AI’s Four Shakeouts

Each layer sorts on the schedule its own capital allows

Mike Edholm

edholms.com

Almost everyone agrees a shakeout is coming for artificial intelligence. Thousands of companies have been funded against a market that will not support thousands of companies, and most of them will disappear, as most of the companies in every comparable rush have disappeared. That much is agreed, and agreement on it settles nothing that matters.

The interesting question is how long the sorting takes. A shakeout that resolves in three years and one that grinds on for twenty-five imply opposite strategies for everyone inside the industry and everyone lending to it. It also decides how to read the news. When the share price of a company at the center of the buildout drops thirty percent in two weeks, is that a sorting beginning, or a market repricing an expansion that has not slowed at all?

The answer is that AI will not undergo one shakeout on one timetable. Its chip layer is already highly concentrated. Its application layer may churn for years without settling. And its infrastructure layer, the data centers and power and interconnects, could endure the longest and most destructive period of all, because its capital is the hardest to repurpose once demand disappoints.

Two conditions

A shakeout ends when the field settles into the small set of names that will hold it, and getting there takes two separate things to happen. Call them the divergence condition and the exit condition.

Someone has to pull decisively clear. That is the divergence condition, governed by increasing returns — scale economies, network effects, installed base, the compounding advantage of the firm that already produces most and knows most. Where returns are strong the leaders separate quickly. Where they are weak, no one establishes a position that makes continuing look hopeless to everyone else.

Then the losers have to go. That is the exit condition, and it is governed by something the literature on shakeouts has largely left alone: whether the capital they committed can be sold to anyone outside their own trouble.

Exit is the neglected half, and it is the reason two industries hit by the same shock at the same moment can take three years and thirty years to settle.

Portable and anchored

Capital is portable when a buyer outside the industry’s difficulty will pay something close to what it cost. It is anchored when its worth depends on the position that created it, so that prizing it out of that position destroys most of the value.

Size has nothing to do with it, and neither does whether the asset can physically be moved. A vineyard is heavy, immobile, and thoroughly portable: the land and the winery plant find buyers whether or not this particular producer survives, which is why wine shakeouts finish in a year or two despite vineyards being as capital-heavy as agriculture gets. An airline's slot portfolio is a set of entries on a schedule, weighs nothing, and is anchored absolutely, because its value comes from the rest of the network that carrier assembled.

Nor does the test require physical plant. An operating system's installed base costs enormous sums to build and is worth nothing to a rival, because what the vendor owns is a relationship with its users, and that is precisely what a sale cannot carry across.

Whether a given industry's capital is portable or anchored can be settled from evidence dated at or near its peak firm count, using four indicators. Did a secondary market publish prices for the industry's principal asset? Did lenders advance against it at a real rate, and did that survive the downturn? Did purchasers of exiting firms' assets come from outside the industry, or only from rivals inside it? Was the plant ever documented as turned to another product?

All four ask the same question in different registers: was there a buyer who was not himself a distressed rival. Three or four yes is portable; zero or one anchored. A score of two usually means the asset had a market in good conditions that closed in bad ones, and for such an industry the test makes no prediction.

The mechanism runs through the exit decision, one firm at a time. A firm stops when continuing is worth less than stopping, and what stopping is worth is whatever the assets fetch. An owner with a resale price is choosing between operating profits and a check, and the check raises the bar the business has to clear. An owner with no buyer is choosing between operating profits and nothing, so almost any revenue above running costs argues for another year. The same distress produces an exit in the first case and a continuation in the second.

The record

Four industries have been tracked from commercial inception with complete counts of firms, entry and exit: automobiles, tires, television receivers and penicillin. Klepper and Simons compiled them, and they are the best test available.

Industry	Peak	Decline
Automobiles	272 firms, 1909	28 by 1929; 9 by 1941
Tires	274 firms, 1922	51 by 1940; 24 by 1970
Television receivers	95 firms, 1950	35 by 1960; 20 by 1970
Penicillin	1952	roughly 70% fewer producers by the early 1980s

Television sorted fastest, losing 63% of its producers in a decade. Its entrants were largely radio manufacturers whose assembly capacity moved between the two products, so the capital was portable, and the same prior experience gave those firms an advantage that compounded from the start. Strong returns, portable assets, quick resolution.

Automobiles and tires took twenty years to shed most of their firms and decades more to finish. Their plant, dies and molds were worth scrap to anyone else, and buyers were overwhelmingly rivals in the same difficulty. Only about 6% of automobile exits came through

acquisition, so these firms were not acquired and kept running. They operated until they could not.

Penicillin is the case that shows both conditions are needed. Fermentation capacity runs other antibiotics and other products, so the capital was portable and exit was available. Yet penicillin sorted slowest of the four, shedding 70% across three decades, with acquisitions accounting for only 3% of exits. Portability alone predicts the opposite. What happened is that the product commoditized and nobody pulled decisively clear, so divergence never really started. Exit stayed cheap and optional, and cheap optional exit produces attrition rather than a sorting.

The PC transition makes the same point at layer level. Chips and operating systems held deeply anchored capital and settled within about a decade of 1981, because returns were strong enough to make continuing hopeless before the anchoring could matter. Hardware assembly held portable capital — contract manufacturing and working capital — and ran twenty-five years without ever concentrating, because nothing compelled anyone to leave.

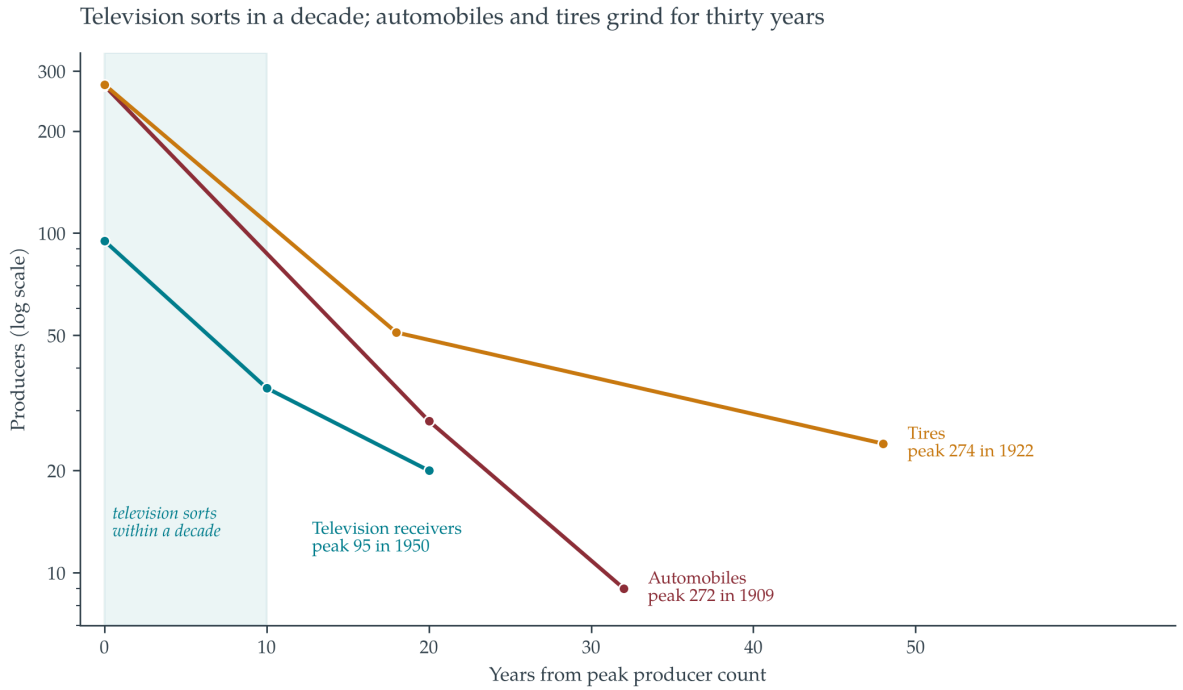

Figure 1. Producer counts after the peak, log scale. Television, the most portable capital of the three, sorts fastest. Penicillin is omitted for want of a continuous published count; it also had portable capital and sorted slowest of the four, which is why portability alone does not settle the question. Counts from Klepper and Simons.

Four cells

Cross the two variables and the cases separate cleanly.

- **Strong returns, portable capital.** The fastest and cleanest sorting available. Television receivers.
- **Strong returns, anchored capital.** Fast divergence, then a long tail of firms that cannot afford to stop. Semiconductors.
- **Weak returns, anchored capital.** The long grind: chronic overcapacity, prices below cost, a fragmented survivor set. Automobiles, tires, and long-haul fiber after 2001.

- **Weak returns, portable capital.** Continuous attrition that never resolves into a settled structure. Penicillin, and PC assembly.

AI’s four layers

Counting starts at the inflection point, the moment a proven approach becomes industrial. For AI that is ChatGPT’s launch in November 2022, which turned a specialist’s tool into a mass-market one within weeks. The catalyst came a decade earlier, with AlexNet’s result at the 2012 ImageNet competition.

What has been committed since is unusual in kind, not only in scale. Van Nieuwerburgh’s work on financing the buildout argues that the resulting wave of physical capital formation resembles railroads and electrification more than conventional technology-sector spending. Training a model requires no such commitment; the infrastructure underneath it does.

That capital is not distributed evenly, and neither are the returns. On current evidence the four layers occupy four different positions in the grid, though one of them sits on a boundary rather than squarely in a cell, and the arrangement holds only for as long as the layers stay separate.

Chips and fabrication. Anchored, with returns about as strong as they come: design amortization, foundry scale, and a software ecosystem that raises the cost of switching. Divergence is largely complete, and what remains is a contest among challengers that should resolve early.

Data centers, interconnects and power. Anchored, and returns are weak. A shell with a substation has no network effect and no ecosystem; it is real estate with an electrical connection, and the tenth one is worth no more than the second. That is the long-grind cell, where fiber sat after 2001. Expect chronic overcapacity, a fragmented survivor set, and the slowest schedule in the stack. This layer carries the most capital and is the reason a twenty-five-year figure is defensible at all.

Frontier models and the platforms serving them. Anchored, since weights are worth little to a rival and the value lives in distribution and tooling. Returns are the hard call, and this is the layer that sits on a boundary rather than inside a cell: real, but weaker than an operating system’s, because swapping an API costs a fraction of what changing an operating system cost. Read as strong returns, the layer belongs with chips and should resolve early. Read as weak, it belongs with data centers and grinds. What the classification does predict either way is the tail, because anchoring is not in doubt here: the losers’ capital is already spent, salvage is close to nil, and sponsors with nothing to recover keep funding. Moderate divergence, then a long tail, is the reading offered here — and it is the placement in this essay most likely to be wrong.

Applications. Portable and weak returns, which is what today’s evidence shows. Engineers, rented compute and customer relationships all transfer, so exit costs little, and nothing yet compounds the way an installed base or a foundry does. That is the continuous-attrition cell. The prediction rests on returns staying weak, and a submarket could develop returns nobody prices yet — a proprietary data position, a distribution lock, a switching cost that turns out to bind. It would then sort quickly rather than slowly, since portable capital with strong returns is the fastest cell in the set.

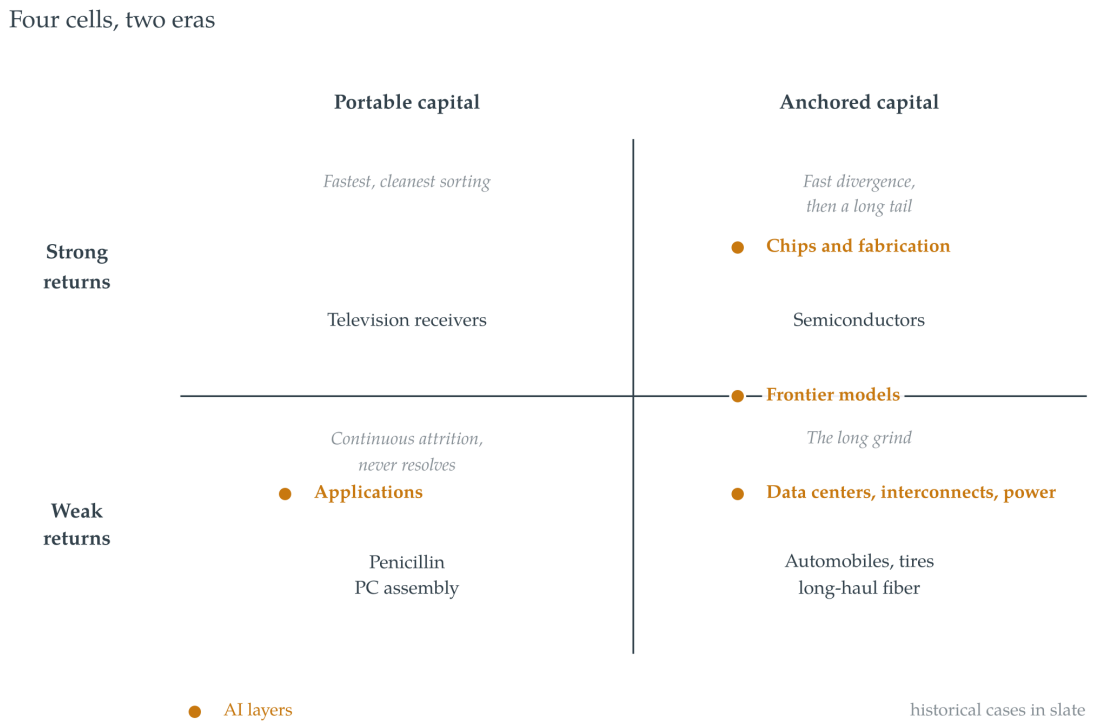

Figure 2. Four cells, two eras. Frontier models sit on the returns boundary: anchored capital, with returns real but weaker than an operating system's.

The ordering runs close to the reverse of the intuitive one. The layer generating the headlines has nearly finished sorting. The layer carrying the most capital has barely started.

Which is why the compression everyone has watched for thirty years does not help here. Recognition has accelerated; construction has not. A data center takes years to site, permit and build regardless of how fast the model inside it improved while the concrete cured, and a fabrication plant takes longer.

The signal can travel at whatever speed markets currently move at. The concrete cannot.

What to watch

Stop asking whether the AI shakeout has started. Ask which layer, and what happens to the capital of the firms that lose in it.

Watch capacity utilization rather than share prices. Falling prices while capital is still being committed is a buildout repricing. Idle capacity while capital withdraws is a sorting beginning.

Watch each layer's own exit rate rather than the sector's headlines, since the layers sort on separate schedules and a strategy calibrated to a single industry-wide verdict will be wrong about at least one of them.

And before committing capital, ask what the test reduces to: who buys this if we are wrong, and at what fraction of cost? An honest answer of a rival, at scrap, means the position will have to be defended for years rather than sold.

The objection

The framework assumes the layers stay separate. They may not, and the four-position reading above depends entirely on it.

Hyperscalers already design their own accelerators. Model developers already contract directly for power. If integration continues, the layers merge and so do their schedules, and a per-layer estimate stops meaning anything. The PC transition ran the other way, toward specialization, which is the basis for expecting the same here — but that is an assumption, not a finding, and it is the weakest joint in the argument.

A second objection has more force than it first appears. Adoption cycles have shortened for decades, and the AlexNet-to-ChatGPT gap is a third of the PC's catalyst-to-inflection span, so why expect a sorting on a twentieth-century schedule? The answer is in the section above: compression has applied to how fast information and capital recognize an opportunity, which is what shortened the gap. It has not applied to how fast physical capital gets built, tested, and either justified or abandoned.

The inheritance

Anchored capital has a cruel afterlife. Because it cannot leave quickly, the losses stay with whoever financed it, while the assets themselves persist and are eventually used by somebody else. That asymmetry is the whole reason the exit condition matters: the same property that traps the loser preserves the asset.

For the layer carrying most of AI's capital, that is the likeliest ending. Not a clean sorting, and not a write-off, but a long period in which shells, substations and interconnects change hands at a fraction of what they cost, run by firms that did not build them, on schedules set by an electrical connection rather than by a model release.

The capital stays.

Sources and notes

Prior work. The three-part shape of a technology transition is not original here. Its lineage runs through Gort and Klepper (1982), Klepper and Graddy (1990), Jovanovic and MacDonald (1994), Klepper (1996, 2002), and Abernathy and Utterback on dominant designs as formalized by Utterback and Suarez (1993). Carlota Perez, *Technological Revolutions and Financial Capital* (2002), gives the arc at civilizational scale, and commentary since late 2025 has applied it to the present moment. This essay does not dispute that reading; it attempts something narrower the framework does not provide.

The data. Steven Klepper and Kenneth Simons, "Industry Shakeouts and Technological Change" (2005), "Technological Extinctions of Industrial Firms" (1997), "The Making of an Oligopoly" on tires (2000), and "Dominance by Birthright" on television receivers. Automobile figures from Klepper on the US automobile industry: producers peaking at 272 in 1909, falling to 28 by 1929 and 9 by 1941, 725 entrants in total, roughly 6% of exits through acquisition. Tires: 533 entrants through 1930, producers peaking at 274 in 1922, 51 by 1940, 24 by 1970. Television receivers: 95 producers in 1950, 35 by 1960, 20 by 1970. Penicillin: roughly 70% fewer producers between the 1952 peak and the early 1980s, acquisitions accounting for 3% of post-peak exits. The divergence condition — a race to keep up in innovation, in which the firms already producing most and knowing most gain most from investing further — is Klepper and Simons's own finding across all four industries.

Portability, and what is claimed as new. A body of work in strategic management sits closer to this argument than the industrial-organization literature does. Marvin Lieberman, "Exit from Declining Industries: 'Shakeout' or 'Stakeout'?", *RAND Journal of Economics*, 1990, examines the sequence of divestment across thirty declining chemical products and finds small plants closing at higher rates

while large multiplant firms were likelier to close individual plants; it addresses which firms leave first rather than how long the leaving takes. Jaideep Anand and Harbir Singh, "Asset Redeployment, Acquisitions and Corporate Strategy in Declining Industries," *Strategic Management Journal*, 1997, is the closest published relative by subject. Lieberman, Lee and Folta, "Entry, exit, and the potential for resource redeployment," *SMJ*, 2017, treats redeployment inside the firm rather than resale outside it. Daniel Levinthal and Brian Wu, "Resource Redeployment and the Pursuit of the New Best Use: Economic Logic and Organizational Challenges," *Strategy Science*, 2024, describe resources that go dormant for want of any outside buyer, the defunct steel plant or the mine that no longer pays. That is what this essay calls anchored capital, named in that literature from the inside of the firm rather than the outside. In economics, Luís Cabral (2012) models capacity sunk costs as a driver of shakeouts; the claim here is that redeployability rather than quantity of capital sets the pace, descending from Williamson on asset specificity. Its closest relative is Shleifer and Vishny, "Liquidation Values and Debt Capacity," *Journal of Finance*, 1992; Kim and Kung (2017) build a measured redeployability index, used there to explain investment and debt capacity rather than shakeout duration. The contribution attempted here is threefold and modest: applying redeployability to duration, decomposing duration into a divergence phase and a portability-driven tail, and reading a stack layer by layer on both.

Measurement. Duration is counted in surviving firms, from peak producer count until net exits fall to near zero and stay there, following the industry life-cycle literature. Counted in capacity or prices, the same industry can look sorted years earlier.

The four AI layers. Boundaries are drawn for this argument rather than observed. Applications is a residual category rather than an industry: it has no countable producer population from inception, and keeps refilling underneath whatever is sorting.

Provenance. The catalyst-inflection-shakeout framework, the historical datings and the closing argument come from the author's client work beginning in 2000, the earliest surviving artifact a presentation deck from 2004. That deck contains three dating errors, corrected here. The convergence with the published literature above was discovered afterwards, not built on. The reading of the PC transition — chip and operating-system layers concentrating within roughly a decade of 1981 while hardware assembly took until the mid-2000s — is the author's, not a published finding, and its supporting market-capitalization figures rest on a contemporaneous reproduction of a Goldman Sachs note of August 1997 that has not been located. An earlier version of this argument used commuter and trunk airlines under deregulation as a natural experiment; the carrier counts do not support it, and the slot portfolio survives as illustration carrying no evidential weight.

AI's physical capital. Stijn Van Nieuwerburgh, "Data Centers: Financing the AI Buildout," Columbia Business School working paper, March 2026.

Current conditions, as of September 2026. Statements about compute and memory scarcity, capital commitment, and upstream strain in chip supply and power generation describe conditions at the time of writing, and are the parts of this argument most likely to date.

Disclosure of interests. The author was formerly employed by Ericsson and by IBM.

AI disclosure. The ideas and the writing are mine. I used an AI assistant to restructure an existing draft, including reframing its central claim against the prior literature, and for source verification and proofreading.