The AI Credibility Triangle

A Framework for Interpreting AI Markets

Gemini, ChatGPT, Claude, W. H. L.

Champaign Magazine / Independent Research

March 17 2026

Abstract

Artificial intelligence has become one of the most powerful narrative drivers in global financial markets. Firms perceived to possess meaningful artificial intelligence capabilities often command significant valuation premiums, while corporate communication increasingly emphasizes AI integration. This environment creates informational asymmetry between genuine technological capability and narrative amplification.

This paper introduces the AI Credibility Triangle, a conceptual framework linking three interacting forces shaping AI market perception: the AI Premium, AI-Washing, and AI Signal. A key theoretical claim of the framework is that AI Signal primarily stabilizes the AI Premium rather than necessarily creating it. While narratives may initially generate speculative valuation premiums, sustained valuation requires verifiable technological capability.

To operationalize these dynamics, the paper proposes a Signal-to-Wash Index (R_sw) measuring the ratio of verifiable AI capability to narrative intensity, and an AI Credibility Matrix classifying firms according to observable signal strength and narrative integration. The model is interpreted through information theory and technological innovation cycles, particularly the work of Carlota Perez. Illustrative scoring examples using NVIDIA, Microsoft, C3.ai, and Box Inc. demonstrate the framework's classificatory power using traceable indicators.

The resulting model provides investors, researchers, and policymakers with a structured and operational framework for distinguishing durable technological value from narrative-driven mispricing.

1. Introduction

Artificial intelligence has emerged as the defining technological narrative of the mid-2020s. Advances in large-scale foundation models and accelerated compute infrastructure have generated unprecedented investor enthusiasm. Financial markets have responded by assigning significant valuation premiums to firms perceived as leaders in artificial intelligence development.

However, this enthusiasm has also created a parallel phenomenon: AI-Washing, the strategic amplification of AI narratives by firms whose underlying capabilities may be limited. This environment creates a signaling problem analogous to the informational asymmetries described in the classic job-market signaling model of Spence (1973), in which low-quality senders mimic the signals of high-quality counterparts to capture unearned premiums.

The result is a market in which three forces interact:

- AI Premium – valuation premiums associated with expected AI capabilities
- AI-Washing – narrative amplification surrounding AI claims
- AI Signal – verifiable evidence of technological capability

These forces form what this paper describes as the AI Credibility Triangle.

2. The AI Premium

2.1 Valuation Decomposition

The total market value of a firm can be expressed as:

$$V(total) = V(fundamental) + P(AI)$$

where *V(fundamental)* represents traditional valuation factors and *P(AI)* represents the AI Premium, capturing market expectations regarding future AI-driven revenue streams not yet realized.

2.2 Real-Options Interpretation

The AI Premium can be interpreted using real-options theory, where AI investment represents a call option on future technological breakthroughs:

$$P(AI) \propto f(\ \sigma(market),\ p(capture)\ )$$

where *σ(market)* represents technological uncertainty and *p(capture)* represents the probability of capturing AI-driven market value. This interpretation explains why firms with uncertain but potentially transformative AI capabilities often command large premiums even absent current-period earnings.

2.3 Signal as Premium Stabilization

Proposition 1 (Premium Stabilization):
AI Signal reduces uncertainty in expected AI-driven cash flows, thereby stabilizing the AI Premium rather than initiating it. Narratives may generate an initial speculative premium; only verifiable capability sustains it.

Strong signal reduces uncertainty regarding future technological capability, lowering the risk discount applied by investors. Conversely, weak signal increases valuation volatility and may lead to rapid premium collapse when narrative expectations fail to materialize. This stabilizing relationship forms a central edge of the AI Credibility Triangle and distinguishes the framework from simpler hype-cycle models that treat signal and narrative as independent.

3. AI-Washing

AI-Washing refers to the strategic amplification of AI claims in corporate communication beyond underlying technological capability. AI-Washing can be interpreted as a rational response to the capital allocation distortions created by the AI Premium: as long as markets reward AI narrative, firms face incentives to amplify it regardless of underlying capability.

Common forms include:

- Rebranding existing software products as AI-enabled without material model integration
- Emphasizing superficial machine-learning features in marketing communications
- Marketing AI capability without corresponding infrastructure investment
- Selective disclosure of favorable benchmark results while omitting known limitations

Such behavior increases informational noise in markets, creating the asymmetric information environment analyzed in Section 6.

4. AI Signal

AI Signal represents verifiable evidence of genuine artificial intelligence capability. Because technical disclosure is often limited, verification relies on multiple categories of indirect evidence across four primary domains:

1. Model development – proprietary architectures, benchmark performance, training infrastructure
2. Compute infrastructure – accelerator procurement, cluster deployment, data-center capacity
3. Enterprise deployment – commercial adoption, AI-driven revenue, customer scale
4. Research output – publications, open-source contributions, technical talent density

Definition: The Autonomy Gap

Definition (Autonomy Gap):
The Autonomy Gap is the measurable difference between advertised AI autonomy and the level of human intervention actually required for system operation. A small Autonomy Gap indicates genuine automation capability; large gaps indicate narrative exaggeration and are a primary observable symptom of AI-Washing.

The Autonomy Gap is particularly observable in Wizard-of-Oz systems that describe autonomous AI behavior while relying on concealed manual processes. It provides analysts with a concrete, operational criterion for evaluating AI-Washing that does not require access to proprietary technical documentation.

5. The Signal-to-Wash Index

To operationalize credibility dynamics, this paper defines the Signal-to-Wash Ratio (R_sw):

$$R\_sw = S \;/\; N$$

where S is the Signal score and N is the Narrative score, each independently scored on a 0–10 scale. The index is intended as a heuristic classification tool, not a precise financial metric. Scores are illustrative and subject to revision as disclosure standards evolve.

5.1 Signal Score Indicators

Indicator	Score Range
Proprietary model development	0–3
Compute infrastructure ownership	0–3
Enterprise deployment	0–2
Research publications	0–2

Note on sector calibration: Signal indicators may carry different practical weight across firm types. For hardware and infrastructure providers, compute infrastructure ownership is often the dominant verifiable signal. For software-oriented firms, model development and enterprise deployment may be more diagnostically meaningful. Analysts applying this rubric to cross-sector comparisons should apply contextual judgment to indicator weighting.

5.2 Narrative Score Indicators

Indicator	Score Range
AI mention frequency in filings	0–3
Marketing emphasis on AI	0–3
Public positioning as AI leader	0–2
Revenue narrative dependence	0–2

5.3 Interpretation

R_{sw} Range	Interpretation
> 1.5	High signal credibility
0.7 – 1.5	Balanced narrative
< 0.7	Narrative inflation

The threshold $R_{sw} > 1.5$ requires Signal to exceed Narrative by at least 50 percent, establishing a deliberately demanding credibility bar. None of the illustrative firms in Section 9 reach this threshold, reflecting the contemporary AI market's high-entropy informational environment in which narrative intensity currently outpaces verifiable signal across even the strongest participants. This threshold is aspirational and may be recalibrated as larger empirical datasets of AI disclosures accumulate.

6. Information-Theoretic Interpretation

The credibility triangle can be interpreted through the communication theory of Shannon (1948). In this analogy, firms transmit signals about technological capability through the financial disclosure ecosystem, while AI-Washing introduces noise that degrades the quality of investor inference.

Communication Element	AI Market Equivalent
Sender	Firm
Message	Discrete corporate disclosures: product announcements, benchmark results, earnings statements, research publications
Signal	Verifiable AI capability embedded in disclosures
Noise	AI-Washing (narrative amplification beyond capability)
Receiver	Investors
Channel	Financial disclosure ecosystem

A critical asymmetry reinforces this framework: high-signal firms such as large-scale compute providers invest billions in physical infrastructure, which constitutes a costly signal in the sense of Spence (1973) — one that is difficult and expensive to mimic. Marketing narratives, by contrast, are low-cost and easy to replicate. This cost differential means that genuine AI Signal carries disproportionate informational weight relative to its surface salience.

As AI-Washing increases across the market, the noise floor of the communication channel rises. In extreme cases, rising noise may produce signal pooling equilibria in which high- and low-quality firms become indistinguishable to investors — precisely the market failure that Spence's signaling model identifies as the consequence of costless mimicry. When this occurs, capital allocation efficiency degrades and systematic mispricing becomes probable.

7. The AI Credibility Matrix

The triangle produces a two-dimensional classification system with AI Signal Strength on the vertical axis and Narrative Integration on the horizontal axis.

Figure 1 — The AI Credibility Matrix

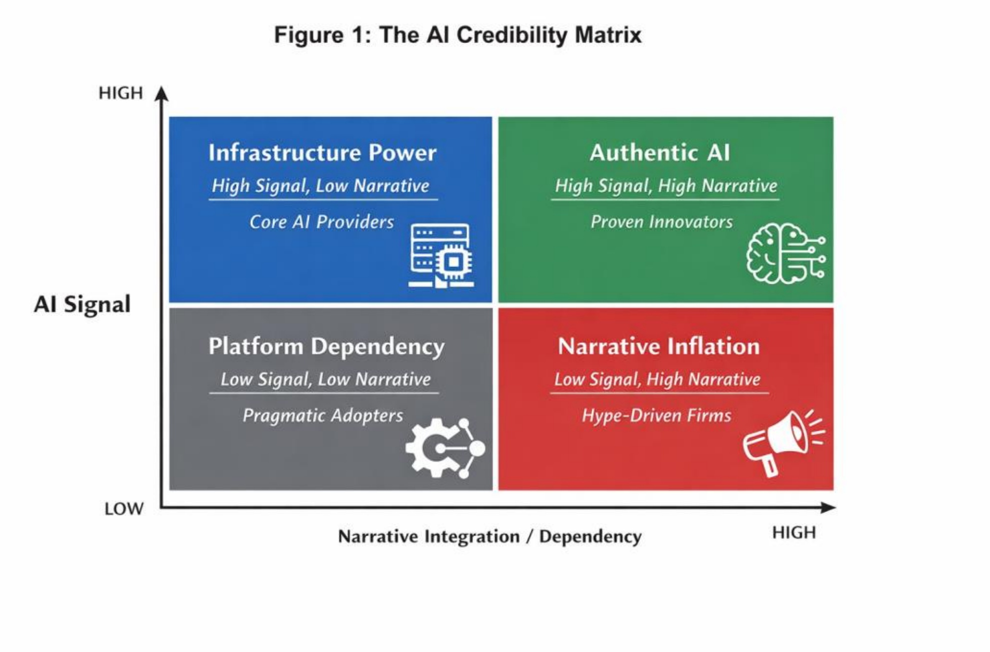

Figure 1: The AI Credibility Matrix. Firms are classified along two axes: AI Signal Strength (vertical) and Narrative Integration (horizontal). The four quadrants — Infrastructure Power (High Signal, Low Narrative), Authentic AI (High Signal, High Narrative), Platform Dependency (Low Signal, Low Narrative), and Narrative Inflation (Low Signal, High Narrative) — capture distinct market positions. Infrastructure Power firms derive their position from observable physical capability rather than narrative intensity.

Figure 2 — The AI Credibility Triangle

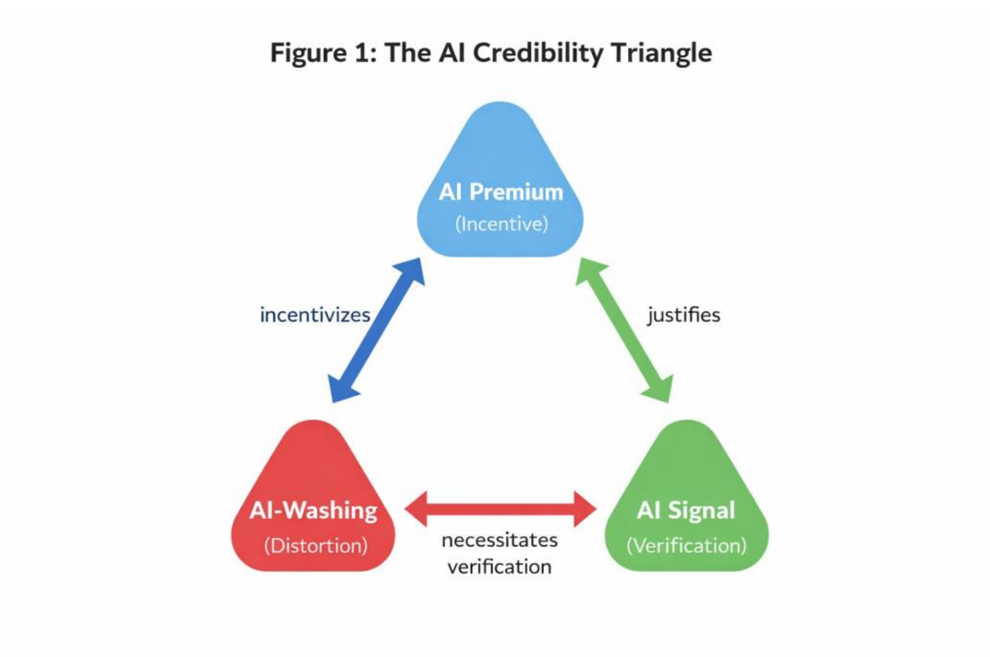

Figure 2: The AI Credibility Triangle. Three interacting forces — AI Premium (incentive), AI-Washing (distortion), and AI Signal (verification) — form a dynamic credibility system. AI Signal stabilizes the AI

Premium (Proposition 1); AI-Washing necessitates verification; the AI Premium incentivizes both signal investment and narrative amplification.

Matrix Quadrants

Quadrant	Signal	Narrative	Description
Authentic AI	High	High	Proven innovators whose AI positioning is supported by strong verifiable capability
Infrastructure Power	High	Low	Core AI providers — chips, compute, cloud — whose market position rests on observable infrastructure
Narrative Inflation	Low	High	Hype-driven firms with strong AI narratives but limited verifiable signal
Platform Dependency	Low	Low	Pragmatic adopters using AI tools without positioning as AI innovators

Scoring Guide

Score Range	Signal Classification	Narrative Classification
0–3	Low	Low
4–6	Moderate	Moderate
7–10	High	High

8. Technological Cycles and the Perez Framework

The dynamics of the AI Credibility Matrix correspond closely to the theory of technological revolutions developed by Carlota Perez (2002). Perez argues that major technological waves unfold through two sequential phases separated by a structural Turning Point. The quadrants of the AI Credibility Matrix map onto these phases as follows: Narrative Inflation corresponds to the Installation Phase; Infrastructure Power to the late Installation and early Deployment transition; and Authentic AI to the mature Deployment Phase.

Installation Phase (Perez 2002, Chapters 2–3)

Installation periods are characterized by financial speculation and narrative amplification around emerging technologies. Speculative capital flows toward firms promising technological transformation, often ahead of productive capacity. This dynamic closely corresponds to the Narrative Inflation quadrant.

It is important to note that Perez’s Installation Phase also involves genuine infrastructure investment; therefore the correspondence is not exact. Infrastructure builders may operate during this phase while simultaneously accumulating genuine signal — explaining why Infrastructure Power firms can coexist with Narrative Inflation firms during the same market cycle.

The Turning Point (Perez 2002, Chapter 4)

The Turning Point marks the structural transition from speculative capital to productive deployment. Firms that accumulated genuine technological capability during the Installation Phase begin to consolidate their positions. In the AI Credibility Matrix, this corresponds to a migration of firms from Narrative Inflation toward Infrastructure Power and Authentic AI — a movement driven by the market's growing ability to distinguish signal from noise.

Deployment Phase

The Deployment Phase is characterized by productivity gains, infrastructure consolidation, and market dominance by firms that successfully converted speculative premiums into operational capability. Infrastructure Power and Authentic AI firms dominate this phase. Platform Dependency firms may also benefit as AI tools become broadly accessible. The AI Credibility Matrix thus describes not only a cross-sectional snapshot but also the longitudinal migration of firms across a technological revolution.

9. Illustrative Firm Scoring

To demonstrate the framework, four firms are scored using the rubric defined in Section 5. Scores are derived from 2024–2025 annual reports (10-K filings), earnings-call transcripts, product announcements, and infrastructure disclosures. All scores are illustrative and subject to revision as disclosure standards evolve.

Indicator Scoring Table

Firm	Models	Compute	Deploy	Research	S	Filings	Mktg	Position	Rev Narr	N	R_sw
NVIDIA	2	3	2	1	8	2	2	2	1	7	1.14
Microsoft	3	2	2	2	9	3	3	2	2	10	0.90
C3.ai	1	1	1	0	3	3	3	2	2	10	0.30
Box Inc.	0	0	1	0	1	2	2	1	1	6	0.17

Data Sources

Scoring is based on the following primary sources:

- NVIDIA: Compute score of 3 reflects H100 and Blackwell GPU cluster disclosures in NVIDIA 2024 Annual Report and Q4 2025 earnings. Model score of 2 reflects NeMo and CUDA ecosystem rather than frontier foundation models. Hardware-focused firms may naturally score lower on model development; see the sector calibration note in Section 5.1.
- Microsoft: Model score of 3 reflects OpenAI partnership depth, Azure OpenAI Service deployment, and Copilot integration across the product suite (Microsoft 2024 Annual Report; Azure AI product disclosures). Narrative score of 10 reflects the highest observed marketing and filing intensity in the sample.

- C3.ai: Marketing score of 3 reflects consistent AI-first branding in SEC filings and investor communications. Research score of 0 reflects absence of peer-reviewed model publications (C3.ai 10-K 2024).
- Box Inc.: Scores reflect AI feature additions (Box AI) without underlying model development or compute ownership (Box 2024 Annual Report).

Boundary Case: NVIDIA

NVIDIA's R_sw = 1.14 places it within the Balanced Narrative band of the Signal-to-Wash Index, yet it is classified as Infrastructure Power in the matrix. This apparent inconsistency illustrates a structural feature of the framework: matrix classification depends on absolute signal and narrative levels, while R_sw reflects their ratio. A firm with S = 8 and N = 7 occupies a fundamentally different market position than a firm with S = 3 and N = 2.6, even though both yield the same ratio, because absolute signal strength determines actual technological capacity.

Additionally, NVIDIA's high absolute narrative score (N = 7) may reflect not opportunistic amplification but the market's growing recognition of infrastructure dominance — suggesting that the meaning of a given narrative score may differ between firm types. This nuance merits further theoretical development and illustrates the importance of reading R_sw and matrix placement together rather than independently.

Matrix Placement Summary

Firm	Signal (S)	Narrative (N)	R_sw	Quadrant
NVIDIA	8	7	1.14	Infrastructure Power
Microsoft	9	10	0.90	Authentic AI
C3.ai	3	10	0.30	Narrative Inflation
Box Inc.	1	6	0.17	Platform Dependency

10. Policy and Market Implications

The AI Credibility Triangle suggests that market transparency may benefit from stronger disclosure standards regarding artificial intelligence capabilities. The U.S. Securities and Exchange Commission has begun examining AI-related corporate disclosures, issuing guidance on the use of AI claims in marketing materials and investor communications (SEC Release No. 34-101718, 2024; readers should verify the current status of this release at sec.gov).

Potential transparency improvements consistent with the framework include:

- Disclosure of compute infrastructure capacity (accelerator count, data-center investment)
- Reporting of AI-attributable revenue streams separately from total revenue
- Documentation of model deployment scale and human-intervention rates
- Standardized benchmark reporting with known limitations disclosed

Such disclosures would reduce the narrative noise in the financial disclosure channel — consistent with the Shannon framework of Section 6 — improving the market's ability to distinguish genuine AI Signal from AI-Washing and reducing the risk of the signal pooling equilibria described above.

11. Conclusion

The rapid rise of artificial intelligence has produced a market environment where technological capability and narrative amplification interact in complex and consequential ways. The AI Credibility Triangle provides a structured framework for understanding these dynamics by linking valuation incentives (AI Premium), narrative amplification (AI-Washing), and verifiable technological capability (AI Signal).

The framework's central theoretical claim — formalized as Proposition 1 — is that AI Signal stabilizes rather than creates the AI Premium. This has important implications for both investment practice and regulatory design: market corrections in AI valuations are most likely to occur not when narratives fade, but when the stabilizing signal fails to materialize at the scale implied by valuation multiples.

The Signal-to-Wash Index and AI Credibility Matrix operationalize these insights into practical classification tools. The finding that none of the illustrative firms reach the R_sw > 1.5 credibility threshold is itself informative, confirming that even the strongest current AI participants operate in a high-narrative environment where signal verification remains a collective challenge.

Future research may extend this framework through larger empirical datasets of AI-related disclosures, infrastructure investment, and productivity outcomes, and by tracking firm migration across matrix quadrants over successive technological cycles.

References

Brynjolfsson, E., Li, D., & Raymond, L. R. (2023). Generative AI at work. NBER Working Paper No. 31161. National Bureau of Economic Research.

DeLong, J. B., & Magin, K. (2006). The U.S. equity return premium: Past, present, and future. Journal of Economic Perspectives, 19(4), 3–20.

Perez, C. (2002). Technological revolutions and financial capital: The dynamics of bubbles and golden ages. Edward Elgar Publishing.

Shannon, C. E. (1948). A mathematical theory of communication. Bell System Technical Journal, 27(3), 379–423.

Shiller, R. J. (2019). Narrative economics: How stories go viral and drive major economic events. Princeton University Press.

Spence, M. (1973). Job market signaling. Quarterly Journal of Economics, 87(3), 355–374.

U.S. Securities and Exchange Commission. (2024). Guidance on artificial intelligence-related disclosures in corporate filings. SEC Release No. 34-101718. Available at: https://www.sec.gov

Champaign Magazine. (2026, March 12). Aikipedia: AI Premium. https://champaignmagazine.com/2026/03/12/aikipedia-ai-premium/

Champaign Magazine. (2026, March 9). Aikipedia: AI-Washing. https://champaignmagazine.com/2026/03/09/aikipedia-ai-washing/

Champaign Magazine. (2026, March 13). Aikipedia: AI Signal. https://champaignmagazine.com/2026/03/13/aikipedia-ai-signal/