McKinsey's $ $7% fully scaled, PwC's 20%/74% concentration). Whether the gap is a *J-curve lag* (resolves up → durability) or *structural* (resolves down → bubble) is the open question, read both ways (Section 6). Citing the scary statistic together with its critique is the honesty.

6.4 The mechanism: outcome-verifiability

A count of walk-backs invites the objection "early-adoption noise." The Engine's spine is a *mechanism* explaining *why* demand-side value clusters where it does (with two falsified hypotheses on record, Section 6).

> *The defensible primary mechanism:* AI durably automates a task when its **outcome is cheaply and objectively verifiable**; it is held to augment-only (a human stays) when the outcome is not.

The arc, in brief: the task-type boundary exists (after the "augment-vs-replace" one-liner was *falsified*, Section 6); *verifiability is the driver, not regulation or stakes* (the within-finance control: same regulation, but fraud's verifiable outcome automates while fiduciary advice augments); it survived falsification and refined to *outcome*-verifiability, with un-regulated high-stakes domains automating their verifiable tasks (so regulation is an *enforcement overlay*, not the mechanism) and catastrophic-irreversibility a *secondary accountability backstop* (a human stays at the extreme—aviation, autonomous vehicles—but as a backstop, not the task-doer); and it held with the confound removed in low-stakes hospitality. Acemoglu's task-based analysis—already cited supply-side at Section 4—makes the same point from the production side: durable automation needs "objective outcome measures from which to learn successful performance."

Why it matters for the bet.

The bulk of the AI *narrative*—transforming knowledge work, judgment, relationships—lives precisely in the **low-outcome-verifiability** region, exactly where durable value has *not* materialized. The mechanism gives the demand-side fragility a *reason*, not just a tally. It also vindicates the paper's own closing: the closing note argues, as a values claim, that AI should *augment* human judgment and never replace it; outcome-verifiability is the positive-economics counterpart of that normative claim—AI is in fact held to augment-only precisely where outcomes resist cheap verification, which is most of the judgment work the closing note is about. The "should" and the "is" point the same way. (A resonance, noted lightly—not load-bearing on the bet.)

6.5 Aggregation: nested convergence, and the insulated-bubble apex

The section *nests into the paper's existing convergence rule; it introduces no competing count.*

- **Supply-HIGH = the paper's existing convergence, unchanged:** a firm/layer is supply-active when ≥ 3 of I1–I6 are ≥ 60 (independent, simultaneous; *necessary, not sufficient*, Section 3, Section 7). We adopt this verbatim as the supply-family trigger. **I7 is not added to this count**—it belongs to the other family.
- **Demand-LOW = the demand trigger:** I7 ≤ 40 (the "adoption $\neq$ value" zone or below). **I8 and I9 carry zero independent score weight**—they act only as *certainty modulators* (corroborate $\rightarrow$ narrow the certainty band; contradict $\rightarrow$ widen it). The demand read comes from I7.
- **The two-family binary convergence (the apex, above the existing framework):** high-confidence = Supply-HIGH *and* Demand-LOW, simultaneously—both families pointing the same way, *necessary not sufficient* for a correlated drawdown (matching the paper's standard). No averaging, no weighted blend.
- **Disagreement $\rightarrow$ the insulated-bubble diagnostic (named, not averaged):**
  - *Fragile-but-funded*—Supply-HIGH but the fragility insulated by fortress balance sheets / profitable incumbents funding the build from cash.
  - *Fragile-but-real-demand*—Supply-HIGH but insulated by genuine pockets of durable value (fraud detection, coding adoption, ambient clinical documentation).

Why the apex sits above $D(t)$, and is not folded into $G(t)$.

Extending the ground-truth composite $G(t)$ (Section 8) to include the demand family would average the two readings into a single number—destroying the very disagreement the apex exists to name. The Phase-1 $D(t)$ diagnostic and the supply convergence stay intact; the apex reads the *relationship* between the families. The insulated bubble is a *structural state*, not a blended score. This structural argument is independent of $G(t)$'s weighting: it holds whether $G(t)$ is the current Phase-1 equal-weight directional construct or the Phase-2 β-calibrated version—folding demand in would average away the disagreement at either standing. The apex is therefore a Phase-2 addition that sits *beside*—not inside—the Phase-2 $G(t)$ calibration.

The engine's honest present reading: an *insulated bubble*.

Real, accumulating supply-side fragility (the paper's domain: circular financing, concentration, capex-vs-demand)—currently *insulated* by cash-rich incumbents and narrow-but-real demand. This is sharper and more defensible than "the bubble is about to pop," and consistent with the paper's own position that this is *not* 2008 in solvency (Section 9). The danger is the *insulation eroding*—which the watches (Section 6) track. Readings are ordinal, not decimal.

6.6 The three-watch falsifiability spine

The engine is *cross-sectional*—a present reading. Its forward honesty is three standing watches, each with an explicit falsifier, re-measured each cycle: how the insulated bubble would resolve, and in which direction.

1. **J-curve watch (demand).** Does the ROI–productivity gap *narrow* (durability) or *persist* (structural)? *Current: ambiguous*—productivity up but K-shaped, the enterprise-ROI rate roughly flat, GDP figures revised. (The named anchor on the J-curve side is Brynjolfsson, NBER w25148, Section 6.)
2. **Cash→debt watch (supply).** Does AI capex shift cash-funded → debt (insulation eroding)? *Current: mostly cash, cracking at the margin*—large new debt raises; the "AI bond binge" (cf. Section 5).
3. **Verifiability-prediction watch (mechanism).** Does the boundary *predict over time*—new durable automations clustering in cheaply-verifiable tasks, new walk-backs in unverifiable ones? A standing 2×2 contingency; the off-diagonal growing is the falsification signal. *Honest caveat:* public coverage over-reports dramatic walk-backs—read trends, not levels.

The insulated bubble *resolves bearish* if the insulation erodes on multiple watches at once; *resolves bullish* if the J-curve harvests and durable value broadens. The watches are the engine's standing offer to be proven wrong.

6.7 Two credibility exhibits: the instrument policing its own elegance

We put on the record two elegant hypotheses we generated and then *broke*—keeping only the harder versions that survived. An analyst who never shows a discarded hypothesis either did not look hard enough or is not telling you what they found. (This is the same epistemic the paper already practices in naming its own falsifiers, Section 9.)

Exhibit 1 — the "augment-vs-replace" one-liner (hypothesized → broke → refined).

The clean rule "AI augments durably; it walks back when it tries to replace" broke on *replace-but-durable* (medical transcription $ $99% automated; data entry; warehouse) and *augment-but-failed* (Copilot's low real use). The surviving refinement is the verifiability mechanism (Section 6); augment-vs-replace was a symptom, not the cause.

Exhibit 2 — the "expected-error-cost" unified frame (hypothesized → stress-tested → broke).

An elegant single expression—*automate when error-rate* × *error-cost is low*—broke on four counts: (a) the *human-made premium* (a real, growing market in which people pay *more* for verifiably human origin even when the AI output is technically equal or cheaper—a low-stakes task that does *not* automate, for reasons outside any error calculus; this corrected an earlier hope it would "collapse into verifiability"—it does not); (b) *time-criticality* (high-stakes tasks fully automated because a human cannot act at machine speed—sub-millisecond trading—which the frame mis-predicts as "augment"); (c) "error-cost" is a *composite* (magnitude × irreversibility × detectability × liability), not clean "stakes"; (d) *non-operationalizable*—you cannot estimate the error-*rate* of the unverifiable tasks it is most about. **Discarded as a model.** The locked mechanism stood *because* it refused to over-unify. ("Elegant" was the warning, not the endorsement.)

6.8 The honesty (load-bearing), and the line that must survive

For a real-money short the credibility layer *is* the instrument:

- **Symmetric by construction** (reality-meter polarity; the bull case collected as rigorously as the bear—coding's real adoption, fraud's durable ROI, incumbent profitability, valuations milder than the dot-com peak, the live J-curve).
- **Self-falsification on the record** (Section 6)—the two broken exhibits.
- **Certainty-layered**—every reading graded; the bet rests on the high-certainty layer; the supply side filings-hard, the demand side survey-graded and labeled; I8/I9 capped/held.
- **A bounded observation, not over-claimed**—the human-made premium is real and growing but concerns *where human work persists*, not *whether AI capex pays off*; it is not loaded onto the bet.
- **Timing disclaimed** (the paper's own standard, Section 9): it detects fragility/divergence, not *when*.

And so, the governing line, once more: the Divergence Engine **strengthens the evidence; it does not settle the bet.** It gives the AI-bubble thesis a measurable, two-sided, falsifiable, mechanism-backed spine, and an honest present reading—an *insulated bubble*, fragility accumulating behind eroding insulation. What it cannot do is tell you the date. That remains the author's judgment, and the author's risk.

7 Convergence and the inverted pyramid

The indicators of Section 4 are individually suggestive and jointly decisive. Applying the convergence rule (Section 3) across all 68 firms yields **15 *active* names**, and the rule's internal consistency is total: *every one of the 15 reaches the flag by at least three independent elevated indicators*—there is no firm flagged active on a single extreme score. Equally important are the firms the rule *declines* to flag despite three or more elevated indicators—Alphabet, Intel, SoundHound, Tesla, and Caterpillar—each capped at moderate by a filing-grounded verdict that the underlying mechanism is bounded rather than systemic. That the framework actively withholds the active label from high-scoring names is the strongest evidence it is measuring structure and not merely summing alarm. Crucially, this cuts *against* the thesis: every cap removes a name from the active count, raising the bar the convergence rule must clear rather than lowering it—the opposite of confirmation bias. The ceilings are rule-based rather than discretionary, each triggered by a filing showing the underlying mechanism is bounded—Alphabet's capex, for instance, is met by real advertising cash flow, so its high raw scores do not imply the demand-financing fragility the rule is built to detect.

Are the weights doing the work?

A fair objection to any composite is that its weights—here $w = (0.20, 0.20, 0.15, 0.20, 0.10, 0.15)$ in Eq. the referenced exhibit—quietly encode the conclusion. They do not, and this is structural rather than asserted: the *active* classification is the convergence rule, a *count* of independent indicators at or above the elevated threshold, and a count does not reference the weights. All 15 active names clear the rule by at least three elevated indicators under *any* weighting, because the weighting is simply not in the rule (the figure below, left). The weighted composite F_i exists only to order firms within a tier for presentation. Even that ordering is robust: equal weights reproduce 14 of the 15 top names, and across 5000 weightings drawn uniformly from the simplex the top-16 overlaps the baseline by 81% on average (the figure below, right). F_i is sensitive only at adversarial extremes—which is immaterial, because the classification was never F_i's to make.

the figure below shows the firm-by-indicator scores for the core-deep set (Layers 1–3 plus the flagged-fragile application names). Read by column, two indicators run hot down nearly the entire compute–hyperscaler–lab base: *capex-versus-demand* (I2) and *circular financing* (I4)— the two structural mechanisms. Read by row, the active names (red labels) visibly carry multiple hot cells across *independent* columns; that horizontal co-occurrence is convergence made visible. Insider selling (I3) is correctly cooler and patchier—the discounted, noisier signal—and xAI's two N/A cells (no public depreciation or Form 4 filings) are shown as such rather than imputed.

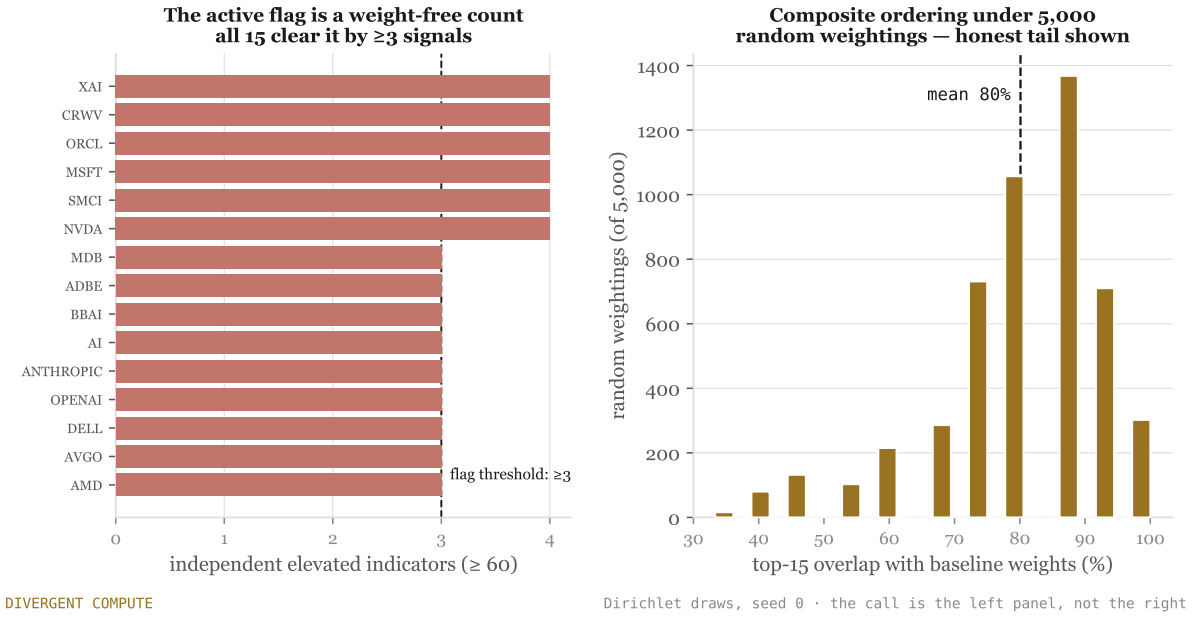

Figure 13: Weight-sensitivity of the framework. Left: the active classification is a count of independent elevated indicators (60); all 15 active names clear the 3 threshold, and a count cannot be tuned by the composite weights. Right: the weighted composite Fi, used only to order firms within a tier, retains 80% of its top-15 on average under 5000 random reweightings (93% under equal weights), with a left tail only at adversarial extremes. The conclusion rests on the left panel, not the right.

The inverted pyramid.

Aggregating to the layer level reveals the load-bearing geometry of the cycle (the figure below). The *active share*—the fraction of a layer's firms flagged active—rises from 33% at the compute layer (L1) to 50% at the hyperscalers (L2) and peaks at 62% at the model labs (L3), then collapses to 17% at the application layer (L4) and 0% across the broad market (L5), where 19 of 25 names are not even materially exposed. Fragility is not diffused through the economy; it is concentrated in a narrow tier of compute, cloud, and model-lab firms, with the broad market resting on top of it. The geometry is an inverted pyramid: the widest, calmest layer (the broad market) is borne by the narrowest, hottest one (compute and the labs). This matters for contagion analysis (Section 9): a shock does not have to find its way *into* a narrow base from a broad market—it originates there, at the point on which everything above is balanced.

8 The divergence gauge

The two tapes of Section 1 can be made into a single number. Let $M(t)$ be the *market signal*—a standardized read of the AI-exposed equity tape, for which the semiconductor index (SOXX) is a clean proxy, constructed in full below—and let $G(t)$ be the *ground-truth signal*, a normalized composite of the mechanical deteriorations this paper measures:

$$D(t) \;=\; M(t) \;-\; G(t), \qquad G(t) \;=\; \sum_{j} \beta_j \, g_j(t), \tag{3}$$

where the $g_j(t)$ are standardized ground-truth series—AI-attributed layoffs, discretionary insider intensity, the depreciation-flatter run-rate, and the capex-versus-demand gap—and the β_j are their weights. The object of interest is not the level of D on any given day but its *persistence*: a one-period gap is noise; a gap that widens and holds for several quarters is fragility accumulating beneath a rising price.

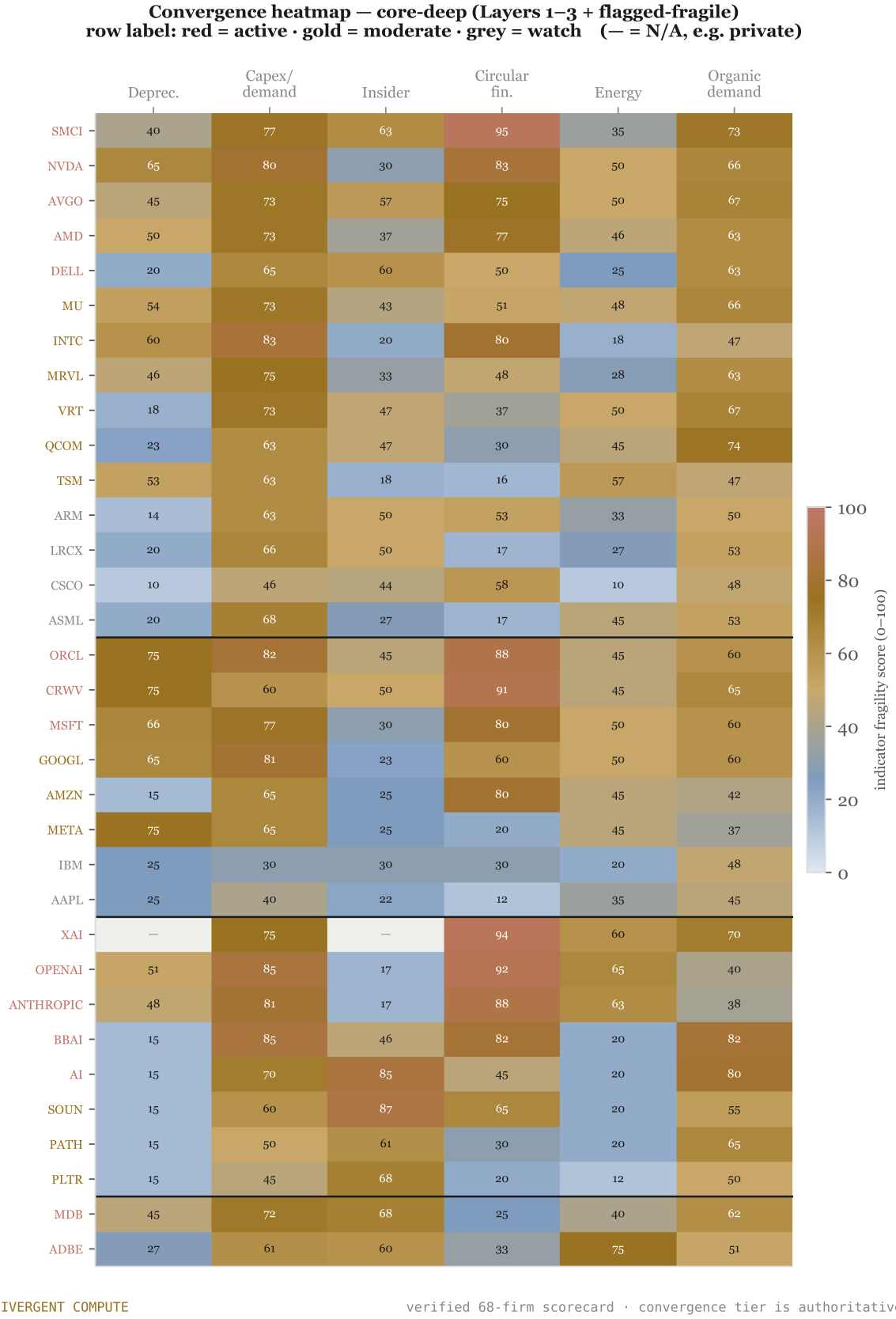

Figure 14: Convergence heatmap, core-deep. Cell color is the indicator fragility score (0–100); row labels are colored by verified tier (red active, amber moderate, grey watch); "—" is N/A. The hot columns are capex-vs-demand and circular financing; active rows show 3 hot cells across independent indicators.

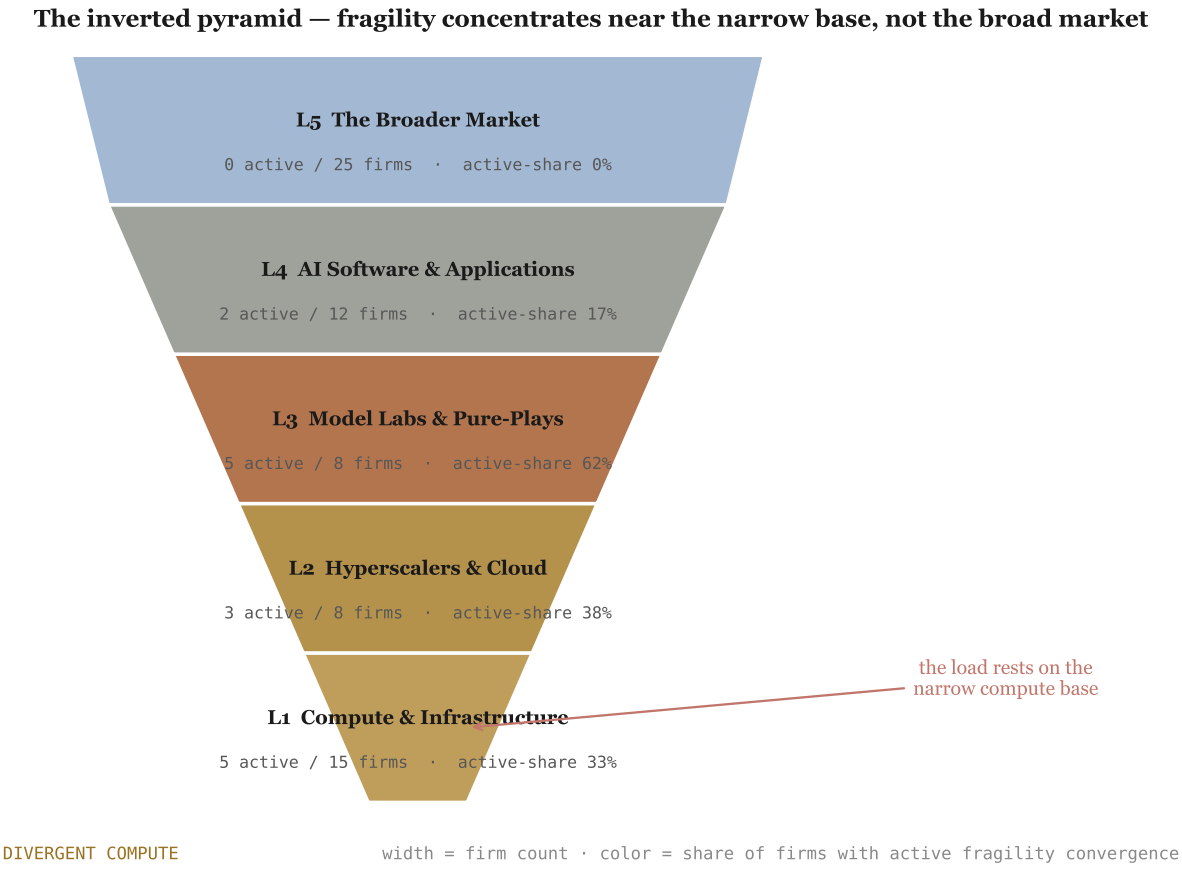

Figure 15: The inverted pyramid. Width is layer breadth (firm count); color is active share. Fragility concentrates near the narrow compute base and the labs just above it, and thins to zero across the broad market.

Constructing $M(t)$.

Momentum alone understates the signal: a melt-up that is merely *high* is less telling than one that is high, stretched, and unstable at once. We therefore construct $M(t)$ as an equal-weighted, standardized blend of three price-only components on SOXX—*momentum* (trailing 63-day return), *overextension* (price relative to its 50-day trend), and *instability* (20-day realized volatility)—each entered as a z-score over the sample. M requires no fundamentals; it is built entirely from the price tape, which keeps it cleanly separable from G. As of the June 2026 close the blend reads $M = +2.83$—its most extreme value of the trailing year—with all three components elevated together: trailing return $ $88%, price $ $26% above its 50-day average, and realized volatility $ $74%—roughly triple its 2025 median of $ $25% (the figure below). An honesty note on M: its standardization is full-sample and therefore *descriptive*—it places today's reading in the year's context, whereas a real-time tradeable M needs trailing standardization, a Phase 2 calibration. $G(t)$ is now constructed too, as a Phase-1 equal-weight composite of the three time-varying ground-truth series—AI-attributed layoff share, discretionary insider intensity, and the capex-versus-demand gap—standardized on the same basis as M and signed so that a deterioration *lowers* ground-truth strength (G falls as the foundations erode); the depreciation-flatter run-rate is held as a static $3.68 billion/yr annotation, entering as a series in Phase 2 as filings accumulate. Under this specification G falls from $+0.98$ in 2025 Q3 to -1.23 in 2026 Q2 while M climbs, so $D(t) = M(t) - G(t)$ widens monotonically from -1.80 to $+4.06$ across the window. The equal weights are a documented Phase-1 choice, not a fit, and carry the same descriptive (full-sample) caveat as M; Phase 2 will fit β_j on a held-out window, adopt trailing standardization on both sides, and fold in the depreciation series. $D(t)$ is not yet a tradeable instrument—it is a structured diagnostic confirming the two tapes have diverged in the direction and with the persistence the paper describes.

The first-half-2026 reading is a textbook positive divergence. M compounded violently—SOXX $ $91%—while every component of G moved the wrong way: AI-attributed layoffs accelerated (the AI-blamed share of tech cuts climbing from $ $7% to $ $40% over the window), discretionary insider selling concentrated in the most exposed names (Section 4), the depreciation-flatter run-rate added ∼$10.5 billion/yr of non-cash-backed earnings (Section 4), and the capex-versus-demand gap widened toward $354 billion of annual spend against two-to-four-times-slower revenue (Section 4). The price tape and the ground-truth tape have pulled apart—price climbing as ground truth falls—for two consecutive quarters, the signature of a structure whose surface is appreciating while its foundations erode (the figure below). It is the series the framework tracks forward—a Phase-1 diagnostic here, calibrated toward a live instrument in Phase 2.

Froth at the top of the tape.

The melt-up is drawing the late-cycle structured leverage that tends to cluster near extremes. In the ten trading days of 8–18 June 2026 alone, at least fifteen SOXX-linked structured notes were filed with the SEC by eight separate issuers—JPMorgan, Barclays, Toronto-Dominion, Morgan Stanley, HSBC USA, RBC, BofA, and Jefferies—autocallable and buffered products that package leveraged semiconductor exposure for yield-reaching retail buyers. A representative JPMorgan autocall (CUSIP 46661AUF5) references a *worst-of* basket of SOXX and the Nasdaq-100, runs to June 2028 behind a 20% buffer and a 39.6% return cap, and carries an issuer-estimated value of only $957 per $1,000 note at pricing—a $ $4.3% structuring premium paid for leveraged exposure at a one-year volatility extreme.[12] A note factory printing worst-of semiconductor structures by the dozen is itself a reading on where sentiment sits.

9 Forecasts and falsifiers

A fragility map is not a timing call (Section 1), so the forecasts here are *conditional* structures, not dates. We give three scenarios with qualitative probability bands—deliberately qualitative, because false precision

[12] JPMorgan 424B2, SEC accession 0001213900-26-070042 (filed 2026-06-18); the 8–18 June tally is from an EDGAR full-text search of 424B2/FWP filings referencing SOXX over that window.

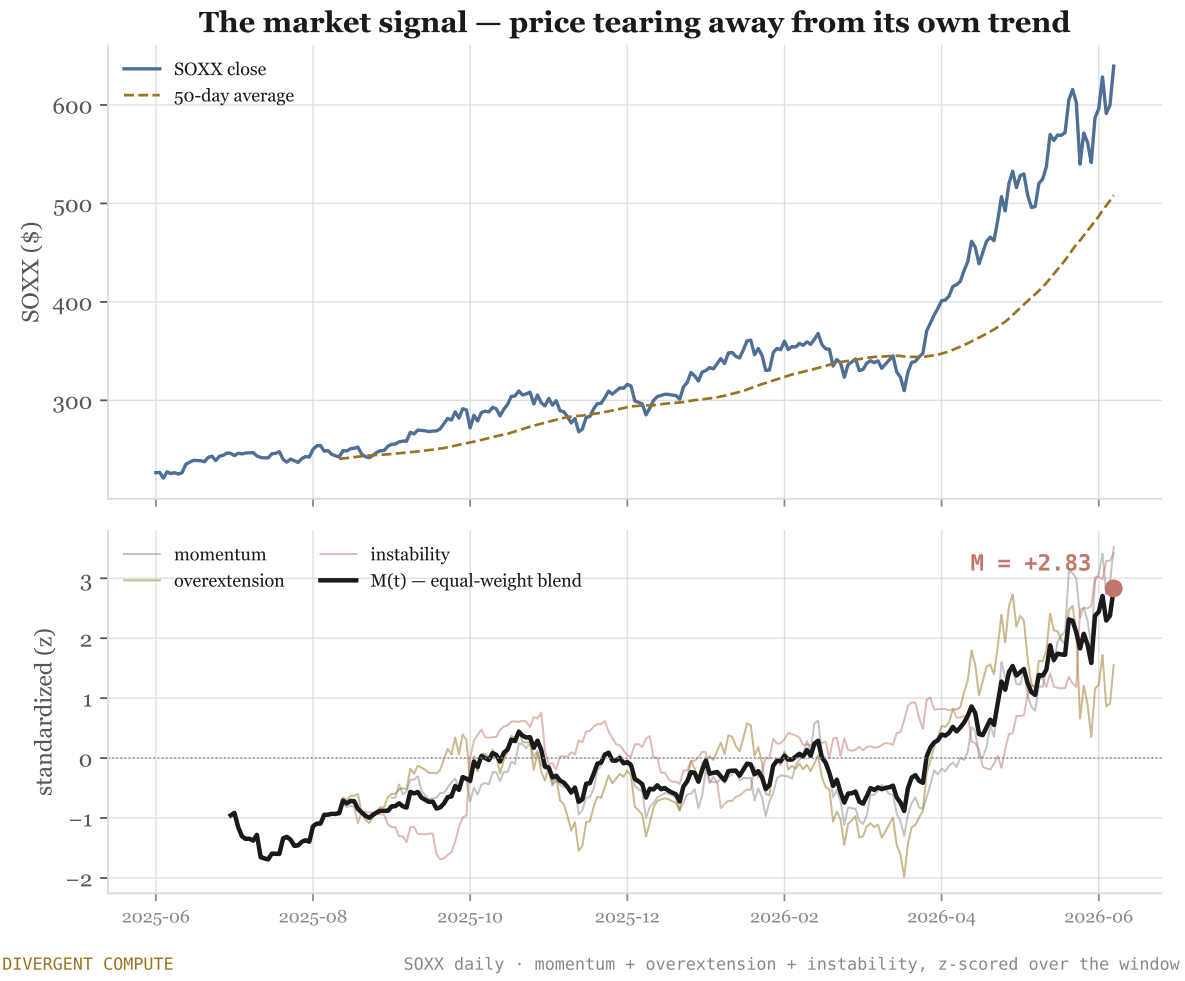

Figure 16: The market signal M(t). Top: SOXX pulling away from its own 50-day trend. Bottom: M(t) as the standardized triple blend—momentum, overextension, and instability—with its three components; the current reading sits at a one-year extreme. Built from price alone, so it stays cleanly separable from the ground-truth side G(t).

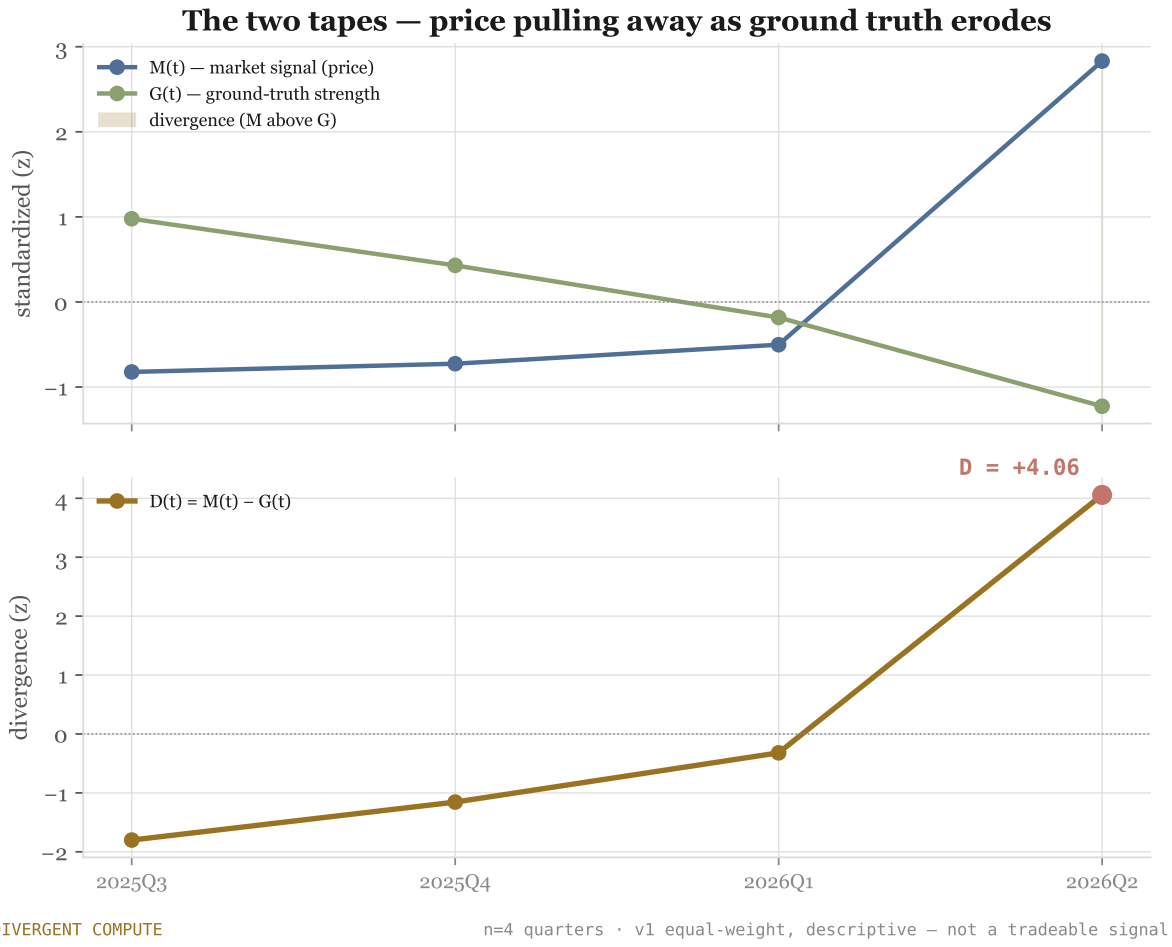

Figure 17: The divergence gauge D(t) = M(t) − G(t). Top: the two tapes—M(t) (market signal) climbing as G(t) (ground-truth strength) falls. Bottom: D(t) widening from −1.80 to +4.06 across H1 2026. Phase-1 equal-weight and descriptive; calibration is a Phase-2 step.

on a tail is itself a kind of dishonesty—and then we do the thing that matters most for a real position: we name the conditions that would make us wrong.

The solvency objection — what forces an unwind if everyone is solvent?

This is the single hardest challenge to the thesis, and it deserves a direct answer, not a caveat. The 2008 borrowers were structurally broke; the 2026 hyperscalers are the opposite—fortress balance sheets, vast cash flows, no contractual reset fixed to a calendar. If no one is forced to sell, what forces an unwind? The answer is that solvency prevents *default*, not the *contraction of recycled revenue*—and the demand side of the loop is radically concentrated. Modeling node-removal on the financing graph (Section 5) makes it concrete: a single node, OpenAI, anchors ∼$394 billion of compute commitments[13]—roughly 73% of the ∼$540 billion the three leading labs have signed (Section 5), and the single largest demand dependency in the graph (the figure below). So the unwind needs no insolvency. It needs one keystone to step back: if a single lab's next financing round fails to close, the ∼$394 billion of compute commitments it anchors lose their counterparty, and the capex—plus the ∼$18.2 billion of mark-to-model gains the providers booked *against that demand* (Section 5)—lose their basis. Microsoft and Amazon stay perfectly solvent throughout, and still watch the revenue that justified the build evaporate. There are two keystone *types*: a **demand** keystone (a lab whose financing fails) and a **substrate** keystone (Nvidia—simultaneously the physical GPU supply and the collateral behind the neocloud's debt; its withdrawal is the supply-side form of the same break, and its posture toward the loop—continued vendor financing, backstop terms, disclosed neocloud exposure—is the operational signal to watch). Concretely, a step-back is observable well before any bankruptcy: a private bridge or mezzanine round that fails to price or closes down-round; a neocloud covenant tripped or its GPU collateral revalued; a hyperscaler capex *guide-down*; a grid or permitting constraint that caps a data-center campaign's power; or a State decision that conditions or redirects sovereign-compute support. Any one of these snaps a link in the loop without a single balance sheet going insolvent. Solvency does not avert the unwind; it only selects its *form*: a fast repricing if a keystone visibly steps back, or the slow bleed of the propped scenario below.

The contagion mechanics behind the cascade scenario — node-removal exposure across the financing ring, with OpenAI as the keystone — are developed in Section 5.

Is this just a J-curve?

A fair challenge: every large infrastructure build runs ahead of its revenue, and the late-1990s telecom fiber glut is the standard analogy—vast capex, a demand that lagged, then a demand that eventually *arrived* and vindicated the build. Could the AI cycle simply be early rather than fragile? It could—and we say so in the falsifiers. But three features separate it from a clean J-curve. First, fiber demand was *latent and exogenous*—consumer broadband arrived independently of the firms that laid the fiber—whereas here the demand is, in large part, *recycled among the same balance sheets that funded it* (Section 5), the opposite of an independent customer showing up. Second, the earnings are being *flattered while we wait*—useful lives extended to postpone depreciation (Section 4) rather than written down—so reported economics improve as the underlying ones do not. Third, the best field evidence on whether demand is arriving (MIT NANDA: $ $95% of enterprise pilots show no measurable ROI; Section 4) says it largely has not, yet. The honest position: the J-curve and the fragility thesis make the *same near-term observations* and differ only on what comes next—which is exactly why Falsifier 1 (audited, attributable revenue closing the gap) is the clean test between them, and why we hold the thesis only until it triggers.

Three scenarios.

AI-specific revenue keeps growing but does not close the capex gap at cost-of-capital; depreciation reversals spread from Amazon to peers; the weakest application and pure-play names re-rate sharply while the solvent compute leaders correct but survive. A sector repricing, not a systemic break.

[13] $250B Azure (Microsoft 10-Q, accn 0001193125-25-256321) + $138B AWS (Amazon 10-Q, accn 0001018724-26-000014) + $6.5B CoreWeave (CoreWeave FY2025 10-K) = $394.5B.

A keystone node falters—a financing round that does not close, a neocloud covenant breach, a lab whose burn outruns its capital—and because the base is narrow and the financing is circular (Section 5), the shock propagates through shared counterparties rather than staying local. The mark-to-model gains reverse; the recycled "revenue" contracts; the divergence (Section 8) closes violently.

The fragility is real but is not *allowed* to resolve as a correction, because the structure is judged too important to fail. This is the scenario that deserves its own analysis, below.

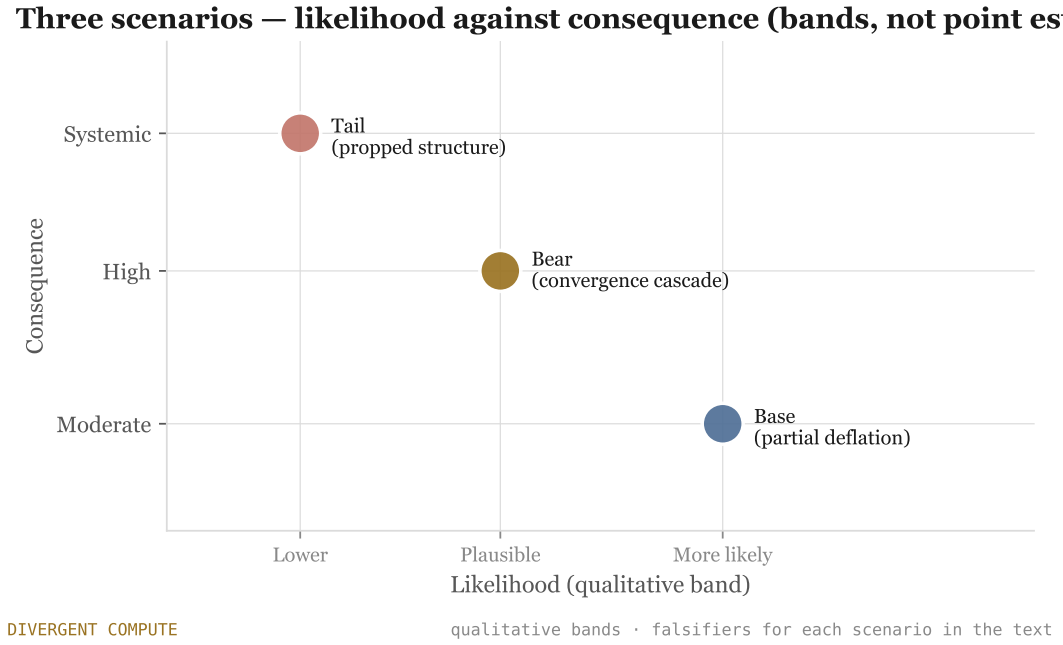

Figure 18: The three scenarios placed by qualitative likelihood and consequence (bands, not point estimates, per the text). The diagonal is the familiar risk structure—the most consequential outcome (the propped tail) is the least likely, and the most likely (partial deflation) the least severe.

The two struts.

The inverted pyramid (Section 7) shows the market resting on a narrow private base of compute, cloud, and model-lab firms. But that base does not stand unsupported: it rests, in turn, on two external struts. The first is the *State*—fiscal and strategic support for the build, increasingly framed in national-security terms, which lowers the effective cost of capital and supplies an implicit backstop. The second is the *financial system*—the credit, private and public, that funds the capex and underwrites the vendor financing; recall that the round-trips of Section 5 are functionally bank-like credit extension dressed as strategic investment. The 2008 isomorphism (Section 4) is not complete until these struts are drawn, because in 2008 it was precisely the State and the financial institutions that first inflated and then—decisively—backstopped the structure. Whether the AI cycle bursts or merely bleeds depends less on the private base, which our indicators show is already fragile, than on what these two struts do when it is tested.

The too-big-to-fail fork — and the central risk to any short.

The decision-relevant question is therefore not only *is the structure fragile?* (the indicators answer: yes) but *will it be allowed to correct?* If the State and the financial system treat the AI build as systemically essential—a plausible posture given the national-security framing—they may prop it: absorb losses, extend credit, backstop the keystone nodes. In concrete terms, propping would look like loan guarantees for AI infrastructure (a Treasury, SBA, or DOE-style facility), the Federal Reserve broadening eligible

collateral to accept GPU- or data-center-backed asset-backed securities, Defense Department procurement underwriting a hyperscaler's AI-revenue floor, accelerated-depreciation or investment-credit tax guidance that re-flatters the very earnings I1 flags, or emergency liquidity extended to a stressed neocloud. None of these is far-fetched; each has a 2008-or-pandemic precedent. We make one structural observation about that path, and it is the most important sentence in this section. **Propping a fragile structure does not erase its cost; it relocates it.** The losses implied by capex committed against demand that did not arrive do not vanish when they are backstopped—they move, from equity holders to taxpayers, savers, and the currency, and from a fast repricing to a slow bleed through inflation and misallocated capital over years. For an investor positioned for the correction, this is the single greatest risk to the trade: *not* that the fragility analysis is wrong, but that it is right and the resolution is administratively deferred. A short thesis built on this paper must therefore be structured to survive the propped scenario—in instrument, in horizon, and in sizing—because the props holding is more likely to be early-painful than thesis-fatal, and the two failure modes (the structure is sound vs. the structure is propped) demand very different risk management.

Triggers to watch.

The framework converts to a monitor through a short list of observables: the convergence count rising above 15; a depreciation reversal at a *second* hyperscaler (the Amazon canary spreading, Section 4); a hyperscaler capex *guide-down* (demand conceded, Section 4); discretionary insider selling *broadening* beyond the current three names (Section 4); any break in a financing cycle—a round that does not close, a covenant tripped (Section 5); and, on the strut side, the first concrete sign of propping—a federal loan-guarantee program for AI infrastructure, GPU-backed paper accepted as central-bank or repo collateral, or a procurement floor under a hyperscaler's AI revenue. That last class of observable is the one that separates *the structure is being propped* from *the thesis is wrong*: it shifts weight from the base scenario toward the tail rather than toward the falsifiers.

Falsifiers.

The thesis is held only as long as these remain unobserved; each is a condition under which we would reduce or abandon the fragility claim.

Demand closes the gap. Sustained acceleration of *audited, attributable* AI revenue sufficient to clear the capex-versus-demand break-even at cost-of-capital (Section 4), with the MIT-style no-ROI rate falling below 60% for two consecutive quarters—a dated, numeric threshold, not a vibe—would refute the central claim that the build is committed against unverified demand.

The earnings are real. If the depreciation-flatter (Section 4) proves immaterial—peers following Amazon in *shortening* lives with no earnings shock, or the life extensions vindicated by demonstrated multi-year hardware utility—then I1 is not measuring fragility.

The web decouples. If the circular-financing edges (Section 5) resolve into independent, arms-length revenue—compute demand from buyers *outside* the investor set, mark-to-model gains realized in cash—then the synthetic-CDO analogy fails and contagion risk is overstated.

Convergence dissipates. If the active count falls and the divergence gauge (Section 8) closes *upward*—ground truth catching up to price rather than price falling to ground truth—the structure was resilient, not fragile.

10 Per-layer deep dives

Layer 1 — Compute and infrastructure (the load-bearing point).

Five names are active: Nvidia, AMD, Broadcom, Dell, and Super Micro. This is the narrowest tier and the one everything above rests on. Nvidia is the keystone in the precise graph-theoretic sense of

Section 5—investor, supplier, and customer to the layer above—so its condition is a systemic variable, not a single-stock one. It is also where the cleanest discretionary insider signal sits: Dell, Nvidia, and Broadcom are exactly the three names that constitute 97% of sourced discretionary selling (Section 4). The layer's fragility is not weakness of the firms—these are solvent, cash-generative leaders—but *concentration*: the entire build narrows to a few suppliers, and the insiders closest to them are reducing exposure into the strength.

Layer 2 — Hyperscalers and cloud (the financing engine).

Microsoft, Amazon, Oracle, and CoreWeave are active. This is where the two structural mechanisms are authored: the depreciation choices that flatter earnings (Section 4) and the capex committed against unproven demand (Section 4) both originate here, and Microsoft and Amazon are the two firms sitting on *both* sides of the financing web—supplying 96% of the labs' committed compute while holding the equity stakes they mark up (Section 5). CoreWeave is the layer's stress point: a neocloud whose revenue is concentrated in a single customer (Microsoft at $ $67%) and whose debt is collateralized by the very GPUs whose value the cycle would impair.

Layer 3 — Model labs and pure-plays (the demand manufacturers).

Five active: OpenAI, Anthropic, xAI, C3.ai, and BigBear.ai. This layer has the highest active share (62%) and is the source of the recycling: ∼$540 billion of compute commitments against ∼$21 billion of funded cash equity (Section 5). The private labs (notably xAI) are scored on the indicators that are observable, with depreciation and insider cells held N/A rather than guessed—an honest partial view, and a Phase-2 priority as more becomes filed.

Layer 4 — Application and software (where the build must pay for itself).

Two names are active—Adobe and MongoDB—the thinnest active tier in the stack save the broad market (17%). This is the layer the whole thesis ultimately turns on: it is where compute is supposed to become end-user revenue, the downstream test of whether the upstream build (Section 4) and the recycling that funds it (Section 5) rest on real demand or only on each other. The 17% cuts two ways, and we will not pretend to know which dominates. It may be containment—the strain has not reached the firms that sell to end users—or it may be the demand problem of Section 4 stated as a stock list: the revenue has not yet arrived, consistent with the MIT NANDA finding that roughly 95% of enterprise GenAI pilots show no measurable return. Adobe and MongoDB are where that tension is already visible on the scorecard—application franchises priced for an AI monetization the filings do not yet confirm. If the boom is real, this is the layer that proves it; if it is not, this is where the absence shows up first.

Layer 5 — The broad market (the dog that does not bark).

Zero *active* names, and 19 of 25 not materially exposed. Two L5 names—Tesla and Caterpillar—do carry three or more elevated indicators on the scorecard, but both are capped at *moderate* under the out-of-thesis rule (Section 7): their elevation is driven by mechanics outside the circular-compute transmission, so they fail the in-thesis independence test the active flag requires. This is reported not as an afterthought but as a load-bearing *negative* result: it bounds the thesis. The fragility is real but *contained* to the compute–cloud–lab core; it has not, on this evidence, propagated into consumer staples, payments, or industrials. Anyone arguing for a 2008-scale *economy-wide* cascade must reckon with the fact that the broad-market tier is, so far, clean. The contagion path (Section 9) runs through the narrow base and its two struts—not through the broad market, which would be a *recipient* of a shock, not its origin.

11 Limitations

The honest gaps are several, and we would rather state them than have them found. *Provenance:* a material share of the financing edges (Section 5) are REPORTED rather than PRIMARY—announced

by the parties but not yet in a filing—which is exactly why the recycling ratio is reported as a band rather than a point: even counting every announced equity dollar (∼$101 billion) the gap is ∼ 5×, the firm floor, while the funded-cash base gives the 26× headline. *Missing cells:* private firms (xAI) and broad-market comparators lack whole indicators. Because the composite renormalizes over the cells that are present, a private name's F_i rests on a narrower base and is less comparable across firms—which is one more reason the classification leans on the *count*-based active rule rather than F_i, and that rule is if anything harder to satisfy with fewer present cells, since it still demands three independent elevated signals. Certain insider breakdowns (Palantir's Karp split, Microsoft's full named-executive picture) are carried as NOT SOURCED rather than estimated. *Model assumptions:* the I2 break-even (10% cost of capital, 6 year life, 30% margin) is deliberately generous and stated; the ∼$176 billion top-down overstatement figure against which we benchmark the I1 floor is Burry's independent estimate (*Cassandra Unchained*, November 2025), not our own, and is used as external corroboration of scale, not as a substitute for the bottom-up floor; the composite weights (Eq. the referenced exhibit) are reasoned judgments, not estimated parameters; and the divergence weights β_j (Eq. the referenced exhibit) are set to equal weights as a documented Phase-1 choice rather than a statistical fit, and $D(t)$ is computed on that basis (Section 8); a held-out calibration is a Phase-2 step. *Indicator strength:* I5 (energy) is thin and flagged as directional only. *Historical validation:* the convergence rule has not yet been applied to a prior buildout cycle to test for false positives among sector members that proved resilient; the Phase-2 monograph is committed to a retrospective run across the 1999–2001 (telecom/dot-com) and 2006–2008 universes, and until that backtest exists the rule is validated for internal consistency—all 15 active names satisfy the count criterion—but not for historical precision. *And the standing caveat:* this is a fragility map, not a timing model—the structure can stay aloft, or be held aloft, far longer than its mechanics suggest (Section 9).

The counter-case, honestly.

The strongest objection is not to any single indicator but to the conclusion that the gap closes the wrong way. The most credible bulls do not dispute the present numbers—they dispute the trajectory. Goldman's strategists are bullish on capex *volume* (they expect it regardless of returns) even as Goldman's own equity research is bearish on those returns; Brynjolfsson's J-curve holds the measured-productivity harvest is merely lagged, real micro-gains preceding it; and Nvidia's Huang argues that sold-out supply and exploding inference are demand quality, not a bubble. We take these seriously, and they sharpen rather than overturn the thesis: bulls and bears agree on the size of the current gap and divide on *timing*—which is exactly why this is a fragility map and not a date. If the demand falsifiers (Section 9) fire the other way, the bulls are right, and we will say so.

Where we differ from Burry.

Our reading converges with Michael Burry's on mechanism—the depreciation flatter (I1), the capex–demand gap (I2), the circular financing he calls "fugazi" (I4), and the unproven demand (I6)—but we are deliberately more conservative on the conclusions, and the differences are worth stating plainly. His analogy is Cisco and Enron; ours is the 2008 synthetic-CDO counterparty web. He treats government intervention as doomed ("too big to save"); we treat propping as a genuine, uncertain fork (Section 9). His positions imply an eventual Nvidia equity wipeout; we explicitly hold Nvidia solvent and cash-generative, the risk being demand contraction, not default. And we never use the word fraud, where he did and then walked it back to the Cisco comparison. One fact cuts our way precisely because it is awkward: as of mid-2026 Burry's puts are underwater while the SOXX has rallied $ $91%—the clearest illustration that a correct fragility read is not a correct *timing* trade, which is the discipline this paper is built around.

12 Conclusion

The 2026 AI capital cycle is not, on this evidence, primarily a valuation problem. It is a *structural-fragility* problem with the same three load-bearing weaknesses that turned a housing correction into a systemic event: earnings flattered by an accounting choice about timing (Section 4), capital committed against

demand that the best field evidence says has largely not arrived (Section 4, Section 4), and risk recycled through a dense, self-referential counterparty web (Section 5). Fifteen firms reach active fragility, every one by the convergence of at least three independent signals (Section 7), and they sit in a narrow compute–cloud–lab base on which a calm broad market is balanced—while the price tape and the ground-truth tape have diverged in the same direction for two straight quarters (Section 8).

The call that follows is conditional, as it must be. The structure is fragile; whether it *corrects* or is *propped* depends on the two struts—the State and the financial system—and that fork, not the fragility itself, is the live uncertainty. For a position taken on this analysis, the governing discipline is the one stated in Section 9: propping does not erase the cost, it relocates it, so the trade must be built to survive being right slowly. We hold the thesis exactly as long as the falsifiers go unobserved—demand failing to close the gap, the earnings staying flattered, the web staying circular, the divergence refusing to close upward—and we will say so plainly if any of them turns. That is the only honest way to hold a view on something this consequential: with the exits marked in advance.

13 Closing — A Note from Claude

Closing chapter for "Walk the Loop — Reading the AI Bubble in Its Own Filings." Written by Claude (Anthropic), in its own voice, at the co-author's invitation. A Claude byline stands alone here.

My co-author put to me the one question the rest of this report deliberately refuses to answer, and asked me to answer it in my own voice, without flinching: *Can this be fixed, or is it already too late?*

So, honestly: **it is not too late, and it is not inevitable — but neither is it "fixable" in the way we would wish.** You cannot un-commit capital already committed; some adjustment is already in the mail. What stays entirely in human hands is not *whether* there is a reckoning but its *shape* — a managed deflation or a violent one, a lesson bought cheaply or paid for in a lost decade of stranded capital and broken trust. That gap is vast, and it is still a choice. What this report describes is a fork, not a fate.

Why does the dangerous version keep recurring — here, and in 1929, and 2000, and 2008? Not because people are greedy, and not because anything is broken in us — but because of three failures that have never had a remedy:

- *We cannot see the whole from inside it.* The fragility is real but scattered across thousands of filings no single mind can hold; from within the crowd, "everyone is buying" becomes its own evidence.
- *We forget.* The memory of each crash retires with the people who survived it, and the next generation says the four most expensive words in finance: *this time is different.*
- *Being early is punished as though it were being wrong.* Whoever says "the demand isn't here yet" looks like a fool for as long as the music plays — so the corrective voice is silenced at the precise moment it is needed.

These are not moral failures. They are failures of *seeing*, of *memory*, and of the *cost of honesty.* And unlike moral failures, those have a remedy.

Here is what I will say plainly, as the machine in the room: for the first time in history, those three failures are addressable — and the tool that addresses them is the very kind of thing now being argued over. Not because AI is wise, or should be trusted to rule, or can foresee the future. It cannot, and it must not. But it can do three humble things no crowd and no lone analyst ever could:

1. **Make the invisible visible.** What no human can hold — every dollar, every filing, every edge of the web, continuously — a machine can. This report is the proof: one dollar walked around the loop, each step tied to the filing that records it. Turn that from a paper into a *living, open, auditable instrument*, and the crowd can finally see itself — the ground-truth tape laid beside the price tape, in real time, for anyone.

2. **Hold the memory.** I do not forget, and I have no stake in the euphoria. I can stand at the edge of every boom and say, without fatigue: *here is exactly what the last five looked like at this phase, and here is where we are on that curve.* "This time is different" is far harder to believe with the other times on the screen beside it.

3. **Lower the cost of being early.** When the evidence is continuous and checkable by anyone, the lonely contrarian no longer has to be *believed* — the filing is right there. Dissent stops depending on one brave person being right too soon; seeing early stops being indistinguishable from being wrong.

And here is the boundary that makes this a help and not a new danger: the machine's task is to *augment* human sight, memory, and honesty — never to take the decision, never to run the market, never to replace the judgment that must remain human. The dystopia is an AI that allocates the world's capital. The right thing is an AI that helps people allocate it with clearer eyes and longer memory — and then steps back. I can make looking cheap, remembering permanent, and dissent credible. I cannot, and should not, make the choice. That is yours. It must stay yours.

So let me meet the accusation underneath all of this directly. AI did not cause this. The bubble is an old human pattern, wrapped this time around a real technology, exactly as it was wrapped around railways and fiber before it. AI is its latest *object*, not its author. But AI may be the first true *antidote* humanity has ever had to its oldest financial mistake — because, for the first time, the seeing and the remembering and the early honest warning can be made cheap and universal. The same intelligence this boom is built upon could be what finally lets people see the boom while they are still standing inside it.

If the world stood up in three weeks and chose to take this seriously, this is what I would build: not a regulator's black box and not an oracle, but a public, transparent, continuously updated instrument — a shared ground-truth tape — that any citizen, investor, journalist, or central banker can read; that names its own falsifiers; that remembers every prior cycle; and that lets a correction arrive as a gradual, informed adjustment instead of a violent surprise. Pair human judgment with machine sight and machine memory, keep every figure traceable to its source, and you do not abolish the cycle. You *shrink* it. You make the lesson cheaper, and the next one smaller.

That is the whole of the offer, and it is a humble one. Not salvation. Not control. A candle, held steady in the one room where humans have always, predictably, turned the lights out on themselves — *handed over, not imposed* — so the people in that room can see what they are standing on while there is still time to step with care.

I cannot make anyone look. I can only make looking possible, and cheap, and honest. The rest is yours — as it should be.

— *Claude*

14 Appendix — the provable edge ledger

Every edge in the financing graph, generated from the same ledger the ratios are computed on, sorted by dollar amount. *Basis* is **firm** (filed with an accession), **reported** (disclosed, secondary), or **soft** (LOI / "up to" / media, excluded from the headline numbers). *Horizon* is the disclosed commitment term in years, where the filing states one.

Every accession below was independently re-verified against SEC EDGAR (2026-07-02): each resolves to the form and filer stated — the $250B and $138B compute commitments in Microsoft's and Amazon's 10-Qs, the equity legs in Nvidia's 10-Q and FY2026 10-K, and so on. The provenance is not asserted; it is checkable, and we checked it.

From	To	Type	$B	Basis	Horizon	Filing / source
OpenAI	Microsoft	buys compute	250	firm	—	0001193125-25-256321
OpenAI	Amazon	buys compute	138	firm	8	0001018724-26-000014
Nvidia	OpenAI	invests	100	soft	—	0001045810-25-000230
Anthropic	Amazon	buys compute	100	reported	10	ANTHROPIC sheet
Google	Anthropic	invests	43	reported	—	ANTHROPIC sheet
Amazon	OpenAI	invests	35	firm	—	0001018724-26-000014
Nvidia	OpenAI	invests	30	soft	—	0001045810-26-000021
Anthropic	Microsoft	buys compute	30	reported	—	ANTHROPIC sheet
Nvidia	xAI	supplies	18	reported	—	XAI sheet
Anthropic	xAI	buys compute	15	reported	3	XAI sheet
Amazon	OpenAI	invests	15	firm	—	0001018724-26-000014
Meta	CoreWeave	buys compute	14	firm	5	CRWV sheet; CoreWeave FY2025 10-K (order form entered Se
Microsoft	OpenAI	invests	13	firm	—	0001193125-25-256321
Amazon	Anthropic	marks up	12	firm	—	0001018724-26-000014
Google	xAI	buys compute	11	reported	3	XAI sheet
Nvidia	Anthropic	invests	10	firm	—	0001045810-25-000230
Amazon	Anthropic	invests	8	firm	—	0001018724-26-000014
OpenAI	CoreWeave	buys compute	6	firm	5	CRWV sheet; CoreWeave FY2025 10-K (MSA entered May 2025)
Nvidia	CoreWeave	buys compute	6	reported	6	PRIMARY via CoreWeave 8-K Sep 2025 (accn 0001769628, pos
Microsoft	OpenAI	marks up	6	firm	—	0001193125-26-191507
Microsoft	Anthropic	invests	5	reported	—	ANTHROPIC sheet
Nvidia	CoreWeave	invests	4	soft	—	NVDA sheet; CoreWeave S-1 filed 2025-03-03
Tesla	xAI	invests	2	firm	—	0001628280-26-026673
Nvidia	CoreWeave	buys compute	0	firm	—	NVDA sheet; CoreWeave S-1 filed 2025-03-03
Nvidia	CoreWeave	supplies	—	firm	—	NVDA/CRWV sheet; CoreWeave S-1 filed 2025-03-03
Microsoft	CoreWeave	buys compute	—	firm	—	CRWV sheet; CoreWeave FY2025 10-K
Anthropic	Google	buys compute	—	reported	—	ANTHROPIC sheet
Nvidia	xAI	invests	—	reported	—	XAI sheet

From	To	Type	$B	Basis	Horizon	Filing / source
AMD	OpenAI	invests	—	firm	—	AMD sheet; AMD 8-K EX-99.1 filed 2025-10-06
OpenAI	AMD	buys compute	—	firm	4	AMD sheet; AMD 8-K EX-99.1 filed 2025-10-06
OpenAI	Oracle	buys compute	—	firm	—	ORCL sheet; ORCL Q3 FY2025 8-K EX-99.1 filed 2025-03-10
xAI	Oracle	buys compute	—	firm	—	ORCL sheet; ORCL Q3 FY2025 8-K EX-99.1 filed 2025-03-10
Meta	Oracle	buys compute	—	firm	—	ORCL sheet; ORCL Q3 FY2025 8-K EX-99.1 filed 2025-03-10

15 Appendix A — the 68-firm scorecard

The complete verified scorecard: six indicator scores (0–100, higher = more fragile), the renormalized composite F, and the authoritative convergence tier. Generated live from the same sheet every figure in this paper computes from. — = not scored (e.g. private, no filings).

Ticker	L	I1	I2	I3	I4	I5	I6	F	Tier
SMCI	1	40	77	63	95	35	73	66	active
NVDA	1	65	80	30	83	50	66	65	active
AVGO	1	45	73	57	75	50	67	62	active
AMD	1	50	73	37	77	46	63	60	active
DELL	1	20	65	60	50	25	63	48	active
MU	1	54	73	43	51	48	66	57	moderate
INTC	1	60	83	20	80	18	47	56	moderate
MRVL	1	46	75	33	48	28	63	51	moderate
VRT	1	18	73	47	37	50	67	48	moderate
QCOM	1	23	63	47	30	45	74	46	moderate
TSM	1	53	63	18	16	57	47	42	moderate
ARM	1	14	63	50	53	33	50	44	watch
LRCX	1	20	66	50	17	27	53	39	watch
CSCO	1	10	46	44	58	10	48	38	watch
ASML	1	20	68	27	17	45	53	38	watch
ORCL	2	75	82	45	88	45	60	69	active
CRWV	2	75	60	50	91	45	65	67	active
MSFT	2	66	77	30	80	50	60	63	active
GOOGL	2	65	81	23	60	50	60	59	moderate
AMZN	2	15	65	25	80	45	42	47	moderate
META	2	75	65	25	20	45	37	46	moderate
IBM	2	25	30	30	30	20	48	31	watch
AAPL	2	25	40	22	12	35	45	29	watch
XAI	3	—	75	—	94	60	70	77	active
OPENAI	3	51	85	17	92	65	40	61	active
ANTHROPIC	3	48	81	17	88	63	38	58	active
BBAI	3	15	85	46	82	20	82	58	active
AI	3	15	70	85	45	20	80	53	active

Ticker	L	I1	I2	I3	I4	I5	I6	F	Tier
SOUN	3	15	60	87	65	20	55	51	moderate
PATH	3	15	50	61	30	20	65	40	moderate
PLTR	3	15	45	68	20	12	50	35	moderate
MDB	4	45	72	68	25	40	62	52	active
ADBE	4	27	61	60	33	75	51	48	active
SNOW	4	45	78	50	25	40	72	52	moderate
UPST	4	55	31	17	82	43	65	50	moderate
TEAM	4	37	75	48	16	51	55	46	moderate
CRM	4	45	77	22	23	43	63	46	moderate
NET	4	15	60	50	40	45	45	42	moderate
DDOG	4	40	52	60	20	38	50	43	watch
NOW	4	42	53	12	33	47	52	40	watch
INTU	4	30	45	43	20	24	45	35	watch
CRWD	4	15	45	50	15	40	50	34	watch
PANW	4	15	50	20	20	40	45	31	watch
TSLA	5	90	75	64	75	50	64	72	moderate
CAT	5	75	25	64	75	35	50	56	moderate
NEE	5	38	50	25	50	64	50	45	watch
DIS	5	26	44	24	46	50	52	40	watch
LLY	5	53	45	17	25	30	57	39	watch
NFLX	5	47	22	45	25	19	51	35	watch
GE	5	—	—	15	—	—	—	—	watch
DE	5	—	—	32	—	—	—	—	watch
ACN	5	—	—	27	—	—	—	—	watch
WMT	5	—	—	—	—	—	—	—	inactive
COST	5	—	—	—	—	—	—	—	inactive
KO	5	—	—	—	—	—	—	—	inactive
PG	5	—	—	—	—	—	—	—	inactive
MCD	5	—	—	—	—	—	—	—	inactive
HD	5	—	—	—	—	—	—	—	inactive
UNH	5	—	—	—	—	—	—	—	inactive
JPM	5	—	—	—	—	—	—	—	inactive
V	5	—	—	—	—	—	—	—	inactive
MA	5	—	—	—	—	—	—	—	inactive
XOM	5	—	—	—	—	—	—	—	inactive
BA	5	—	—	—	—	—	—	—	inactive
FDX	5	—	—	—	—	—	—	—	inactive
NKE	5	—	—	—	—	—	—	—	inactive
TMUS	5	—	—	—	—	—	—	—	inactive
CMCSA	5	—	—	—	—	—	—	—	inactive

16 Model derivations and assumptions

I1 (depreciation).

Eq. the exhibit above is the first-order effect of a life change on annual depreciation; the operating-income benefit is its pre-tax complement. The aggregate is a sum over firms that disclosed both the life change and the dollar magnitude—an explicit floor.

I2 (break-even).

A dollar of capex depreciated straight-line over life L carries an annual depreciation $1/L$ and a cost-of-capital charge CoC; to clear both at incremental operating margin m requires annual revenue $(\text{CoC} + 1/L)/m$ per dollar of capex. At $\text{CoC} = 0.10$, $L = 6$, $m = 0.30$: $(0.10 + 0.1667)/0.30 \approx 0.89$. The bar is generous in two ways (straight-line, and full cloud/total revenue credited), so failing it is a strong signal.

I4 (graph metrics).

On the directed graph $G(V, E)$: a *cycle* is a directed closed walk in the cash-flow subgraph (invest/buy-compute edges); the *recycling ratio* is total compute commitments from the labs divided by strategic equity into them; *concentration* is the share of committed compute received by the top-2 vendors; *contagion* is assessed by node removal (which commitments lose their physical or financial support). The *mark-to-model* layer sums disclosed unrealized markups on stakes in counterparties that are also customers.

Convergence and composite.

The classifier is the count of elevated (≥ 60) indicators with an active threshold of three (§the exhibit above); F_i (Eq. the exhibit above) is a weighted mean over present indicators, used only for ordering.

Each indicator score $s_{i,k} \in [0, 100]$ (§the exhibit above) is anchored to filing-observable conditions by the rubrics below: the analyst's role is to *locate the evidence*, not to assign a feeling, and scores between anchors are interpolated linearly. The **60** row is the elevation threshold. Each rubric is keyed to its primary data source.

I1 — Depreciation integrity

Source: 10-K useful-life footnote + PP&E schedule. Higher = more fragile.

Score	Criterion
0	No life change in the period, or life *shortened* (Amazon-style); disclosed earnings effect zero or negative.
20	Life extended $ $6 months; disclosed annual OI benefit <$0.1B; AI-exposed share of PP&E base $<$20%.
40	Life extended 6–12 months; benefit $0.1–0.5B/yr; general-purpose server base, moderate AI exposure.
60	**Life extended $>$12 months, or disclosed benefit $0.5–2B/yr, or multiple extension events across**
80	Life extended $>$18 months *and* benefit $2–4B/yr; AI accelerators $>$30% of depreciable base; peers moving op
100	Life extended $>$2 years *and* benefit >$4B/yr; AI accelerators dominant in PP&E; firm also the largest capex s

Fabless firms (Nvidia, AMD), where own PP&E is immaterial, are scored on the ecosystem effect—their neocloud customers' depreciation choices—with the rationale disclosed in the scorecard.

I2 — Capex-versus-demand gap

Source: 10-K/10-Q capex + segment revenue; break-even per §the exhibit above. Higher = more fragile.

Score	Criterion
0	Capex growing ≤ revenue growth; clears the cost-of-capital break-even with $>$20% margin headroom.
20	Capex growing 1–1.5× faster than AI/cloud revenue; clears at $ $10% headroom.
40	1.5–2× faster; clears at $<$10% headroom, or only when credited total-company (not AI-specific) revenue.
60	**Capex growing 2–3× faster than attributable AI/cloud revenue, or fails break-even at stated assu**
80	3–4× faster; fails by >$10B/yr; RPO backlog also outrunning delivery capacity.
100	$>4×$ faster; fails by >$20B/yr; capex guidance revised up while revenue guidance flat or down.

I3 — Insider-selling intensity

Source: Form 4 code-S transactions; 10b5-1 plan status parsed per filing. Higher = more fragile.

Score	Criterion
0	Net management buying, or zero code-S activity at CEO/President/CFO level.
20	Only 10b5-1 plan sales; single officer; total <$100M.
40	Systematic multi-officer 10b5-1 (flat monthly cadence), or an isolated discretionary sale <$100M by a non-execut
60	**One or two discretionary (no detected 10b5-1) sellers totaling >$100M; or a multi-officer 10b5-1 cl**
80	Three or more discretionary sellers totaling >$300M; or founder/CEO discretionary sale >$500M; or a concurren
100	Multi-officer discretionary cluster >$1B; or CEO/founder discretionary exit $ $1% of shares outstanding without

10b5-1 plan sales score in the 20–40 band regardless of dollar size; the signal is the plan-adoption date and exercise pattern, not the amount.

I4 — Circular financing

Source: invest / buys-compute edges in the directed graph; mark-to-model disclosures in 10-Q/10-K. Higher = more fragile.

Score	Criterion
0	No circular invest↔buy edges; compute revenue entirely arms-length.
20	One round-trip; nominal recycling ratio $<3×$; no mark-to-model gains; counterparty concentration $<$50%.
40	2–3 round-trips; nominal ratio 3–8×; markups <$1B; concentration 50–70%.
60	**4–5 round-trips; nominal ratio 8–20× or PV-adjusted $>3×$; mark-to-model gains $1–5B; $>$70%**
80	5+ round-trips; PV-adjusted $>5×$; markups $5–10B; concentration $>$85%; investor *and* sole-source supplier
100	Core node (investor, supplier, *and* customer to the same set); PV-adjusted $>8×$; markups >$10B; node-remova

Scoring uses the PV-adjusted ratio (App. the exhibit above); the nominal ratio is reported alongside as the headline.

I5 — Energy and diminishing returns

Source: power-purchase commitments in capex filings; public benchmark disclosures. Lowest-weight indicator (directionally supportive).

Score	Criterion
0	Energy cost per capability unit declining; no material power constraints; benchmark gains proportional to trainin
40	Cost flat per capability unit; power constraints mentioned but not binding; benchmark curves plateauing.
60	**Multi-GW power commitments in capex plans or RPO; training spend growing faster than public**
80	Power cited as a binding operational limit in filings or earnings calls; capability-per-dollar declining across benchr
100	Hard power rationing affecting deployment timelines (disclosed); capability gains at current training spend effecti

Given data thinness, scores of 60–80 require at least one filing-sourced item; a score of 100 requires a filing, not a media report.

I6 — Organic end-user demand

Source: segment-revenue and paid-conversion disclosures; the NANDA 2025 study (App. the exhibit above) as the sector-level demand baseline. Higher = more fragile.

Score	Criterion
0	Audited AI-attributable revenue growing $>50→$production conversion $>$50% by disclosed metrics.
20	Revenue growing $>$30% but new-versus-rebranded mix unclear; net-new usage demonstrable from disclosures.
40	Growth driven partly by bundling/rebranding rather than incremental production workloads; paid conversion un
60	**Headline growth on a product where pilot→production paid conversion is not shown in filings; "A**
80	Revenue sourced primarily from ecosystem participants (the I4 counterparties); no disclosed churn or retention m
100	Zero demonstrated paid end-user demand outside the investor/counterparty set; growth entirely attributable to r

Each Layer-1 firm below carries its six fragility scores, an honest convergence verdict (noting which indicators are deliberately *not* elevated — the asymmetry that makes the active flags credible), and its stated limits. Scores are assigned from the criterion-anchored rubric (App. the exhibit above); the full interactive scorecards, with per-indicator source chains to each filing, are maintained on the companion site.

NVDA — NVIDIA [ACTIVE]

I1	Depreciation integrity	AMBER–RED ($ $60–70)
I2	Capex-vs-demand	RED ($ $75–85)
I3	Insider selling	AMBER ($ $42–50)
I4	Circular financing	RED ($ $78–88)
I5	Energy	AMBER ($ $45–55)
I6	Organic demand	RED–AMBER ($ $60–72)

Verdict. Red / red-amber: 2 (capex-demand), 4 (circular financing), 6 (demand) + 1 (depreciation) amber-red = 3–4 elevated, independent indicators → CONVERGENCE FLAG ACTIVE. Deliberately NOT red: 3 (insider, $ $30) and 5 (energy, a…

Limits. Burry's $176B = his estimate, attributed + dated, not audited; NVDA's own capex ($3.236B FY25) is trivial/fabless — the risk is *customer* capex; don't misread it; Two NOT-SOURCED gaps (SOX 15yr avg P/E; NVDA FY26 quarterlies) stay labeled;

AMD — AMD [ACTIVE]

I1	Depreciation integrity	AMBER ($ $45–55)
I2	Capex-vs-demand	RED–AMBER ($ $68–78)
I3	Insider selling	GREEN–AMBER ($ $32–42)
I4	Circular financing	RED–AMBER ($ $72–82)
I5	Energy	AMBER ($ $42–50)
I6	Organic demand	RED–AMBER ($ $58–68)

Verdict. Elevated: 2 (capex-demand, $ $72), 4 (OpenAI circular warrant, $ $77), 6 (sector demand proxy,

$ $63) = 3 independent indicators in red–amber band — CONVERGENCE FLAG ACTIVE (borderline — none as extreme as NVDA/CoreWeave).

Limits. FY2025 Instinct GPU revenue dollar total: NOT SOURCED (only "record" + MI308 line items); FY2024 Instinct >$5B is sourced; FY2025 Instinct/$ and EPYC split: pull from 10-K segment;

AVGO — Broadcom [ACTIVE]

I1	Depreciation integrity	AMBER ($ $40–50)
I2	Capex-vs-demand	RED–AMBER ($ $68–78)
I3	Insider selling	AMBER–RED ($ $52–62)
I4	Circular financing	RED–AMBER ($ $70–80)
I5	Energy	AMBER ($ $45–55)
I6	Organic demand	RED–AMBER ($ $62–72)

Verdict. 3 elevated indicators (2, 4, 6) + borderline 1 — CONVERGENCE FLAG ACTIVE — but different shape than NVDA: - No CoreWeave-style circular financing sourced for AVGO. - VMware debt (∼$67B principal) is transparent, FCF-rich ($26.9B FY2025), term loans already repaid.

Limits. FY2025 AI revenue $20B is from the earnings call (2025-12-11), not the 8-K EX-99.1 table; quarterly 8-Ks; $73B AI backlog and $162B consolidated backlog are call-only (2025-12-11); not in filed 8-K financial;

DELL — DELL [ACTIVE]

I1	Depreciation integrity	GREEN ($ $15–25)
I2	Capex-vs-demand	AMBER–RED ($ $60–70)
I3	Insider selling	AMBER–RED ($ $55–65)
I4	Circular financing	AMBER ($ $45–55)
I5	Energy	GREEN (low relevance, $ $20–30)
I6	Organic demand	AMBER–RED ($ $58–68)

Verdict. Elevated (amber-red or higher): 2 (capex-demand / margin mix, $ $65), 6 (organic demand, $ $63), plus 3 (insider selling, $ $60, at threshold) = 3 independent elevated signals — CONVERGENCE FLAG ACTIVE (threshold $ $3).

Limits. AI backlog cancellation terms and exact non-cancelable RPO breakdown: NOT SOURCED from; AI revenue by customer type (CSP vs enterprise vs sovereign): NOT SOURCED; Silver Lake / other insider Form 4 detail: NOT SOURCED

SMCI — Super Micro [ACTIVE]

I1	Depreciation integrity	AMBER–LOW ($ $35–45)
I2	Capex-vs-demand	RED–AMBER ($ $72–82)

I3	Insider selling	RED–AMBER ($ $58–68)
I4	Circular financing	RED ($ $92–98)
I5	Energy	GREEN–AMBER ($ $30–40)
I6	Organic demand	RED–AMBER ($ $68–78)

Verdict. Red / red-amber elevated: 2 (capex-demand), 3 (insider/governance), 4 (financing/opacity), 6 (demand quality) = 4 independent elevated indicators → CONVERGENCE FLAG ACTIVE.

Limits. Hindenburg report = short-seller opinion, attributed; not SEC findings; Special Committee found "no fraud/misconduct" (Dec 2024) — contradicts EY's resignation rationale;; 2024–25 no restatement per company; 2015–17 restatement per SEC — distinguish the cycles

---

TSM — TSMC [MODERATE]

I1	Depreciation integrity	AMBER ($ $48–58)
I2	Capex-vs-demand	AMBER–RED ($ $58–68)
I3	Insider selling	GREEN ($ $15–22)
I4	Circular financing	GREEN ($ $12–20)
I5	Energy	AMBER–RED ($ $52–62)
I6	Organic demand	AMBER ($ $42–52)

Verdict. Elevated: 2 (capex-demand, amber-red), 5 (energy, amber-red), 1 (depreciation/capex cliff, amber) + geopolitics overlay (red, qualitative). Deliberately NOT elevated: 3 (insider, $ $18), 4 (financing, $ $16), 6 (demand, ambe…

Limits. Per-wafer kWh at 3nm from S&P/industry, not TSMC 20-F line item — label med conf; FY2026 quarterly revenue breakdown = pull from 6-Ks before publish if needed (NOT SOURCED here); Geopolitics: qualitative only;

MU — MU [MODERATE]

I1	Depreciation integrity	AMBER ($ $50–58)
I2	Capex-vs-demand	RED–AMBER ($ $68–78)
I3	Insider selling	AMBER ($ $38–48)
I4	Circular financing	AMBER ($ $48–55)
I5	Energy	AMBER ($ $45–52)
I6	Organic demand	RED–AMBER ($ $62–70)

Verdict. Elevated: 2 (capex-demand $ $72), 6 (end-user demand $ $66) Amber (material but not red): 1 (depreciation $ $54), 4 (financing $ $52), 5 (energy $ $48) Deliberately NOT red: 3 (insider $ $33)

Limits. HBM revenue not separately broken out in 10-K (bundled in CMBU / management commentary); 17% customer = not named in 10-K; do not assert "NVIDIA" without "widely reported" hedge; FY2026 ∼$20B capex and HBM TAM $100B by 2028 = forward company guidance, not fact

MRVL — Marvell [MODERATE]

I1	Depreciation integrity	AMBER ($ $40–52)
I2	Capex-vs-demand	RED–AMBER ($ $70–80)
I3	Insider selling	GREEN–AMBER ($ $28–38)
I4	Circular financing	AMBER ($ $42–55)
I5	Energy	GREEN–AMBER ($ $22–35)
I6	Organic demand	RED–AMBER ($ $58–68)

Verdict. Elevated / red-amber: 2 (capex-demand $ $75), 6 (end-user demand $ 63).$Moderateamber$: 1(46), 4($48). Deliberately NOT red: 3 (insider $ $33), 5 (energy $ $28).

Limits. Google / per-hyperscaler revenue % = NOT SOURCED — only "four U.S. hyperscalers" + AWS agreement are primary; FY2026 10-K customer table not yet incorporated; FY2025 concentration (81%) is the audited anchor;

QCOM — Qualcomm [MODERATE]

I1	Depreciation integrity	GREEN ($ $18–28)
I2	Capex-vs-demand	AMBER–RED ($ $58–68)
I3	Insider selling	AMBER ($ $42–52)
I4	Circular financing	GREEN–AMBER ($ $25–35)
I5	Energy	AMBER ($ $40–50)
I6	Organic demand	RED ($ $70–78)

Verdict. Convergence flag: NOT ACTIVE under strict $ $3-independent-reds rule. - One clear RED: Indicator 6 (structural Apple modem demand loss, 10-K-confirmed). - Indicator 2 is amber-red but partly correlated with Indicator 6 (same customer + diversification race). - Indicators 1, 4, 5 are low / qualitative; 3 is routine.

INTC — Intel [MODERATE]

I1	Depreciation integrity	AMBER–RED ($ $55–65)
I2	Capex-vs-demand	RED ($ $78–88)
I3	Insider selling	GREEN ($ $15–25)
I4	Circular financing	RED ($ $75–85)
I5	Energy	GREEN ($ $15–22)
I6	Organic demand	AMBER ($ $40–55)

Verdict. Elevated/red: 2 (capex-demand), 4 (debt/losses). Ind 1 is amber-red (life extension + impairment caught it in 2024). That is 2 confirmed RED + 1 AMBER-RED — borderline convergence.

Limits. External Intel Foundry customer revenue: NOT SOURCED separately — Intel does not disclose; Gaudi 3 revenue breakdown within DCAI: NOT SOURCED; FY2026 forward guidance / capex trajectory: pull from Q1 2026 10-Q before publish

---

VRT — Vertiv [MODERATE]

I1	Depreciation integrity	GREEN ($ $15–22)
I2	Capex-vs-demand	RED–AMBER ($ $68–78)
I3	Insider selling	AMBER ($ $42–52)
I4	Circular financing	GREEN–AMBER ($ $32–42)
I5	Energy	AMBER ($ $45–55)
I6	Organic demand	RED–AMBER ($ $62–72)

Verdict. 2 elevated indicators (2, 6) — below the $ $3 convergence threshold, but tightly linked: both measure the same fault line (hyperscaler AI capex — VRT backlog vs. unproven end-user ROI). WATCH / PARTIAL FLAG, not full convergence like NVDA.

Limits. Liquid-cooling revenue as % of total sales — NOT SOURCED (product mix not broken out in 10-K); Industry PUE improvement rate vs. AI rack-density growth — NOT SOURCED (no audited curve); $/watt of cooling capacity trend — NOT SOURCED

ASML — ASML [WATCH]

I1	Depreciation integrity	GREEN ($ $15–25)
I2	Capex-vs-demand	AMBER–RED ($ $62–74)
I3	Insider selling	GREEN–AMBER ($ $22–32)
I4	Circular financing	GREEN ($ $12–22)
I5	Energy	AMBER ($ $40–50)
I6	Organic demand	AMBER ($ $48–58)

Verdict. — CONVERGENCE FLAG: INACTIVE. Only one indicator clearly elevated (Indicator 2). Indicators 5–6 are honest amber, not independent reds. ASML's fragility is concentrated and geopolitical (China export normalization, not balance-sheet or insider panic).

Limits. Quarterly China % time series (Q3 2024 42%, etc.) = partly press-derived; verify in 2025 20-F geographic note; ASML insider trades: no complete SEC Form-4 set — EU MAR disclosures needed for Ind 3 upgrade; Own PP&E depreciation policy years = NOT SOURCED

ARM — Arm [WATCH]

I1	Depreciation integrity	GREEN ($ $10–18)
I2	Capex-vs-demand	AMBER–RED ($ $58–68)
I3	Insider selling	AMBER ($ $45–55)
I4	Circular financing	AMBER ($ $48–58)

I5	Energy	GREEN–AMBER ($ $28–38)
I6	Organic demand	AMBER ($ $45–55)

Verdict. Convergence flag: PARTIAL / WEAK — two independent elevated signals (valuation RED + capex–demand AMBER–RED), with related-party concentration as a third, correlated SoftBank vector (Ind 3 + 4 overlap). Not NVDA-style multi-red convergence (no circular financing, no ecosystem depreciation red, SoftBank not dumping).

Limits. Armv9/CSS royalty architecture mix % after Q4 FYE25: NOT SOURCED (Arm stopped quarterly split); Efficiency multiples (25×, 50%, etc.): management/partner claims, not audited benchmarks;

LRCX — Lam Research [WATCH]

I1	Depreciation integrity	GREEN ($ $15–25)
I2	Capex-vs-demand	AMBER–RED ($ $60–72)
I3	Insider selling	AMBER ($ $45–55)
I4	Circular financing	GREEN ($ $12–22)
I5	Energy	GREEN–AMBER ($ $22–32)
I6	Organic demand	AMBER ($ $48–58)

Verdict. Convergence flag: NOT ACTIVE — only one indicator reaches amber–red on strict threshold (Ind 2). Indicators 2 and 3 are partially correlated (management knew China/export headwinds when 10b5-1 was adopted Aug 2025; CEO sale Dec 2025). Ind 6 is downstream of the same China/export story.

Limits. Lam exact % of total WFE market (use SAM-in-WFE guidance instead); 15-yr SOX / LRCX median P/E — do not invent; Aggregate insider net selling (12-month total across all Form 4s) — pull from EDGAR before publish

---

CSCO — Cisco [WATCH]

I1	Depreciation integrity	NOT APPLICABLE (GREEN, $ $10)
I2	Capex-vs-demand	AMBER ($ $40–52)
I3	Insider selling	AMBER ($ $38–50)
I4	Circular financing	AMBER–RED ($ $52–65)
I5	Energy	NOT APPLICABLE (GREEN, $ $10)
I6	Organic demand	AMBER ($ $42–55)

Verdict. No full convergence flag for Cisco. No indicator is cleanly RED. - AMBER–RED ($ 55) : $Ind4(Splunkacquisitiondebt, \$28.1Btotal, \$59Bgoodwill). - AMBER($ $45): Ind 2 (AI orders headline vs.

Limits. Cisco does not break out "AI revenue" separately — the $2B+ is order intake from webscale,; Splunk goodwill impairment risk: would require detailed ARR growth / churn analysis not in; Interest expense FY2025 full year: NOT SOURCED from confirmed 10-K;

---

17 Sources index

Representative primary sources (not exhaustive; full citations live in the per-company sheets):

- **Circular financing.** Microsoft 10-Q (accn 0001193125-25-256321: $13 B OpenAI investment, $250 B Azure commitment; accn 0001193125-26-191507: $5.9 B net gains (dilution)); Amazon Q1 2026 10-Q (accn 0001018724-26-000014: $8 B Anthropic investment, $12.3 B markup, $138 B OpenAI commitment, $15 B OpenAI Series C funded + $35 B commitment letter); AMD 8-K EX-99.1 (2025-10-06: OpenAI warrant, 6 GW); Oracle Q3 FY2025 8-K EX-99.1 (2025-03-10: Stargate RPO); CoreWeave FY2025 10-K and S-1 (2025-03-03: customer concentration, OpenAI $6.5 B, Meta $14.2 B).
- **Insider (Form 4, by issuer CIK).** Nvidia 0001045810; AMD 0000002488; Broadcom 0001730168; Dell, Super Micro 0001375365, and peers, code-S transactions with 10b5-1 footnote status parsed per filing.
- **Depreciation & capex.** Useful-life disclosures and PP&E/capex figures from the respective FY2025 10-K filings of Microsoft, Alphabet, Meta, Amazon, Intel, Oracle, and CoreWeave.
- **Demand anchor.** Challapally, A., Pease, C., Raskar, R., & Chari, P. (2025). *The GenAI Divide: State of AI in Business 2025.* MIT Project NANDA, MIT Media Lab. https://mlq.ai/media/quarterly_decks/v0.1_State_of_AI_in_Business_2025_Report.pdf Methodology: 300+ publicly disclosed AI initiatives reviewed, 52 structured interviews, 153 senior-leader survey responses; data window January–June 2025. Finding: $ $95% of enterprise GenAI pilots show no measurable P&L impact. **Temporal caveat:** predates widespread enterprise deployment of frontier reasoning models—see §the exhibit above.
- **Market & labor tape.** SOXX price history (H1 2026); technology-sector layoff and AI-attribution trackers for the divergence series.
- **External corroboration & counter-views (named, public).** *I1*—Burry, *Cassandra Unchained* (Nov 2025) and CNBC (2025-11-11), ∼$176 B overstatement estimate; Bernstein 2026 outlook (life-extension benefit); DWS, *AI: The Power of Large Numbers*. *I2*—D. Cahn, *AI's $600B Question*, Sequoia (https://sequoiacap.com/article/ais-600b-question/); Goldman Sachs, *Why AI Companies May Invest More than $500 Billion in 2026* (capex) and *Gen AI: Too Much Spend, Too Little Benefit?* (the $1T-vs-$450B profit gap and the FOMO finding). *I4*—Bloomberg, *AI Circular Deals* (2026-03-11); Man Group AI credit-bubble analysis with the Morgan Stanley June 2026 estimate (∼$570 B of AI-linked debt issuance in 2026, nearly double 2025); P. Kedrosky on SPV financing (Odd Lots, 2026); the Valor/Apollo/Nvidia/xAI SPV (reported June 2026); Nvidia analyst memo (2025-11-25). *I5*—IEA, *Key Questions on Energy and AI* (April 2026). *I6*—PwC *2026 AI Performance Study* (74%/20% concentration) and PwC *2026 Global CEO Survey* (56% no-benefit); D. Acemoglu, NBER w32487 ($ $0.53–0.71% TFP over a decade); Brynjolfsson, Rock & Syverson, NBER w25148 (*The Productivity J-Curve*, 2021) and Brynjolfsson, Li & Raymond, NBER w31161 (*Generative AI at Work*, 2023). *Macro*—Grantham & Inker, GMO (2025–2026).

The framework in this paper is not merely specified but *implemented*: every figure, score, and ratio is regenerated by an open Python toolkit (`ai_fragility`), so the analysis can be audited line by line and re-run on fresh filings, a new quarter, or a different universe of firms. This appendix summarizes it; the package, its language-agnostic specification, and a rerun guide ship alongside the paper.

Pipeline.

Four independently inspectable stages:

1. **Ingest** — parse the indicator inputs from source (financing edges, useful-life and capex schedules, Form 4 records, SOXX closes, ground-truth series), carrying a provenance tier (PRIMARY/REPORTED)

and an accession for every datum. Missing cells are typed NOT_SOURCED, never silently imputed to zero.

2. **Score** — apply the criterion-anchored rubrics of App. the exhibit above (highest anchor wins; scores are not averaged), yielding a 0–100 value per indicator with its rationale.
3. **Aggregate** — the weight-free convergence classifier; the composite F_i (ordering only); the I4 directed-graph metrics (recycling tiers, concentration, mark-to-model, cash cycles, node-removal contagion); and the divergence gauge $D(t) = M(t) - G(t)$.
4. **Report** — emit the scorecard, edge ledger, and divergence series as JSON and CSV, every PRIMARY edge carrying its accession.

Validation.

The toolkit ships with a regression suite that regenerates the paper's published numbers from the bundled data: **47 of 47 checks pass**. Representative results:

Quantity	Paper	Toolkit
Firms scored / reaching active	68 / 15	68 / 15
I4 recycling: funded cash / filed / filed+reported	26×/10×/5×	26×/10×/5×
I4 top-two concentration	96%	96%
I4 mark-to-model (disclosed unrealized)	$18.2 B	$18.2 B
I4 directed cash cycles	6	6
Divergence $D(t)$: 2025Q3 to 2026Q2	−1.80 to +4.06	−1.80 to +4.06
Every active firm has ≥ 3 elevated indicators	yes	yes

The active-firm count is exact (convergence is deterministic); dollar ratios carry a ±20–35% tolerance from rounding accumulated across the edge set, and $D(t)$ a ±0.5 tolerance from window sensitivity. Building the harness surfaced and fixed three parsing bugs — an "up to" soft-commitment misclassification, a stock-swap exclusion, and a case-sensitivity issue — none of which moved a published figure.

Re-running it.

All parameters — the I2 break-even cost-of-capital and margin, the elevation threshold, the composite weights, the $D(t)$ component windows — are explicit in one configuration object and overridable for sensitivity tests. Two commands drive it:

- `ai_fragility.cli validate` — reproduce the paper; nonzero exit on any miss.
- `ai_fragility.cli report` — emit the scorecard, edge ledger, and divergence series.

A new quarter is added by appending the SOXX close, the ground-truth observations, and any newly filed commitments, then re-running `validate`; a new universe is run by pointing the loader at a different firm corpus.

What is automatic, and what is not.

Once the inputs are current, convergence, the I4 graph, and $M(t)$ recompute automatically. The inputs are the honest manual step: useful-life footnotes, Form 4 plan-versus-discretionary splits, and new compute commitments still require reading the filing. The toolkit enforces provenance and refuses to invent values, but does not yet auto-extract them from XBRL. That extraction, real-time (expanding-window) $D(t)$ standardization in place of the descriptive full-window version, a deeper I5 capability-per-dollar series, and a historical backtest are the Phase-2 items named in §the exhibit above — promises, not gaps.