Can AI Companies Justify Their Valuations?

A Framework for Revenue, Margins, and Compute Costs

Joel Gonzalez Gomez

Independent working paper

October 2025

Abstract

The 2025 artificial intelligence investment cycle produced a central question for investors, founders, and corporate strategists: can AI companies justify their valuations, or are public and private markets capitalizing an economic transformation before the cash flows have appeared? This paper develops a valuation framework for AI-related companies based on three primitives: revenue growth, margin structure, and compute intensity. The framework distinguishes between three layers of the AI economy: infrastructure providers that sell compute, hyperscalers that convert capital expenditure into AI capacity, and application/software companies that attempt to convert AI capability into recurring revenue or workflow automation. Using a point-in-time dataset constructed only from information available by 31 October 2025, I compare approximate sales multiples, recent growth signals, capital intensity, and margin profiles. Market capitalizations are anchored to 30 September 2025 and revenue denominators use the latest annual fiscal period reported by each company by the cutoff date. I then derive a simple discounted cash flow model showing the revenue growth and free-cash-flow margins required to justify high revenue multiples. The results are intentionally not framed as a stock recommendation. Instead, the paper argues that AI valuations can be justified only when at least one of three conditions holds: sustained supernormal revenue growth, durable software-like margins despite inference costs, or infrastructure-like capital deployment that creates defensible compute scarcity. The key implication is that AI valuation should be analyzed less as a generic technology multiple and more as a payback problem: how quickly can a firm turn AI-related capital, compute, and product capability into high-quality cash flow?

Keywords: artificial intelligence, valuation, compute costs, gross margins, capital expenditure, software economics, public markets, AI infrastructure.

JEL-style classification: G12, G32, L86, O33.

Contents

1 Introduction

Artificial intelligence became one of the dominant valuation narratives of 2025. The market did not merely reprice a group of technology firms. It repriced an economic belief: that AI would alter the production function of software, enterprise services, advertising, cloud infrastructure, and knowledge work. NVIDIA represented the hardware and accelerator layer. Microsoft, Alphabet, Amazon, and Meta represented the hyperscale infrastructure and platform layer. Palantir and ServiceNow represented enterprise software companies trying to convert AI into workflow productivity, operating leverage, and customer expansion.

The investment question is simple but difficult: can these companies justify their valuations? A conventional answer might compare price-to-earnings or revenue multiples. That approach is insufficient because the AI cycle is not a uniform software cycle. It combines a software-like demand story with an infrastructure-like cost structure. Traditional SaaS scaled with high gross margins because the marginal cost of serving one more customer was low. Many AI applications, by contrast, carry real variable costs in the form of model inference, compute, human review, and data infrastructure. At the same time, AI infrastructure companies and hyperscalers require capital expenditures at a scale closer to industrial capacity buildouts than traditional software distribution.

This paper develops a framework for analyzing whether AI valuations can be justified. The framework rests on three questions. First, how fast can revenue grow? Second, what steady-state margin can the company sustain after compute, depreciation, and support costs? Third, how much capital must be deployed before the revenue and margin story becomes visible? These questions matter because the same valuation multiple can mean different things depending on where a company sits in the AI stack. A semiconductor company with 75% gross margins and 100% revenue growth is not economically equivalent to a hyperscaler spending more than 20% of revenue on capital expenditure, nor to an AI application startup growing rapidly with 25% gross margins.

The empirical part of the paper builds a small point-in-time dataset of AI-linked public companies using only financial statements, earnings releases, guidance, and market-value references available by 31 October 2025. This design matters: no fourth-quarter or full-year 2025 results released in 2026 are used. The data are not intended to produce a definitive market call. They discipline the framework. The core pattern is clear: the highest multiples are found where investors expect either extreme revenue growth, very high operating leverage, or strategic control of scarce AI infrastructure. But high multiples become fragile when compute costs prevent gross margins from converging toward software-like levels.

The paper contributes in three ways. First, it separates the AI valuation debate into layers: infrastructure, hyperscale platforms, and applications. Second, it derives a simple valuation identity connecting revenue multiples to growth, discount rates, and free-cash-flow margins. Third, it introduces a compute-adjusted margin framework showing why gross margin quality is a central issue for AI-native companies.

2 AI Valuation as a Layered Problem

A central mistake in the AI valuation debate is treating all AI exposure as the same. AI is not a single business model. It is a stack. The economic exposure of an accelerator manufacturer is different from that of a cloud platform, which is different from that of an enterprise AI workflow company.

2.1 Infrastructure providers

Infrastructure providers sell the hardware, networking, software systems, or cloud capacity required to train and deploy AI models. Their valuation depends on demand for compute, capacity constraints, pricing power, and customer concentration. NVIDIA is the clearest public-market example. Its fiscal 2025 annual report, available well before this paper's cutoff, recorded revenue of \$130.5 billion and gross margin of 75.0% (NVIDIA, 2025a). The latest quarter available by 31 October was the second quarter of fiscal 2026: revenue reached \$46.7 billion, up 56% year-over-year, Data Center revenue reached \$41.1 billion, and GAAP gross margin was 72.4% (NVIDIA, 2025b).

The infrastructure layer can justify high multiples if three conditions hold: capacity remains scarce, customers continue expanding AI workloads, and margins remain high despite supply-chain scaling. However, this layer also carries cyclicality risk. If customers overbuild capacity or if accelerator supply normalizes, revenue growth and pricing power may decelerate.

2.2 Hyperscalers and platform companies

Hyperscalers occupy the second layer. Microsoft, Alphabet, Amazon, and Meta convert capital expenditure into AI infrastructure and then attempt to monetize it through cloud workloads, advertising optimization, productivity software, consumer products, or model-driven services. Their valuation question is not just whether AI revenue grows. It is whether AI-related capital expenditure produces adequate returns on invested capital.

Information available by the cutoff shows both strong monetization signals and an unusually capital-intensive buildout. Microsoft reported fiscal 2025 revenue of \$281.7 billion, up 15%, while Azure surpassed $75 billion and grew 34% (Microsoft, 2025). Alphabet's third-quarter revenue rose 16% to $102.3 billion, Google Cloud grew 34%, and management raised expected 2025 capital expenditure to $91–93 billion (Alphabet, 2025). Amazon's third-quarter net sales rose 13% to $180.2 billion and AWS grew 20%; however, trailing-twelve-month free cash flow fell to $14.8 billion as property-and-equipment purchases increased sharply (Amazon, 2025). Meta's third-quarter revenue rose 26% to $51.2 billion, while management expected 2025 capital expenditure of $70–72 billion (Meta Platforms, 2025).

For hyperscalers, the AI valuation test is an infrastructure payback problem. Investors must ask whether incremental AI capital expenditure creates incremental revenue, margin expansion, or strategic protection large enough to compensate for the cash deployed.

2.3 AI application and workflow software

The third layer consists of application and workflow software companies. Palantir and ServiceNow provide public examples available before the cutoff. Palantir's second-quarter 2025 revenue exceeded $1.0 billion and grew 48% year-over-year; the company raised its full-year 2025 revenue guidance to approximately $4.14–4.15 billion (Palantir, 2025). ServiceNow reported third-quarter total revenue of $3.407 billion, up 22%, and subscription revenue of $3.299 billion, up 21.5% (ServiceNow, 2025).

The application layer faces the most important margin question. AI software may grow faster than SaaS, but it may also have higher variable costs. Bessemer's 2025 AI benchmarks distinguish between "AI Supernovas," which can reach roughly $40 million of ARR in year one and $125 million in year two but average around 25% gross margins, and "AI Shooting Stars," which grow rapidly with approximately 60% gross margins (Bessemer Venture Partners, 2025). This suggests that AI application valuation cannot rely on growth alone. It must also examine whether gross margins can converge toward software-like levels.

3 Data and Descriptive Evidence

3.1 Data construction

The dataset combines public financial statements, earnings releases, company guidance, and public market-capitalization references for a selected group of AI-linked companies (Yahoo Finance, 2025). The information cutoff is 31 October 2025. Approximate market capitalizations are anchored to 30 September 2025, before the late-October earnings releases, while revenue denominators use the latest annual fiscal period reported by each company by the cutoff. This point-in-time design avoids importing later full-year results into an October 2025 paper. It also means the revenue periods differ across companies, so the tables provide an illustrative valuation snapshot rather than an exact apples-to-apples ranking.

Table 1: AI-linked public companies: point-in-time valuation snapshot

Company	Layer	Latest annual revenue ($B)	Sep. 30 market cap ($B)	Sales multiple
NVIDIA	Infrastructure / semiconductors	130.5	4,542	34.8x
Microsoft	Hyperscaler / AI platform	281.7	3,850	13.7x
Alphabet	Hyperscaler / AI platform	350.0	2,975	8.5x
Amazon	Hyperscaler / cloud + retail	638.0	2,480	3.9x
Meta	Consumer AI platform / ads	164.5	1,845	11.2x
Palantir	Enterprise AI software	2.865	434.8	151.7x
ServiceNow	AI workflow software	10.984	191.6	17.4x

Table 2: Operating signals available by the October 2025 cutoff

Company	Latest disclosed growth signal	Infrastructure intensity
NVIDIA	Q2 FY26 revenue +56%	3.1% actual
Microsoft	FY25 revenue +15%; Azure +34%	22.9% actual
Alphabet	Q3 revenue +16%; Google Cloud +34%	26.3% full-year guide
Amazon	Q3 net sales +13%; AWS +20%	18.2% TTM estimate
Meta	Q3 revenue +26%	43.2% full-year guide
Palantir	Q2 revenue +48%	Not reported on a comparable basis
ServiceNow	Q3 total revenue +22%	Not reported on a comparable basis

Market capitalizations are approximate public reference values and are highly date-sensitive. "Infrastructure intensity" uses the most relevant actual, trailing-twelve-month estimate, or full-year guidance available by the cutoff and is therefore not strictly comparable across companies. For Amazon, the estimate is inferred from trailing-twelve-month operating cash flow and free cash flow disclosed in the third-quarter release. The key valuation variable is:

$$M_i = \frac{V_i}{R_i}, \tag{1}$$

where V_i is the equity market value of firm i, and R_i is the latest reported annual revenue. Sales multiples are imperfect because they ignore margins, growth, cash, debt, and business mix. Their advantage is that they make visible how differently the market capitalizes AI exposure across layers.

3.2 Sales multiples and capex intensity

Figure 1 shows approximate revenue multiples in the point-in-time dataset. The dispersion is large. Palantir screens as an extreme multiple company because investors are capitalizing unusually rapid expected growth and operating leverage against a relatively small annual revenue base. NVIDIA also trades at a high revenue multiple, but with a much larger revenue base and already-visible infrastructure economics. Hyperscalers trade at lower sales multiples because their revenue bases are enormous and include mature business lines.

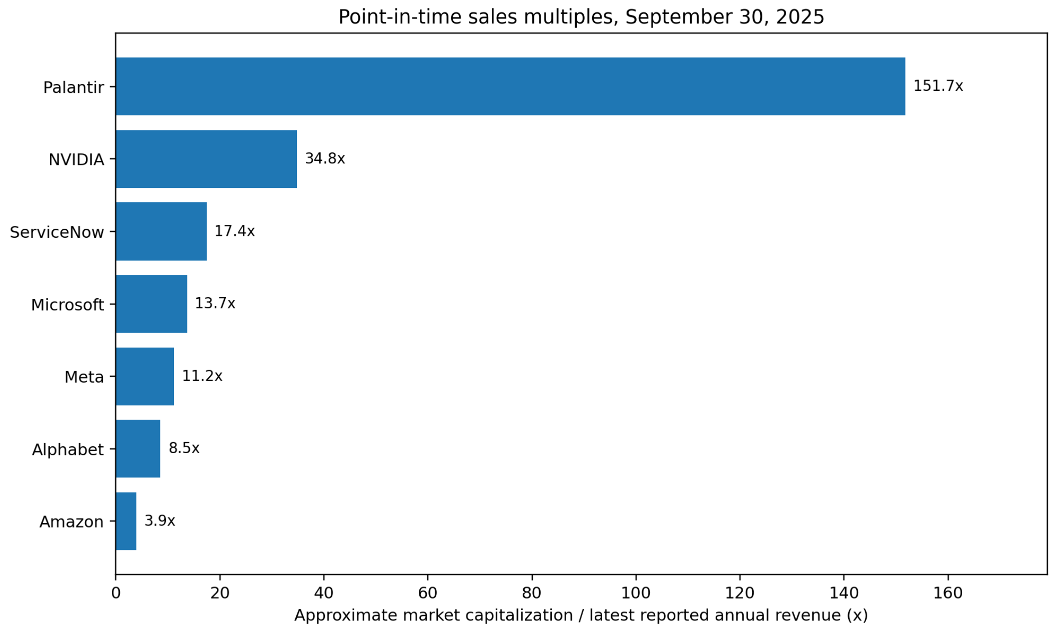

Figure 1: Approximate market capitalization to latest reported annual revenue, anchored to 30 September 2025.

Figure 2 shows selected infrastructure-investment intensity measures available by the cutoff. The periods and definitions differ, so the figure is diagnostic rather than a strict ranking. It nevertheless captures what makes the AI investment cycle economically distinct from traditional software: the leading platform companies are not merely adding features; they are building industrial-scale compute capacity.

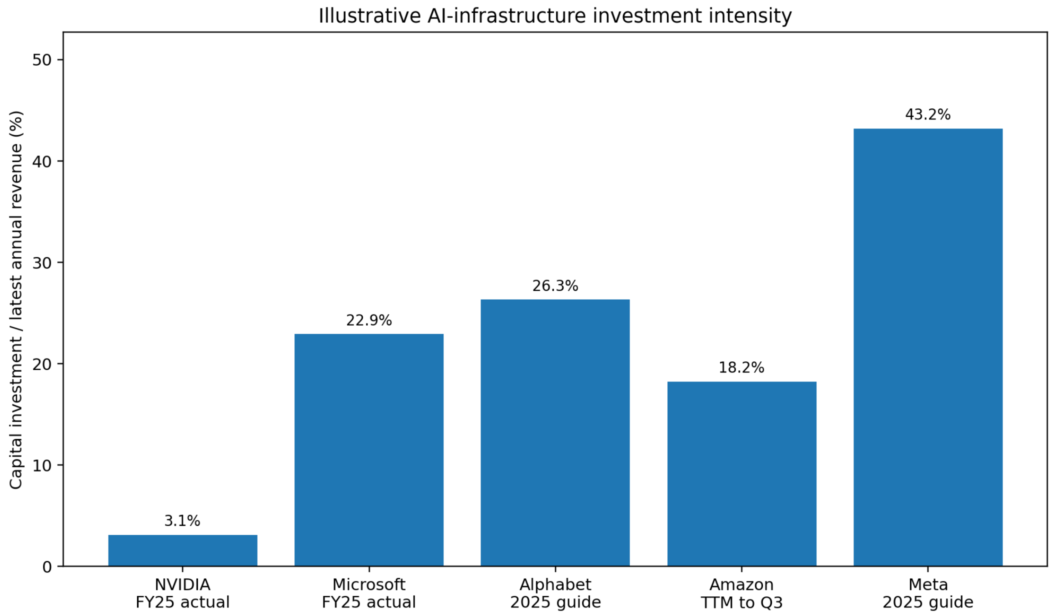

Figure 2: Illustrative infrastructure investment intensity using actual, trailing-twelve-month, or guided values available by 31 October 2025.

3.3 Benchmarking AI software against SaaS

Table 3: Stylized business-model benchmarks

Category	Growth	Gross margin	Capex intensity	Notes
Traditional mature SaaS benchmark	20%	80%	5%	Stylized mature software profile
Bessemer AI Supernovas	200%	25%		Average early AI supernova gross margins around 25%; ARR/FTE around $1.13M
Bessemer AI Shooting Stars	300%	60%		Average early AI shooting stars gross margins around 60%; faster than SaaS growth
Hyperscaler AI infrastructure buildout	15%		20%	Stylized public cloud buildout: capex intensity often more than 20% for AI-heavy infrastructure

The benchmark table highlights a central tension. AI software can grow faster than SaaS, but early AI gross margins can be far below mature SaaS levels. The difference between a

25% gross-margin AI application and an 80% gross-margin SaaS business is not cosmetic. It changes the valuation mathematics.

4 A Valuation Framework

4.1 Revenue multiple identity

Let R_0 denote current revenue, g_t expected revenue growth, m_t free-cash-flow margin, r the discount rate, and T the explicit forecast horizon. Revenue evolves as:

$$R_t = R_{t-1}(1 + g_t). \tag{2}$$

Free cash flow is:

$$FCF_t = m_t R_t. \tag{3}$$

The enterprise value is:

$$V_0 = \sum_{t=1}^{T} \frac{m_t R_t}{(1+r)^t} + \frac{TV_T}{(1+r)^T}, \tag{4}$$

where terminal value is:

$$TV_T = \frac{m_T R_T (1 + g_{term})}{r - g_{term}}. \tag{5}$$

Dividing by current revenue gives the implied sales multiple:

$$\frac{V_0}{R_0} = \sum_{t=1}^{T} \frac{m_t \prod_{k=1}^{t}(1 + g_k)}{(1+r)^t} + \frac{m_T \prod_{k=1}^{T}(1 + g_k)(1 + g_{term})}{(r - g_{term})(1+r)^T}. \tag{6}$$

Equation 6 shows that high sales multiples require high growth, high eventual margins, low discount rates, or high terminal expectations. If compute costs keep free-cash-flow margins low, then a high revenue multiple requires even more aggressive growth.

4.2 Compute-adjusted margin

For AI companies, gross margin should be decomposed. Let p be revenue per task or unit of usage, c_c compute cost per unit, c_h human-review or support cost per unit, and c_o other delivery cost. Then gross margin can be written as:

$$GM = 1 - \frac{c_c + c_h + c_o}{p}. \tag{7}$$

For SaaS, the marginal delivery cost is often low relative to subscription revenue. For AI, inference cost and human review may be material. This creates a margin hurdle:

$$GM_{AI} \to GM_{SaaS} \quad \text{only if} \quad \frac{c_c + c_h + c_o}{p} \downarrow . \tag{8}$$

This condition is one of the central valuation tests for AI-native software. If compute costs fall, model efficiency improves, and human review declines, AI gross margins can improve. If usage growth raises compute cost proportionally with revenue, valuation multiples should be lower than SaaS multiples, all else equal.

4.3 Capex payback condition

For hyperscalers, the relevant question is not only gross margin but payback on capital expenditure. Let I_0 denote incremental AI capital expenditure and ΔFCF_t the incremental free cash flow generated by AI services. A capital program is value-creating if:

$$\sum_{t=1}^{T} \frac{\Delta FCF_t}{(1+r)^t} - I_0 > 0. \tag{9}$$

Equivalently, the AI buildout is justified when the present value of incremental AI cash flows exceeds the present value of the compute investment. This is the core economic question for hyperscalers.

5 Numerical Simulation

5.1 Growth-margin tradeoff

The first simulation asks what revenue multiple is implied by different growth and free-cash-flow margin assumptions. I assume a 10-year forecast horizon, a 10% discount rate, and a 3% terminal growth rate. The output is the sales multiple implied by equation 6.

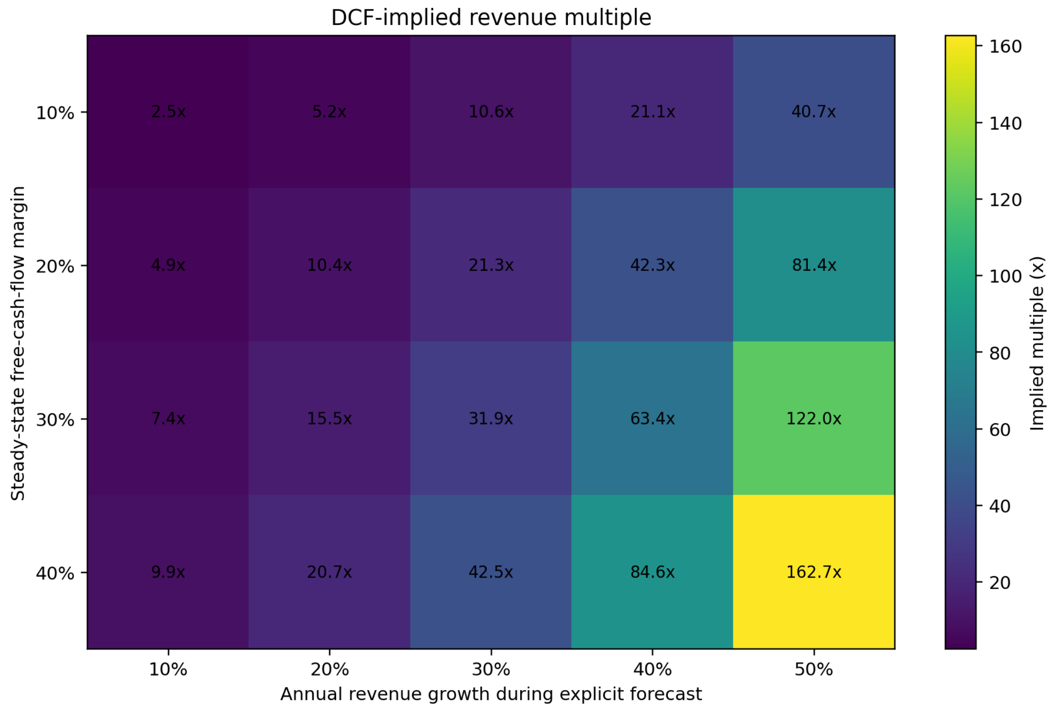

Figure 3: DCF-implied revenue multiples under different growth and margin assumptions.

The simulation illustrates why high AI multiples are demanding. At low free-cash-flow margins, even strong growth may not justify very high revenue multiples. At higher free-cash-flow margins, high growth becomes much more valuable. This is why margin quality is so important in AI valuation.

5.2 Gross margin sensitivity to compute and human review

The second simulation decomposes gross margin into compute cost and human-review cost. The result shows that AI application businesses can move from low-margin to high-margin profiles if compute cost falls and human review intensity declines.

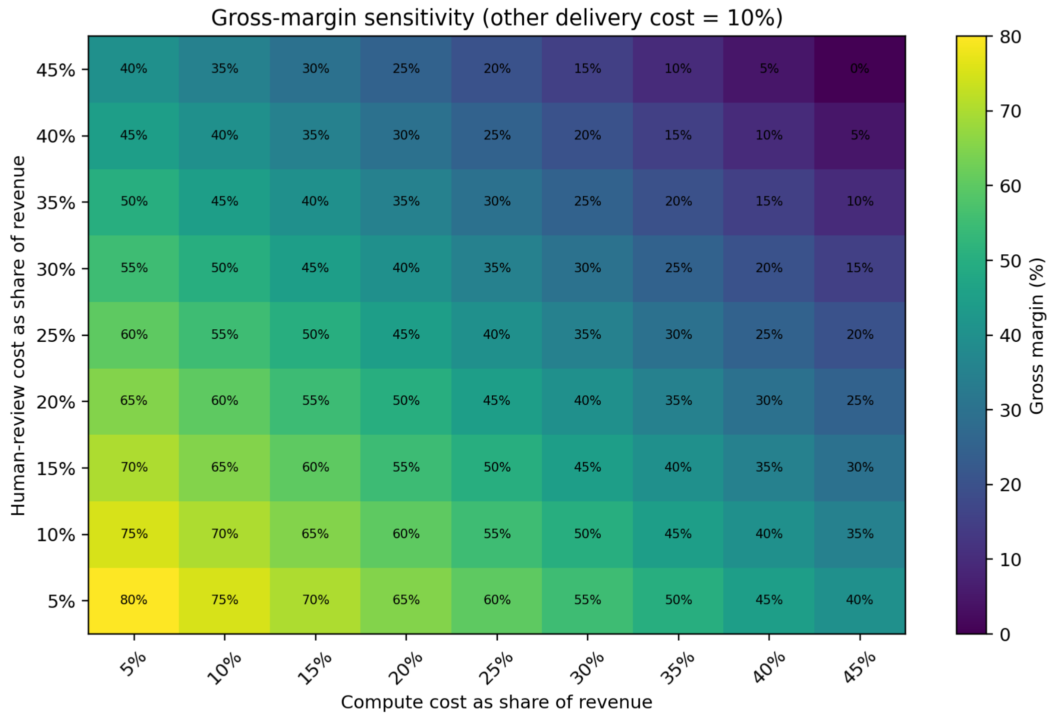

Figure 4: Gross margin sensitivity to compute cost and human-review cost.

The figure formalizes a key business-model point. AI-native companies may initially resemble services businesses because human review and compute absorb much of revenue. They become more software-like only when reliability improves, human intervention declines, and inference cost falls relative to price.

6 Can AI Companies Justify Their Valuations?

The answer is conditional. Some AI valuations can be justified, but not by the generic claim that "AI is big." They require a specific path from capability to revenue and from revenue to cash flow.

6.1 Case 1: Infrastructure scarcity

NVIDIA's latest disclosed profile by the October cutoff shows the clearest near-term fundamental support. Second-quarter fiscal 2026 revenue grew 56

6.2 Case 2: Hyperscaler payback

The hyperscaler case is more complex. Microsoft, Alphabet, Amazon, and Meta have deep distribution, large cash flows, and strategic reasons to invest in AI infrastructure. But their AI economics are embedded within broader businesses. For these firms, AI valuations are justified if capital expenditure produces measurable revenue acceleration, margin protection, or strategic control. Otherwise, the market may be capitalizing capex before payback is observable.

6.3 Case 3: AI application software

Application-layer valuations require the strongest margin discipline. Palantir's latest reported quarter before the cutoff combined 48

7 Discussion: A Three-Pillar Test

I propose a three-pillar test for AI valuation.

7.1 Pillar 1: Revenue durability

AI revenue should be evaluated based on recurrence, retention, expansion, and workflow depth. Revenue linked to experimental usage or novelty should receive a lower quality score than revenue embedded in mission-critical workflows.

7.2 Pillar 2: Margin convergence

Investors should ask whether gross margins converge toward software levels or remain structurally burdened by compute and human review. The relevant question is:

$$\lim_{Q\to\infty} GM_{AI}(Q) \geq GM^*, \tag{10}$$

where Q is usage volume and GM^* is the margin threshold required to justify the valuation multiple. If scale does not improve margin, the company is less software-like than its valuation may imply.

7.3 Pillar 3: Compute capital efficiency

For infrastructure-heavy companies, the key metric is AI revenue per dollar of capital deployed:

$$ACE = \frac{\Delta R_{AI}}{I_{AI}}, \tag{11}$$

where ACE is AI capital efficiency, ΔR_{AI} is incremental AI revenue, and I_{AI} is incremental AI investment. The market can tolerate high capex if ACE is high or improving. It becomes fragile when capex rises faster than monetization.

8 Limitations

This paper has several limitations. First, the dataset is selective and illustrative rather than exhaustive. Second, AI revenue is not consistently disclosed across public companies. Third, market capitalization is volatile and sensitive to date selection; the snapshot uses approximate values anchored to 30 September 2025. Fourth, the latest annual revenue periods differ across companies, and the infrastructure-intensity measures combine actual, trailing-twelve-month, and guided values. Fifth, capex is often multi-purpose: cloud infrastructure can support AI, search, advertising, consumer products, and internal workloads simultaneously. Sixth, private AI company valuations are less transparent than public-market valuations.

These limitations do not eliminate the usefulness of the framework. They reinforce the central point: AI valuation requires separating revenue growth, margin structure, and capital intensity rather than relying on narrative alone. The October 2025 information cutoff should also be understood as part of the research design: later fourth-quarter and full-year results could materially change the snapshot but are intentionally excluded.

9 Conclusion

AI companies can justify high valuations, but only under demanding economic conditions. The framework developed in this paper shows that valuation depends on the interaction between revenue growth, margin quality, and compute intensity. Infrastructure providers can justify high valuations when compute scarcity produces visible revenue growth and high gross margins. Hyperscalers can justify AI investment when capital expenditure creates durable incremental cash flow. Application-layer AI companies can justify premium multiples when they combine rapid growth with margins that converge toward software economics.

The central conclusion is that AI valuation is a payback problem. The market is capitalizing a future in which AI changes software, cloud, advertising, and enterprise workflows.

Whether that capitalization is justified depends on how quickly firms convert AI capability into high-quality revenue and cash flow. The most attractive AI companies will not merely show AI exposure. They will show evidence of durable demand, improving gross margins, and efficient conversion of compute into monetizable output.

A Variable definitions

- V_i: market capitalization or enterprise value of firm i.
- R_i: annual revenue of firm i.
- $M_i = V_i/R_i$: revenue or sales multiple.
- GM: gross margin.
- FCF_t: free cash flow in period t.
- m_t: free-cash-flow margin in period t.
- I_{AI}: AI-related capital expenditure or infrastructure investment.
- $ACE = \Delta R_{AI}/I_{AI}$: AI capital efficiency.

B Data notes

All company-level figures are gathered from public financial reports, earnings releases, company guidance, or public market-capitalization references available by 31 October 2025. Revenue, capex, and margin values are expressed in U.S. dollars. Approximate market capitalizations are anchored to 30 September 2025. Revenue denominators use each company's latest reported annual fiscal period by the cutoff date. No fourth-quarter or full-year 2025 results released in 2026 are included. The dataset is designed for strategic and valuation analysis, not a securities recommendation.

References

Bessemer Venture Partners. 2025. *The State of AI 2025*. Available at https://www.bvp.com/atlas/the-state-of-ai-2025.

NVIDIA Corporation. 2025a. *Form 10-K for fiscal year ended January 26, 2025*. Available at https://www.sec.gov/Archives/edgar/data/1045810/000104581025000023/nvda-20250126.htm.

NVIDIA Corporation. 2025b. *Second Quarter Fiscal 2026 Financial Results*, August 27. Available at https://nvidianews.nvidia.com/.

Microsoft Corporation. 2025. *Annual Report 2025*. Available at https://www.microsoft.com/investor/reports/ar25/index.html.

Alphabet Inc. 2025. *Third Quarter 2025 Results*, October 29. Available at https://abc.xyz/investor/.

Amazon.com, Inc. 2025. *Third Quarter 2025 Results*, October 30. Available at https://ir.aboutamazon.com/.

Meta Platforms, Inc. 2025. *Third Quarter 2025 Results*, October 29. Available at https://investor.atmeta.com/.

Palantir Technologies Inc. 2025. *Second Quarter 2025 Results*, August 4. Available at https://investors.palantir.com/.

ServiceNow. 2025. *Third Quarter 2025 Financial Results*, October 29. Available at https://investor.servicenow.com/.

Yahoo Finance. 2025. *Historical valuation measures and market-capitalization references*, values accessed for 30 September 2025.

Curcio, Eliseo. 2025. *Introducing LCOAI: A Standardized Economic Metric for Evaluating AI Deployment Costs*. arXiv:2509.02596.

Fang, Xinmin, Lingfeng Tao, and Zhengxiong Li. 2025. *Anchoring AI Capabilities in Market Valuations: The Capability Realization Rate Model and Valuation Misalignment Risk*. arXiv:2505.10590.

Maslej, Nestor et al. 2025. *Artificial Intelligence Index Report 2025*. Stanford Institute for Human-Centered AI.