MEKETA

The AI Boom

What Could Come Next

WHITEPAPER | SEPTEMBER 2026

CONTRIBUTORS
Alison Adams, PhD
Frank Benham, CFA, CAIA

Artificial intelligence has moved from research curiosity to one of the topics dominating global capital markets in just a few years. Since ChatGPT was released in November 2022, AI has reordered equity market leadership, redirected hundreds of billions of dollars of annual capital expenditure toward data centers, chips, and power, and drawn public and private credit markets into financing a buildout whose returns remain unproven. While forecasts for AI-related investment may vary considerably depending on assumptions, it is expected that another $4 trillion to $8 trillion could be invested in AI-related infrastructure over the next five years.[1] Each year of the boom has produced a new narrative and a new analytical framework, from scaling laws to the DeepSeek moment to agentic demand. What none of the narratives or frameworks have settled is where the boom goes from here, and which indicators can help distinguish one path from another.

This paper examines several different economic frameworks to help investors answer that question. Much of the current commentary reduces the AI boom to a single ending, a leverage-driven bust. But a deeper look into how major technological innovations have evolved in the past could provide investors with a richer analytical framework to evaluate potential outcomes of the AI boom. We ask "What if the speculative boom/bust narrative is inaccurate?" "If so, what other outcomes could be possible?" We separate four end states instead: a financial break, a rotation of value away from the companies currently priced to capture it, a transition in which a financial break gives way to a longer phase of falling prices and broadening adoption, and a demand disappointment that requires no credit event.

Key Takeaways

- The AI boom has more than one way to end. We see multiple possibilities, including a financial break, a rotation of value away from the companies currently priced to capture it, a transition into a longer deployment phase, and a demand disappointment that requires no credit event.
- A financial break would not by itself settle the question of AI's economic significance. In past technology buildouts, the collapse of the claims on new infrastructure marked the boundary into a longer adoption and usage phase. The infrastructure was generally completed and eventually used, even where shareholders of the companies that built it did not always profit.
- A combination of economic frameworks can be used to consider different questions. These include whether the financing structure holds, which firms capture the gains, what follows a financial break, and whether demand grows as unit costs fall. Treating them as competing forecasts obscures that they are describing different parts of the same structure.
- The extent to which efficiency gains will drive greater AI use will bear not only on the financial prospects of current AI investments, but also on how AI impacts the broader economy. Whether falling unit costs expand total AI usage rather than simply reduce spending on AI will depend on how far usage extends past current applications, and the most aggressive capital expenditure plans rest on the proposition that AI demand has no obvious ceiling.
- Any single indicator can support opposite conclusions about events, so multiple indicators are worth monitoring in each context. Investors may be better served by identifying which path the evidence supports than by assuming the boom must end the way the most recent comparable episode did.

The AI Boom

A key theme of the last few years has been the stupendous magnitude of capital expenditures related to AI. Over the past three and a half years, Alphabet, Meta, Amazon, and Microsoft have spent approximately $1.1 trillion on AI-related investments.[2] Markets absorbed the SpaceX initial public offering in June 2026, which raised roughly $75 billion and ranks as the largest on record.[3] Public markets may next be asked to fund the two leading AI frontier laboratories. Anthropic and OpenAI each filed confidentially in June 2026, and valuations for both companies are estimated to be over $1 trillion.[4]

Over the past five years, as AI technology continued to advance rapidly, investors have grappled with a continuous stream of AI investment narratives and analytical frameworks. Investor expectations have shifted repeatedly over this period as the economics of AI deployment have evolved. Past technological and economic booms have all ended, but each had its own path. They did not always culminate in collapse or in ways that contemporaries recognized at the time. This backdrop naturally leads to the question of how the current boom will end, and specifically whether the spending and the financing structure supporting it will be justified and rewarded.

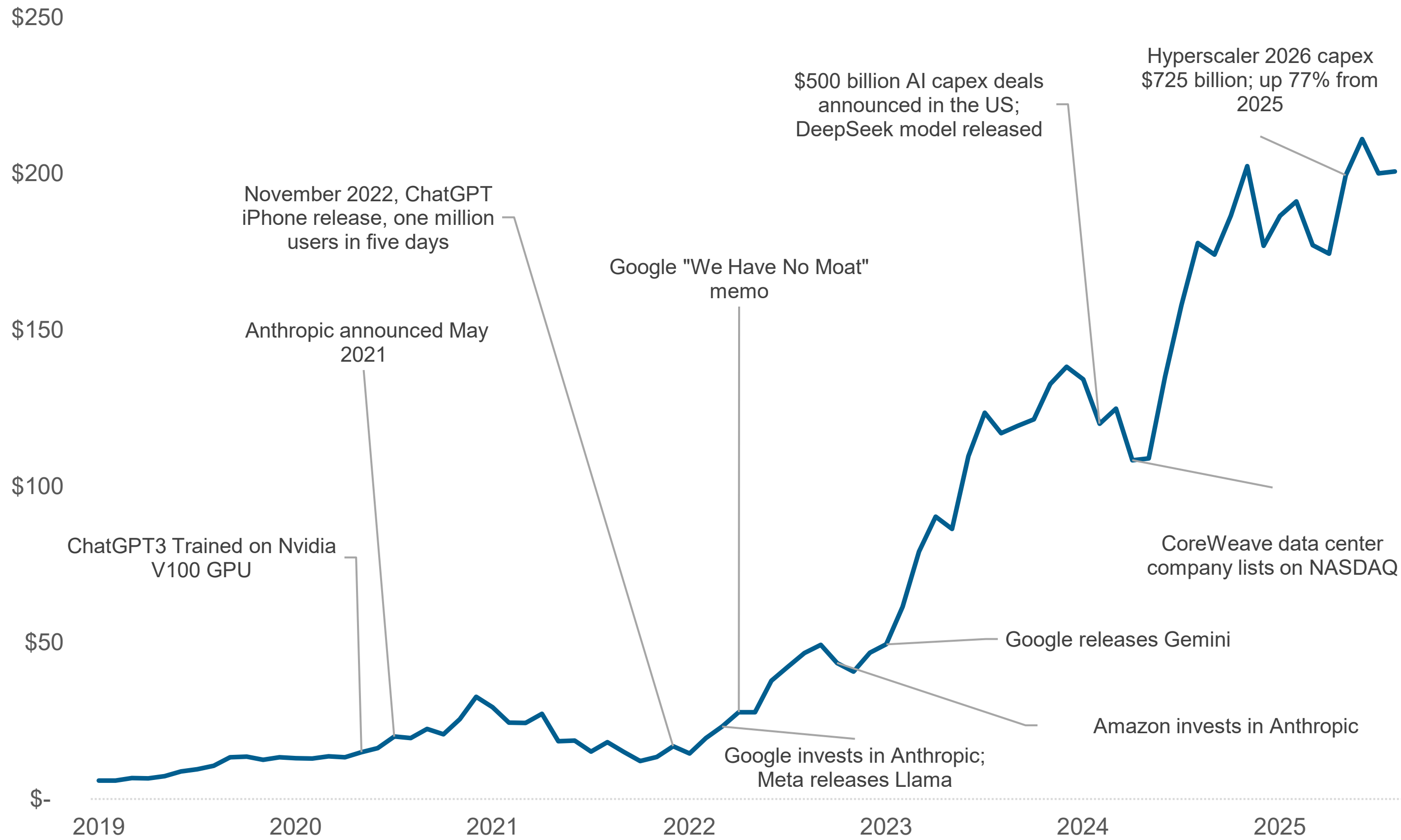
How Might the Boom End?

Investors may look to past financial bubbles and financial crises as a framework to guide how the AI boom might progress and end. The dynamics of investor enthusiasm, frothy valuations, speculation, and leverage are familiar variables and many institutional stress tests and models capture these financial characteristics. In particular, Hyman Minsky argued that during extended periods of stability, borrowers and lenders accept progressively thinner margins of safety, so that debt service comes to depend not on operating cash flow but on refinancing or asset sales. In Minsky's account, this reversal is generated internally and requires no external shock to trigger it, since the same stability that encourages thinner cushions eventually leaves too little room to absorb an ordinary disappointment in cash flow.

The dot-com era is a closer comparison than the earlier manias Minsky's readers usually cite, though it is often invoked imprecisely. The dot-com bubble was funded largely with equity rather than debt, and its collapse destroyed substantial market value without producing a banking or lending crisis. However, the telecommunications infrastructure buildout that ran alongside it is the closer analog to the more recent buildout of AI infrastructure. Telecom providers raised debt against network capacity constructed before demand for it existed while equipment makers extended credit to the customers buying their equipment. When telecom traffic growth failed to meet projections, the refinancing market closed. Global Crossing and WorldCom both filed for bankruptcy in 2002.[5] The current buildout shares several of these features, including debt raised against assets built ahead of demand, special purpose financing structures, circular arrangements between suppliers and customers, and broad retail investor participation.

The analogy has limits. A railroad or a telecom carrier built infrastructure that competitors then used on similar terms. AI changes the economics of the businesses that adopt it, which means the technology bears on what a portfolio already owns as much as on which AI-related securities it holds. A framework built only around the debt on a builder's balance sheet will not capture that.

Figure 1
AI Timeline and Nvidia's Share Price



Source: Bloomberg, data as of July 2026. The term hyperscaler refers to the largest operators of cloud computing infrastructure, principally Amazon Web Services, Microsoft Azure, Google Cloud, and Meta's internal infrastructure, which build and run data centers at a scale far above that of conventional enterprise IT.

Four end states are worth separating, since each carries different implications for how a portfolio is positioned. The first is a financial break, in which the financing structure supporting the buildout fails before the assets it funded generate sufficient returns. The second is a rotation, in which the capital expenditure proves justified in aggregate but the value migrates away from some established companies that investors currently own.[6] The third is a transition in which a break in the financing structure proves to be the boundary into a longer phase of falling costs, widespread adoption, and increasing significance for the broader economy. The fourth is a demand disappointment, in which usage fails to grow into the capacity being built, without any accompanying credit event.

These outcomes are not mutually exclusive, and they may occur one after another. For example, a demand disappointment could trigger a financial break, and a financial break could clear the way for a transition. The economic frameworks discussed below each speak to one of these paths, and each is better understood as answering a different question than competing for the same answer.

Minsky and Kindleberger address the financing structure and whether it can withstand its own leverage. Schumpeter, and later Aghion and Howitt, address which firms capture the gains and which are displaced. Perez addresses sequence, distinguishing the installation of a new technology from its deployment. Jevons addresses demand, and specifically whether falling unit costs expand usage faster than they compress revenue. Taken together, they describe a wider range of outcomes than the single bust narrative that dominates much current commentary.

Minsky and Kindleberger: Does the Financing Structure Break?

The AI capital cycle is shifting toward credit and special purpose vehicles (SPVs).[7] Against the roughly $1.1 trillion in AI-related capital expenditure by the four largest hyperscalers cited above, some now estimate that the total data center buildout may reach $3 trillion (see Figure 2), and investors are monitoring rising debt levels more closely.[8] The classic debt bubble worry is that leverage rises so rapidly and to such an extent that cashflows are not able to service the interest on the debt. A "Minsky moment" arrives when borrowers can no longer service debt out of operating cash flow and are forced to sell assets to meet their obligations.[9]

With the likes of Alphabet and Meta burning through free cash flows as they ramp up their capital expenditures, investors are now applying a discount to company valuations as debt levels rise. At the same time, private credit firms such as Apollo, Blackstone, BlackRock, Blue Owl, and PIMCO are deploying investment structures to finance the construction of new data centers and power generation capabilities. Morgan Stanley estimates that as much as $800 billion in private financing could be invested in the next two years.[10] Anthropic and OpenAI have both taken investment from Nvidia, Alphabet, and Microsoft. Those same laboratories buy chips from Nvidia and lease compute from Microsoft's Azure and Amazon's AWS, which places their largest investors on the other side of their largest purchases.[11] OpenAI has announced plans to spend $1 trillion on compute, much of it with the same firms that have invested in it.[12]

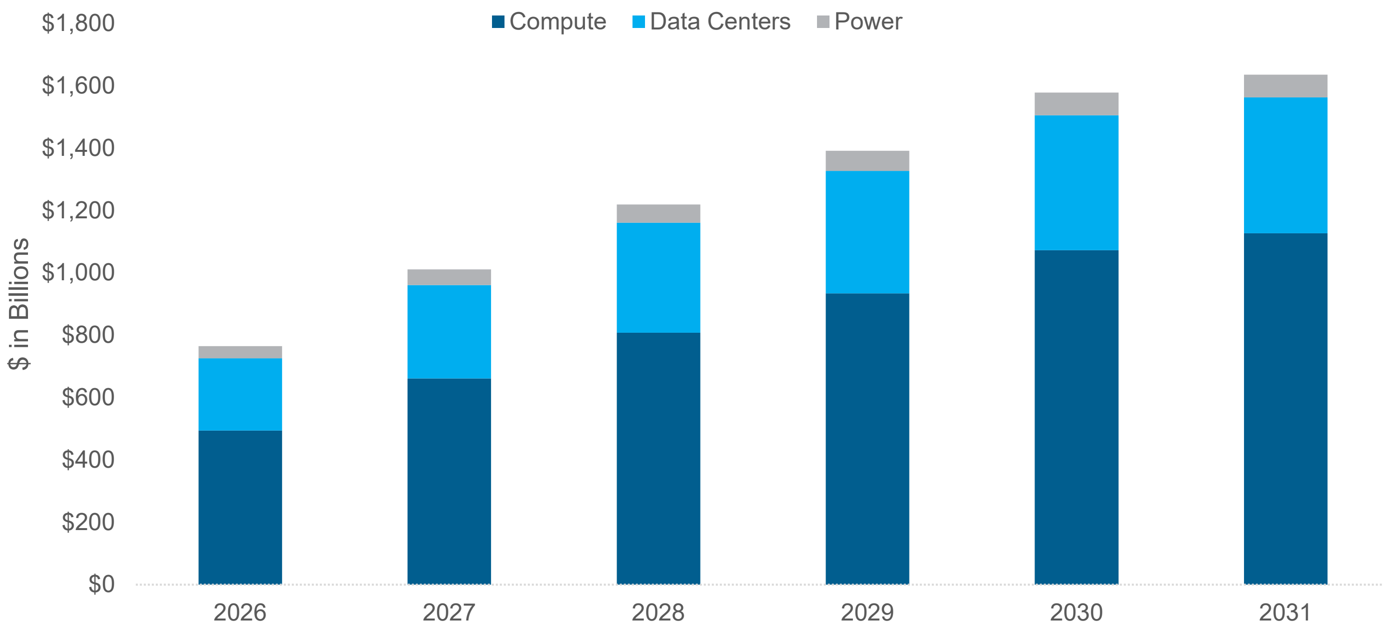
Minsky classified borrowers into three financing postures, and Kindleberger set that classification inside a historical framework for financial bubbles. At first, operating cash flow covers both interest and principal, so the borrower is not dependent on credit markets to meet obligations. The next stage is speculative finance where cashflows cover the interest payments but not the principal, so the borrower must roll over maturing debt and access to new credit depends on continued earnings. This may give way to even more speculative (i.e., "Ponzi") financing where cash is insufficient to service either interest or principal, and the borrower must issue new debt or sell assets simply to service existing obligations. When new rounds of debt issuance fail and creditors seek the return of their capital, the bubble bursts. Kindleberger described a similar sequence at the level of the market rather than the firm: displacement, boom, euphoria, distress as insiders begin to sell, and revulsion.[13] Minsky's Ponzi stage maps onto the transition from euphoria to distress, when refinancing, not earnings, is sustaining positions.[14]

Applying that classification to the current buildout is difficult. The hyperscalers have so far largely financed capital expenditure from operating cash flow and investment grade bond issuance, which places them in Minsky's first category.[15] If obligations are measured to include special purpose vehicles and the multi-year compute commitments that sit outside those balance sheets, the picture is harder to read.

Circular financing in AI companies has become a concern in some quarters of the market. The specific concern is that capital supplied by one participant returns to that same participant as revenue. For example, Nvidia invests equity in CoreWeave, which uses those proceeds, alongside debt collateralized by GPUs, to purchase Nvidia chips.[16] Nvidia recognizes that purchase as revenue, which supports the market value and operating cash flow that funds the next investment. And Nvidia has separately contracted to buy any capacity CoreWeave cannot sell, leaving it supplier, investor, and buyer of last resort all at once (see Figure 3). Hence the same dollar can be recognized as revenue at more than one point in that circuit. This implies that reported revenues may overstate actual demand.

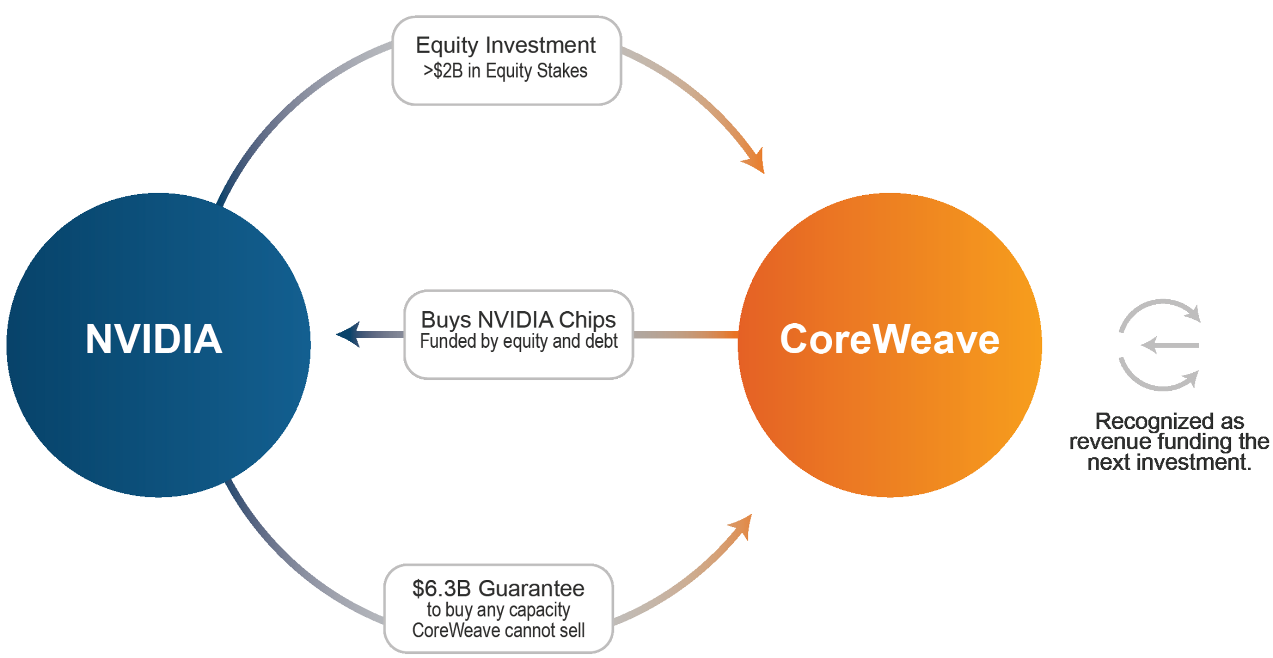
The same pattern recurs across the ecosystem over longer circuits. For example, Alphabet has invested billions in Anthropic, and Anthropic in turn commits much of that capital to purchasing Google cloud capacity, returning a portion of the investment to Alphabet as revenue. Nvidia has increasingly used its balance sheet to finance participants in the AI ecosystem, including capital that ultimately funds purchasers of its own hardware.[17] This can be viewed as analogous to vendor financing schemes used by companies like Cisco during the dot-com era, and therefore as a possible warning signal of the increasing enmeshment of AI companies.[18] These arrangements sit closest to Minsky's Ponzi category, since revenue booked by one participant originates as an investment or commitment made by another participant in the same chain. Competitive pressure, chip scarcity, and cheap credit make each individual deal rational for the counterparties involved even as the aggregate structure grows more fragile.

Figure 2
Projected Global AI Infrastructure Investment



Sources: Goldman Sachs, "Tracking Trillions: The Assumptions Shaping the Scale of the AI Build Out," April 2026. Financial Times, R. McMorrow et al., "Big Tech AI Spending Spree Tops $1 Tn," July 30, 2026. Figures shown are third party projections and do not represent actual results. Actual amounts may differ materially, and estimates of AI-related capital expenditure vary considerably across sources.

Figure 3
Illustration of Intermixed Nature of Nvidia and CoreWeave



Source: The Economist, "Nvidia Is The Central Bank of AI," September 3, 2026. CoreWeave is one of several specialized providers, sometimes called neoclouds, that own and rent GPU capacity without operating a broader cloud services business.

Obligations that sit outside the reported balance sheet complicate the picture further, and they take more than one form. Some are borrowings by special purpose vehicles established to fund specific data center projects, which do not consolidate onto the sponsor's balance sheet. Some are leases that have been signed but have not yet commenced, which are not recognized until the underlying asset is made available for use. Others are multi-year commitments for cloud and compute capacity that are structured as service agreements and therefore fall outside lease accounting altogether. Moody's estimated that the five largest hyperscalers held approximately $969 billion in total undiscounted future lease commitments at year-end 2025, of which roughly $662 billion related to leases not yet commenced.[19] Aggregate estimates that combine these categories have reached approximately $1.65 trillion, though such figures vary considerably with what each source chooses to include.[20] Separately, until OpenAI and Anthropic go public, there is no public balance sheet against which to measure their commitments.

Private credit lenders financing SPV data center projects are lending in part against the investment grade ratings of the hyperscaler counterparties, and that security depends on those counterparties' ability to service their debt as both reported and unreported obligations rise.[21] Minsky's argument is that this is the ordinary sequence, not an aberration, and that the cushion thins during good times precisely because good times have persisted.

Schumpeter: Who Wins and Who Loses?

A second framework sets aside the financing question and asks who captures the gains. Joseph Schumpeter argued that economic growth is driven by creative destruction, the process by which new technologies and new forms of organization displace established ones rather than compete with them on price.[22] Such technologies have an impact that not only is creative in terms of new businesses but also destructive for legacy companies that cannot adapt. Schumpeter argued that the competition that matters in capitalism is not price competition but competition from new commodities and technologies, new sources of supply, and new forms of organization.[23] AI is a new technology that not only challenges legacy companies in terms of price and organization structures but also may have the ability to undermine the foundations of business models.

Nvidia is usually offered as an example, though its rise illustrates only the enabling side of this process. Nvidia was founded in 1993, but global demand for GPUs did not take off until the onset of the AI boom, which propelled the company into the list of ten largest companies listed in the S&P 500 Index in 2023.[24] Nvidia's market capitalization reflects investors' expectations that the firm will capture substantial gains from the AI boom. Destruction of the old economic structure is a separate question, and its evidence lies elsewhere.

Chegg is the clearest AI-related example to date of the destruction side of Schumpeter's model. Its business rested on a decade of investment in a proprietary library of textbook and homework solutions, sold to students for roughly $20 a month. Management and investors treated that accumulated library as the company's "moat" (i.e., a durable advantage). Five months after ChatGPT was released, Chegg's management told analysts that the free chatbot was suppressing new customer growth, making Chegg the first listed company to attribute a revenue shortfall to generative AI. Paying subscribers fell from a peak of about 7.8 million to roughly 3.2 million by early 2025, revenue declined about 30% year over year, and the company cut close to half its workforce. No credit event, financing failure, or leverage was involved. A decade of accumulated proprietary data simply ceased to be a defensible asset once a general-purpose model could reproduce its function at no marginal cost to the user.

The creative destruction in this episode was the work of an ecosystem. Nvidia supplied the hardware, but the displacement of incumbents came from the model developers and platform owners, including OpenAI, Anthropic, Alphabet, Microsoft, and Amazon, whose products change the economics that legacy companies face. For many established companies, the move to invest in AI capabilities is far beyond a question of return on investment; it represents an existential imperative.

While Schumpeter and Minsky both put credit creation at the center of their respective theories, one treats it as the engine of renewal and the other as the source of fragility. Schumpeter's account places credit at the center of the renewal process, since it is lending that hands new entrants the purchasing power to displace incumbents before those incumbents have any reason to expect it.[25] He described the result as a process that destroys the old economic structure from within.

Software-as-a-service (SaaS) companies have experienced periods of sharp multiple compression since 2024, driven by investor concern that AI tools will substitute for the workflows their products automate.[26] Listed professional services and information companies have seen similar pressure, as investors have questioned whether business models built on billable expertise or proprietary content libraries will retain their pricing power. AI has lowered barriers to entry in several markets, which increases the ability of new entrants to displace established companies. A leaked internal Google memo made this argument from inside one of the incumbents, contending that neither Alphabet nor OpenAI held a durable barrier against fast-improving open-weight models.[27]

Index leadership has been unusually persistent since the Global Financial Crisis. Turnover among the largest S&P 500 constituents has been low by the standards of earlier decades, and the largest companies have retained their positions for longer than in prior periods of comparable concentration. Schumpeter's work reminds us that this persistence is not a permanent feature, as innovation and entrepreneurship have repeatedly challenged and even destroyed legacy companies.[28] Evidence of the rise and fall of even the largest companies can be seen in the history of the stock market over the past half-century (see Figure 4).

This may help explain the impetus to invest heavily in AI-related capital expenditures by many companies. For example, Alphabet, once the newcomer, is now a legacy company that views itself as being in an existential race to reinvent and restructure its business model by building data centers, power generators, and developing custom-built TPUs.[29] The relevant point for investors is that Alphabet is spending because it believes that its current business model will not survive otherwise. This is a Schumpeterian judgment, and it explains why the spending continues through periods when a purely cyclical investor would expect it to pause.

Figure 4

Top Ten Companies by Market Capitalization by Decade

1970	1980	1990	2000	2010	2020
IBM	IBM	IBM	GE	ExxonMobil	Apple
AT&T	AT&T	Exxon	ExxonMobil	Microsoft	Microsoft
General Motors	Exxon	GE	Pfizer	Apple	Amazon
Exxon	Amoco	Philip Morris	Citigroup	Walmart	Alphabet
GE	Schlumberger	Procter & Gamble	Cisco	Berkshire Hathaway	Facebook (Meta)
Mobil	Shell Oil	Royal Dutch Petroleum	Walmart	GE	Berkshire Hathaway
Texaco	Mobil	Bristol-Myers Squibb	Microsoft	Procter & Gamble	Visa
DuPont	Chevron	Merck	AIG	Johnson & Johnson	Johnson & Johnson
Procter & Gamble	Atlantic Richfield	Walmart	Merck	Pfizer	Tesla
Eastman Kodak	GE	AT&T	Intel	Coca-Cola	Walmart

Source: S&P Dow Jones Indices, FactSet. Annual year-end top ten companies by market capitalization.

Framed that way, the spending can be individually rational and collectively excessive at the same time. Each firm is choosing sensibly given the alternative of losing its position and potentially forfeiting its business entirely. When many competitors reason the same way, the prudent individual choice aggregates into a buildout larger than any plausible path for demand might justify.

Philippe Aghion and Peter Howitt formalized Schumpeter's insight into a mathematical growth model, and their work supplies the mechanism that Schumpeter left implicit.[30] In their model, the pace of innovation-driven growth and the share of the resulting rents captured by incumbents depend on the incentives facing potential innovators. Those incentives are shaped by the competitive structure of the market as much as by the size and frequency of the innovations themselves. Later work by Aghion with several co-authors qualified the relationship, finding an inverted-U pattern: competition tends to spur innovation among firms already close to the technological frontier, where the prospect of escaping rivals is realistic, and to discourage it among firms further behind.[31] Applied to the AI boom, that qualification implies the outcome is not settled by the technology's capabilities. It depends on whether the firms currently investing are close enough to the frontier that competition drives them forward, or far enough behind that it discourages them from investing at all.

Aghion and Howitt have also shown that long-run economic growth stems from purposeful research and development by rational agents seeking profit.[32] That mechanism bears on how the established hyperscalers are positioned, since each is defending an existing market and margin structure while competing against firms whose cost bases and business models differ from their own.

For investors holding broad public equity exposure, the question is whether the companies currently priced as AI's beneficiaries can defend their claim on it. That claim depends on whether barriers to entry hold. The current capital expenditure is an attempt to build barriers that scale, capital intensity, and possibly regulation might sustain. Whether they hold is an open question, and it is a different question from whether AI proves economically significant.

Perez: Does a Break Lead to Adoption?

Carlota Perez's historical analysis identifies five technological revolutions since 1771, beginning with the Industrial Revolution and running through the age of information and telecommunications.[33] She identified a recurring sequence: an installation period in which new infrastructure is built and financial capital drives speculative investment, a turning point, and a deployment period in which the technology diffuses across the wider economy. Her claim about the turning point is the relevant one for this paper. In each of her cases it involved a financial collapse, and that collapse marked the boundary between installation and deployment rather than the end of the technology's economic significance.

Perez's analysis does not contradict Minsky so much as continue from him. Minsky describes the mechanism by which an installation period ends, while Perez analyzes what follows it. Perez described a period of change where social norms, regulations, and financing may need to be reconfigured to allow for broader adoption of the technology. For example, existing rules (for safety, labor, and intellectual property rights) typically must be rewritten to incorporate the new technology. This period of change may not be a single identifiable event, and it may unfold over several years.[34] The subsequent deployment period is characterized by falling prices, broader adoption, and benefits that accrue increasingly to users of a technology. Investment does not stop, but it slows toward the rate of general economic growth, so a slowdown in capital expenditure may signal deepening deployment.

The telecommunications buildout described earlier is the clearest recent case, and it looks different from this vantage point. While many carriers did not survive, the fiber they laid stayed in the ground. Bandwidth costs fell far enough to support businesses that could not have existed at the prices the carriers needed to earn, and much of the consumer internet of the following decade was built on capacity that its original owners could not monetize. What failed was the claims on the infrastructure, not the infrastructure itself.[35]

This distinction matters more than the timing of any bust. A collapse in the valuations of AI infrastructure providers would be consistent with the technology proving transformative regardless of what it means for the shareholders of those firms. Which claims survive the turning point is the question that matters for positioning. In Perez's historical cases, the physical infrastructure was generally completed, and later operators eventually used it, while many of the equity claims on the firms that built it were not honored. Some companies failed outright, others survived heavy drawdowns and recovered, and the returns available to a diversified holder depended greatly on the holding period and the price paid at entry.[36]

Jevons: Does Demand Hold as Costs Fall?

A fourth framework addresses demand directly. Writing about Britain's coal supply in 1865, William Stanley Jevons observed that as steam engines became more efficient, demand for coal rose rather than fell. Cheaper power widened the range of uses to which it could be put. Investors could consider the cost of AI tokens, the unit in which model output is priced and sold, through a similar lens.[37] The same mechanism, of efficiency gains driving costs lower and increasing demand, has run through frontier AI models. DeepSeek's R1 release in January 2025 tested the market's confidence in that relationship, temporarily erasing close to $600 billion of Nvidia's market value in a single session on the assumption that cheaper training would translate to weaker chip demand.[38] That premise was subsequently questioned, since it turned out that DeepSeek had trained on Nvidia chips and reported training costs understated the full expense. The more useful observation is what happened afterward. Chip demand did not fall, and both spending on training and consumption of inference capacity continued to rise.[39] The episode is therefore consistent with the Jevons mechanism operating so far, though it does not establish that the mechanism will continue to hold.

As AI models have improved, demand has risen, and for now the economics of serving that demand appear to be on an increasingly favorable track. Anthropic's annualized revenue run rate is projected to reach $120 billion by the end of 2026,[40] and margins on its inference infrastructure, the difference between what a token sells for and what the compute to generate it costs, have risen from 38% to over 70%.[41] Demand is currently strong enough that scarcity is showing up in prices. Amazon recently raised by 20% the price of its EC2 capacity blocks, which let developers reserve GPU capacity in advance.[42]

The cost of producing a token, meanwhile, continues to fall, and it is falling through several independent channels at once. Successive chip generations deliver substantially more output per unit of energy. Custom chips designed for specific workloads can be more efficient than general purpose hardware. Open-weight and smaller models offer capability at a fraction of the cost of frontier systems. Each of these would compress the price of AI output on its own, and demand has so far grown faster than price has fallen.[43]

The comparison has a limit worth stating. Coal was an input to nearly everything, so efficiency gains in steam power met a demand curve with considerable depth behind it.[44] Whether AI tokens behave the same way depends on how far demand extends past current uses. If the applications now driving consumption represent most of what the technology will be used for, falling costs may compress revenue faster than volume replaces it, and the capacity currently being built could exceed what the market is willing to absorb. That is the demand disappointment case, and it does not require a credit event to produce poor outcomes for the companies that funded the buildout.

The strongest argument against it is that AI demand may not be bounded by human users at all. AI agents can generate demand for other agents, and machine-initiated usage may have no obvious ceiling other than the cost of the tokens it consumes.[45] That proposition is plausible and, as yet, untested. It is also the single assumption on which the most aggressive capital expenditure plans depend. Utilization rates against committed capacity are therefore a more informative measure than token volumes alone.

Utilization rates against committed capacity are therefore a more informative measure than token volumes alone.

What This Means for Investors

This paper does not forecast which of the four end states will occur. The frameworks examined answer different questions, and together they describe a wider range of outcomes than a single bust narrative. What can be done is to identify which evidence would distinguish one path from another, and to watch that evidence closely.

Each path leaves a different signature. A financial break would appear first in the financing structure: capital expenditure rising further above operating cash flow at the hyperscalers, growth in commitments held off the balance sheet, widening spreads on debt issued through special purpose vehicles, and renegotiation of the long-term contracts on which neocloud financing depends. A rotation would appear in dispersion, with the gap widening between the companies building AI infrastructure and the companies deploying it, while benchmark concentration leaves broad equity exposure more sensitive to that gap than it appears. A transition would appear as falling prices alongside rising volumes, capital expenditure decelerating toward the rate of general economic growth, and margin compression among current leaders occurring while adoption continues to broaden. A demand disappointment would appear in utilization, specifically in the gap between committed capacity and capacity actually in use, and in whether margins hold as competition increases.

Several of these indicators are shared, and the same movement can support opposite conclusions. A slowdown in capital expenditure is the clearest example. Read alone, it suggests the buildout has outrun the demand for it. Read alongside falling prices and rising usage, it suggests deployment is deepening and spending is normalizing. Read alongside deteriorating credit metrics, it points to financing constraints. No single series settles the question, which is why these measures warrant watching together.

For institutional portfolios, concentration is the exposure that every one of these paths runs through. The companies currently priced as the primary beneficiaries of AI occupy an unusually large share of public equity benchmarks. Because of that scale, an event that reshuffles the industry would register in broad equity allocations whichever path produced it. Rapid technological change has rarely left early leaders intact. Some companies will likely underperform and a few may not survive as independent businesses, while the eventual winners may capture gains that more than offset those losses over a long horizon. That observation does not resolve into a single recommendation, since the appropriate response to a rotation differs from the appropriate response to a credit event. It does suggest that concentration deserves examination as a risk distinct from any view on whether AI proves transformative, and that these two questions, which much commentary runs together, are separable.

The frameworks reviewed here are best understood as complements rather than as competing predictions. Minsky and Kindleberger describe how a period of heavy, credit financed investment becomes fragile. Schumpeter, with Aghion and Howitt supplying the mechanism, describes which firms capture the resulting gains and which are displaced. Perez describes what tends to follow the break and suggests that a collapse in the claims on new infrastructure has historically been compatible with the infrastructure itself proving valuable. Jevons describes the demand condition on which the entire structure depends. Investors watching for a single Minsky moment may find that the more consequential developments unfold across several years and are visible only in aggregate.

End Notes

1 Source: Goldman Sachs, "Tracking Trillions: The Assumptions Shaping the Scale of the AI Build Out," April 2026. AI-related spending estimates vary considerably depending on time frame, companies included, sector(s), and US only versus global estimates. Most estimated spending reports have raised their total estimates since 2025.

2 Source: Financial Times, R. McMorrow et al., "Big Tech AI Spending Spree Tops $1 Tn," July 30, 2026. Including Alphabet, Microsoft, Amazon, and Meta. Market estimates will vary depending on timing and what companies are being considered.

3 Source: Reuters and Fortune reporting, June 12, 2026. SpaceX raised approximately $75 billion at a listing valuation of roughly $1.77 trillion, the largest initial public offering on record. SpaceX merged with xAI in February 2026, and its IPO filing described the combined company as an "AI services and infrastructure company."

4 Sources: Financial Times, G. Hammond, "Anthropic investors bet on $2tn valuation in record IPO," August 12, 2026; Wall Street Journal, K. Clark, "OpenAI Considers Pre-IPO Funding Round at More Than $1.2 Trillion Valuation," September 16, 2026. As of this writing, official prospectuses for Anthropic and OpenAI are not yet available. Listing timing for both likewise remains unsettled.

5 Source: Yale Journal on Regulation, J. Sidak, "The Failure of Good Intentions: The WorldCom Fraud and the Collapse of American Telecommunications After Deregulation," October 2002. WorldCom overstated the growth of internet traffic to support continued capital raising for network expansion. That episode returns later in this paper in a different role, since the same buildout that bankrupted the carriers produced infrastructure that was eventually used. Whether an episode looks like a bust or like the start of something depends heavily on where in the sequence it is observed.

6 It is too early to say which established companies will gain or lose market share in the AI boom.

7 A special purpose vehicle (SPV) is a separate legal entity created to hold a specific asset or fund a specific project.

[8] Source: Bloomberg, "The $3 Trillion AI Data Center Build-Out Becomes All-Consuming For Debt Markets," August 2026.

[9] A Minsky moment names a point in a credit cycle where a sudden sharp collapse in asset prices occurs. The collapse in asset prices marks the end of a credit cycle. Source: H. Minsky, "Stabilizing an Unstable Economy," 1986. Paul McCulley coined the term "Minsky moment" in 1998 to describe the Russian financial crisis.

[10] Source: Bloomberg, C. Sam et al., "A Guide to the Circular Deals Underpinning the AI Boom," July 28, 2026.

[11] Source: Financial Times, "Wall Street Giants Partner with Nvidia on $500 Bn AI Financing Deal," August 2026. "Compute" refers to the raw computing power, rented by the hour from data center operators, that is needed to train and run AI models.

[12] Source: Wall Street Journal, "OpenAI Unveils Plans for Seemingly Limitless Expansion of Computing Power," September 23, 2025.

[13] Source: C. Kindleberger, "Manias, Panics, and Crashes: A History of Financial Crises," 1978.

[14] The two frameworks differ on where the reversal originates. Kindleberger's sequence begins with displacement, an external event that changes profit expectations. Minsky's does not require one, since in his account the fragility accumulates through ordinary behavior during good times and the turn can come from a disappointment that would have been absorbed easily at an earlier point in the cycle. The distinction has practical consequences. If a shock is required, the analytical task is to identify what it might be. If it is not, the task is to measure how much cushion remains. Of note, Minsky referred to the first stage as hedge finance, the second stage as speculative finance, and the third stage as Ponzi finance.

[15] This picture may be gradually changing. Oracle, Meta, Alphabet, and Amazon all issued public debt between September 2025 and March 2026. After over a decade of share buybacks and returning cash to investors, in 2026 Alphabet issued new shares to help finance its capital investment program in AI.

[16] A graphics processing unit (GPU) is the specialized processor used to train and run AI models. GPU-collateralized debt is borrowing secured against the chips themselves, in the way that equipment finance is secured against machinery. The value of that collateral depends on how quickly successive chip generations render earlier ones obsolete.

[17] Sources: The Economist, "Nvidia Is the Central Bank of AI," September 3, 2026. Financial Times, B. Elder, "Nvidia's $200bn 'Balance Sheet-As-Service," August 25, 2026.

[18] At the height of the dot-com era, Cisco was best understood as the primary infrastructure supplier for the buildout of the internet. While it is common for a supplier to extend credit to a customer to secure sales, Cisco's vendor financing to companies without revenues contributed to the steep decline in the Cisco balance sheet as its customers defaulted.

[19] Source: Moody's Ratings, "US hyperscalers' off-balance-sheet lease commitments approach $1 trillion," David Gonzales and Alastair Drake, February 2026.

[20] Sources: Nikkei Asia, K. Yamada, "Five US Tech Giants; Hidden Debts Soar to $1.65 tn on Opaque AI Funding," July 21, 2026; Tom's Hardware, "AI Tech Companies Have 'Hidden Debt' Worth Around $1.65 Trillion," July 22, 2026.

[21] Because the sponsor typically does not consolidate the SPV onto its own balance sheet, borrowing at the SPV level does not appear as debt of the sponsor, though the sponsor may retain economic exposure through contractual commitments.

[22] Source: J. Schumpeter, "Capitalism, Socialism and Democracy," 1942.

[23] Source: J. Schumpeter, "Capitalism, Socialism and Democracy," 1942.

[24] Source: Bloomberg; data pulled August 2026. Nvidia was founded in 1993 by Jensen Huang, Chris Malachowsky and Curtis Priem.

[25] Source: J. Schumpeter, "The Theory of Economic Development" (1911, English translation 1934).

[26] Sources: Forrester, "SaaS As We Know It Is Dead: How to Survive the SaaS-pocalypse," February 11, 2026; "The 'SaaS-pocalypse' Has A Remarkable Number of Survivors," September 2, 2026. In early 2026, SaaS companies lost approximately $1 trillion in market capitalization. However, many stocks have since partially recovered.

[27] Source: SemiAnalysis, "We Have No Moat, And Neither Does OpenAI," May 4, 2023. An open-weight model is one whose trained parameters are published, allowing anyone to run, modify, or build on it without access to the developer's infrastructure. This differs from open-source software, since the training data and code are usually not released. The memo argued that open-weight models were improving faster than expected and were closing the gap on proprietary systems, leaving frontier developers with little durable advantage.

[28] Source: J. Schumpeter, "Capitalism, Socialism and Democracy," 1942.

[29] A tensor processing unit (TPU) is a chip Google designed specifically for machine learning workloads, as an alternative to buying general-purpose GPUs from Nvidia.

[30] Source: Econometrica, P. Aghion and P. Howitt, "A Model of Growth Through Creative Destruction," 1992.

[31] Source: Aghion, Bloom, Blundell, Griffith, and Howitt, "Competition and Innovation: An Inverted-U Relationship," Quarterly Journal of Economics, 2005.

[32] Source: Econometrica, P. Aghion and P. Howitt, "A Model of Growth Through Creative Destruction," 1992.

[33] Source: Carlota Perez, "Technological Revolutions and Financial Capital: The Dynamics of Bubbles and Golden Ages," 2002.

[34] Source: Carlota Perez, "Technological Revolutions and Financial Capital: The Dynamics of Bubbles and Golden Ages," 2002.

[35] Source: Yale Journal on Regulation, J. Sidak, "The Failure of Good Intentions: The WorldCom Fraud and the Collapse of American Telecommunications After Deregulation," October 2002.

[36] Consider, for example, the experience of two fiber companies from the dot-com era: Global Crossing and Corning. Global Crossing's equity went to zero. Corning shares fell roughly 99% from its 2000 peak, remained a going concern, and later recovered.

[37] A token is the basic unit in which a model receives input and produces output, roughly corresponding to a fragment of a word, and it is the unit in which AI capability is priced and sold.

[38] Source: Bloomberg, "Nvidia's $589 Billion DeepSeek Rout Is Largest in Market History," January 27, 2025.

[39] Inference is the computation performed each time a trained model responds to a request, as distinct from training, which is the one-time computation that produces the model. While training costs are incurred up front, inference costs recur with every use.

[40] Source: Financial Times, G. Hammond, "Anthropic Tells Investors It Will Be Profitable for a Second Straight Quarter," September 13, 2026. An annualized revenue run rate projects a full year of revenue from a shorter period, typically a single month multiplied by twelve. It is not audited revenue and it does not account for seasonality or contract churn.

[41] Source: SemiAnalysis, D. Nishball et al., "AI Value Capture - The Shift to Model Labs," April 30, 2026. Note that these margins represent an estimate from SemiAnalysis, and other third-party estimates may differ.

[42] Source: Business Insider, E. Kim, "Amazon Just Made a Key AI Cloud Service More Expensive," June 26, 2026.

[43] Source: The Wall Street Journal, J. Mitovich, et al, "The Chip Rally Is at $5.7 Trillion and Counting. How Much Further Can it Go?" May 27, 2026.

[44] As locomotives became more powerful and efficient, the cost of transporting coal also fell. This helped lower energy costs even outside the transportation sector.

[45] Source: Financial Times, "The Rise of Physical AI: Can Robots Save US Manufacturing?" August 30, 2026.

Important Information

This report (the "report") has been prepared for the sole benefit of the intended recipient (the "recipient").

Significant events may occur (or have occurred) after the date of this report, and it is not our function or responsibility to update this report. The information contained herein, including any opinions or recommendations, represents our good faith views as of the date of this report and is subject to change at any time. All investments involve risk, and there can be no guarantee that the strategies, tactics, and methods discussed here will be successful.

The information used to prepare this report may have been obtained from investment managers, custodians, and other external sources. Some of this report may have been produced with the assistance of artificial intelligence ("AI") technology. While we have exercised reasonable care in preparing this report, we cannot guarantee the accuracy, adequacy, validity, reliability, availability, or completeness of any information contained herein, whether obtained externally or produced by the AI.

The recipient should be aware that this report may include AI-generated content that may not have considered all risk factors. The recipient is advised to consult with their Meketa advisor or another professional advisor before making any financial decisions or taking any action based on the content of this report. We believe the information to be factual and up-to-date but do not assume any responsibility for errors or omissions in the content produced. Under no circumstances shall we be liable for any special, direct, indirect, consequential, or incidental damages or any damages whatsoever, whether in an action of contract, negligence, or other tort, arising out of or in connection with the use of this content. It is important for the recipient to critically evaluate the information provided.

Certain information contained in this report may constitute "forward-looking statements," which can be identified by the use of terminology such as "may," "will," "should," "expect," "aim," "anticipate," "target," "project," "estimate," "intend," "continue," or "believe," or the negatives thereof or other variations thereon or comparable terminology. Any forward-looking statements, forecasts, projections, valuations, or results in this report are based upon current assumptions. Changes to any assumptions may have a material impact on forward-looking statements, forecasts, projections, valuations, or results. Actual results may therefore be materially different from any forecasts, projections, valuations, or results in this report.

Performance data contained herein represent past performance. Past performance is no guarantee of future results.

