JIRRS
Journal International Review of Research Studies

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

THE AI INVESTMENT CYCLE: STRUCTURAL ANALOGIES WITH THE DOT-COM BUBBLE AND EVIDENCE FOR A MATURE TECHNOLOGICAL EXPANSION

Andriny Dias Mascarenhas
Faculdade Presidente Antônio Carlos, Brazil
Corresponding author: andrinymascarenhas993@gmail.com

Abstract

The rapid expansion of Artificial Intelligence (AI) investment since 2022 has prompted widespread comparisons with the dot-com bubble of the late 1990s. This article critically examines whether the current AI investment cycle shares the structural characteristics that defined the dot-com collapse, focusing on six analytical dimensions: the retrospective anatomy of the dot-com bubble and its infrastructure failures; the theoretical frameworks used to understand speculative technology cycles; the empirical evaluation of the three central pillars supporting current AI investment; the structural differences between the AI cycle and the dot-com era; the emerging energy infrastructure constraint that may represent the principal bottleneck to future growth; and the investment implications arising from these developments. Drawing upon peer-reviewed literature, Federal Reserve analyses, International Energy Agency reports, Brookings Institution research, and contemporary financial market data, the article argues that the AI cycle differs fundamentally from the dot-com era in terms of infrastructure maturity, investor composition, revenue generation, and user adoption. At the same time, it identifies energy availability, grid expansion, and infrastructure financing as the principal unresolved risks facing the sector. The evidence suggests that while localized speculation and valuation excesses may exist, the underlying economic foundations of the AI cycle differ substantially from those that characterized the collapse of the internet bubble. The article concludes that the most significant challenge facing the AI ecosystem is not demand creation but the capacity of supporting infrastructure to scale alongside rapidly growing computational requirements.

Keywords: artificial intelligence; dot-com bubble; speculative investment; technology cycles; scaling laws; infrastructure economics; energy transition; Jevons paradox.

JIRRS
Journal International Review of Research Studies

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

1. Introduction

In 1995, a live webcam pointed at a coffee pot in the Computer Laboratory of the University of Cambridge became one of the earliest demonstrations of the internet's ability to transform ordinary activities into shared digital experiences. Although trivial in itself, the episode symbolized a broader transition: the internet was evolving from a specialized research network into a platform with significant commercial and social potential. During the years that followed, investor enthusiasm surrounding the emerging digital economy intensified dramatically, contributing to one of the most significant speculative episodes in modern financial history (Shiller, 2000; Couper et al., 2003).

Between January 1995 and March 2000, the NASDAQ Composite Index increased by approximately 572 percent, reaching a peak of 5,048.62 points as investors poured capital into internet startups, telecommunications companies, and infrastructure projects (Goldman Sachs, 2019). The dominant narrative of the period held that the internet would rapidly transform commerce, communication, media, and information exchange. In retrospect, this expectation was largely correct. However, investors significantly overestimated the speed at which adoption would generate sustainable profits and underestimated the infrastructure constraints limiting commercialization. The resulting correction erased trillions of dollars in market value and became one of the most frequently studied examples of a technology-driven speculative bubble (Shiller, 2000; Ofek & Richardson, 2003; Lansing, 2008).

More than two decades later, the emergence of generative artificial intelligence has triggered a comparable wave of public attention, corporate investment, and market optimism. The release of ChatGPT in November 2022 demonstrated the commercial viability of large language models and initiated one of the fastest technology adoption cycles ever recorded. Within a short period, AI systems transitioned from experimental research tools to widely deployed products used by consumers, corporations, software developers, educators, and governments (Omnibound, 2026). Simultaneously, major technology firms committed hundreds of billions of dollars toward AI infrastructure, data center expansion, semiconductor procurement, and model development.

The similarities between the current AI investment cycle and the dot-com era have generated intense debate among economists, investors, and policymakers. Critics argue that rapidly rising valuations, unprecedented capital expenditures, and highly optimistic growth projections resemble the conditions that preceded the collapse of the internet bubble. Supporters counter that such comparisons overlook fundamental

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

structural differences, including the maturity of digital infrastructure, the existence of established revenue streams, and the demonstrable utility of AI systems across multiple sectors of the economy (Kaplan et al., 2020; Hoffmann et al., 2022). Consequently, the central question is not whether speculation exists, but whether the current cycle rests upon stronger economic foundations than those that characterized the late 1990s.

This distinction is important because speculative behavior and technological progress are not mutually exclusive phenomena. The literature on financial bubbles consistently demonstrates that transformative technologies frequently generate periods of excessive optimism, aggressive investment, and inflated expectations before their long-term economic impact becomes fully understood (Minsky, 1992; Shiller, 2000; Lansing, 2008). Historical examples ranging from railroads and electrification to the internet itself reveal that markets often struggle to distinguish between unrealistic short-term expectations and genuine long-term technological value. As a result, the existence of speculation alone does not necessarily imply that an underlying technological transformation lacks economic significance.

A second consideration concerns technology adoption and productivity. Research on digital transformation indicates that once technologies demonstrate measurable economic benefits, adoption can accelerate rapidly as firms seek efficiency gains and competitive advantages (Brynjolfsson & McElheran, 2016). The widespread deployment of generative AI tools across knowledge-intensive industries suggests that current adoption patterns may reflect more than temporary enthusiasm. Unlike many internet startups of the late 1990s, which relied primarily on expectations of future profitability, many of today's AI investments are being undertaken by highly profitable incumbent firms with established customer bases and substantial operating cash flows.

At the same time, the AI cycle faces challenges that differ significantly from those encountered during the dot-com era. The internet boom was constrained by limitations in telecommunications infrastructure, particularly the inability of residential connectivity to scale as quickly as backbone investment. Contemporary AI systems operate within a mature digital environment, but they depend on rapidly expanding computational resources and increasingly large quantities of electricity. Recent analyses by the International Energy Agency project substantial growth in data center power consumption over the coming decade, while Brookings Institution researchers have identified energy infrastructure as a potentially critical factor shaping future AI deployment (IEA, 2025; IEA, 2026; Muro et al., 2025). Consequently, the principal bottleneck facing AI may not be user adoption but the ability of physical infrastructure to supply the energy required for continued computational expansion.

This article argues that meaningful comparisons between the AI cycle and the dot-com bubble must focus on structural conditions rather than surface-level similarities. While both periods exhibit strong investor enthusiasm, large-scale infrastructure investment, and elevated expectations regarding technological transformation, they differ substantially in terms of infrastructure maturity, investor composition, revenue generation, and adoption dynamics. The article further argues that the most significant unresolved risk facing the AI ecosystem is not demand creation but the capacity of energy infrastructure to scale alongside computational requirements.

To investigate these issues, the study combines insights from financial economics, technology diffusion theory, infrastructure economics, and contemporary AI research. The discussion begins with a methodological overview and a reconstruction of the dot-com bubble as a case study in infrastructure overinvestment and adoption constraints. It then examines the theoretical foundations of speculative technology cycles before evaluating the three central pillars supporting the current AI investment thesis: continued model improvement, the necessity of data center expansion, and widespread user adoption. The analysis subsequently explores the principal structural differences between the two cycles, examines the emerging energy bottleneck, and evaluates the investment implications arising from these developments.

The article is organized as follows. Section 2 presents the methodology. Section 3 reconstructs the dot-com bubble as a case study in infrastructure overinvestment and adoption constraints. Section 4 provides the theoretical framework integrating behavioral finance, financial instability theory, and technology-cycle analysis. Section 5 evaluates the three central pillars of the current AI investment thesis. Section 6 identifies the key structural differences between the AI cycle and the dot-com bubble. Section 7 examines the energy infrastructure bottleneck. Section 8 discusses investment implications and risk assessment. Section 9 presents the conclusions.

Before proceeding, it is important to clarify the meaning of the expression “mature technological expansion,” which appears in the title and constitutes a central concept of this study. The term does not imply that artificial intelligence has reached maturity in terms of profitability, competitive equilibrium, investment efficiency, or long-term monetization outcomes. Rather, it refers primarily to the relative maturity of the broader digital ecosystem surrounding AI, including cloud-computing infrastructure, global internet connectivity, semiconductor supply chains, enterprise software integration, and observable patterns of user and organizational adoption (Bresnahan & Trajtenberg, 1995; Brynjolfsson & McElheran, 2016).

Accordingly, the argument advanced in this article is not that the AI sector itself has reached a mature stage of economic development. Instead, the study proposes that the current AI investment cycle operates within a substantially more mature technological and infrastructural environment than the one that existed during the dot-com era. This distinction is important because technological maturity, infrastructure maturity, market maturity, and investment-cycle maturity are related but analytically distinct concepts. Similar distinctions have been discussed in the literature on general-purpose technologies, technology diffusion, and long-run technological transitions (Bresnahan & Trajtenberg, 1995; Rogers, 2003).

2. Methodology

This study adopts a narrative literature review approach combined with comparative historical-structural analysis. Narrative reviews are particularly suitable for synthesizing interdisciplinary evidence, integrating findings from multiple domains, and evaluating emerging phenomena that have not yet generated a sufficiently mature body of standardized empirical research (Baumeister & Leary, 1997; Grant & Booth, 2009; Snyder, 2019). Given the ongoing nature of the contemporary artificial intelligence (AI) investment cycle, this approach was considered appropriate for examining the similarities and differences between the current AI expansion and the dot-com bubble of the late 1990s and early 2000s.

The review protocol was designed to improve transparency, reduce selection bias, and ensure a balanced representation of competing perspectives. Literature was collected from multiple academic and institutional sources, including Google Scholar, Scopus, Web of Science, SSRN, the National Bureau of Economic Research (NBER), the International Energy Agency (IEA), the Organisation for Economic Co-operation and Development (OECD), the International Monetary Fund (IMF), and publications from the Federal Reserve System. Additional evidence was obtained from institutional reports, financial disclosures, and industry analyses when peer-reviewed evidence was unavailable or insufficient to address rapidly evolving developments within the AI sector.

The search process was conducted between January and April 2026. Principal search terms included combinations of “AI bubble,” “artificial intelligence investment,” “dot-com bubble,” “technology bubbles,” “technological revolutions,” “general-

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

purpose technologies," "AI productivity," "AI adoption," "AI infrastructure," "data-center investment," "electricity demand and AI," "technology diffusion," "innovation cycles," and "speculative finance." Additional studies were identified through backward and forward citation tracking.

The temporal scope primarily covered literature published between 2000 and April 2026. Earlier foundational works were included when necessary to establish theoretical concepts central to the analysis, including speculative bubbles, financial instability, technology diffusion, and rebound effects. Consequently, seminal contributions such as those of Minsky (1986), Shiller (2000), Jevons (1865), and Rogers (2003) were retained due to their enduring relevance to the study's conceptual framework.

Sources were selected according to four inclusion criteria: (1) direct relevance to artificial intelligence, technology investment cycles, speculative bubbles, infrastructure economics, or technology diffusion; (2) empirical, theoretical, or institutional contribution to the topic; (3) recognized authority or citation impact within the relevant field; and (4) sufficient methodological transparency to permit evaluation of evidentiary quality. Exclusion criteria included duplicate publications, opinion pieces lacking supporting evidence, promotional materials, blog posts, and sources containing unverifiable claims.

Because the AI investment cycle remains an evolving phenomenon, some recent information was obtained from industry reports, financial disclosures, and journalistic sources. These sources were not treated as equivalent to peer-reviewed scholarship. Instead, evidence was evaluated according to a hierarchical framework. Peer-reviewed academic literature and institutional reports constituted the primary evidentiary base, while financial reports, industry analyses, and journalistic sources were used primarily to document recent market developments, infrastructure investments, adoption trends, and capital-expenditure announcements. Whenever possible, claims derived from non-peer-reviewed sources were cross-checked against academic or institutional evidence.

The analytical framework employed in this review follows principles commonly associated with comparative historical analysis (Mahoney & Rueschemeyer, 2003). The AI investment cycle and the dot-com era were compared across seven structural dimensions: infrastructure maturity, investor composition, revenue generation, adoption dynamics, infrastructure bottlenecks, capital structure, and sources of uncertainty. Evidence relating to each dimension was synthesized comparatively to identify areas of similarity, divergence, and unresolved uncertainty.

JIRRS
Journal International Review of Research Studies

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

Given the rapidly evolving nature of AI markets, all contemporary data, market indicators, adoption statistics, and infrastructure projections referenced in this review should be interpreted as reflecting information available up to April 2026, which serves as the study's data cut-off date. Consequently, the findings should be understood as provisional and subject to revision as new evidence emerges.

2.1 Review Protocol

To improve transparency and reduce selection bias, a structured narrative-review protocol was adopted. Although the present study does not constitute a systematic review, source selection followed predefined procedures designed to ensure relevance, evidentiary quality, and thematic balance.

The review process consisted of four stages: identification, screening, eligibility assessment, and final inclusion. Initial searches across academic databases, institutional repositories, and reference lists produced a broad pool of potentially relevant publications. Sources were then screened for relevance to at least one of the study's central themes: artificial intelligence economics, speculative bubbles, technology cycles, infrastructure investment, technology diffusion, productivity effects, or energy constraints.

After screening, sources were evaluated according to methodological transparency, empirical contribution, theoretical relevance, and institutional credibility. Preference was given to peer-reviewed literature and authoritative institutional reports, although selected industry and financial sources were retained when addressing rapidly evolving developments for which academic literature remained limited.

Table 1 summarizes the review process.

Table 1. Narrative Review Selection Process

Review Stage	Number of Sources
Records initially identified	184
After preliminary screening	103

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

Eligible for detailed evaluation	71
Final sources included in the review	45

Source: Elaborated by the authors.

The final corpus included peer-reviewed journal articles, books, institutional reports, governmental publications, financial analyses, and selected industry reports. This combination was considered necessary due to the interdisciplinary nature of the topic and the rapidly evolving characteristics of the contemporary AI investment cycle.

2.2 Evidence Hierarchy and Source Classification

Because this study examines an ongoing technological and investment cycle, evidence originates from multiple types of sources with varying degrees of methodological rigor, temporal stability, and evidentiary value. To improve transparency and reduce the risk of treating heterogeneous evidence as equivalent, all sources were classified according to a hierarchical framework.

The primary evidentiary base consists of peer-reviewed academic literature, which provides theoretical foundations, empirical findings, and methodological rigor. Institutional reports published by internationally recognized organizations constitute the second level of evidence, particularly when addressing macroeconomic trends, infrastructure development, energy demand, and technology adoption.

Financial reports, industry analyses, and journalistic sources were included primarily to document recent developments that may not yet be fully represented in the academic literature. However, such sources were interpreted with greater caution and were not considered equivalent to peer-reviewed research when supporting major analytical conclusions.

Table 2 summarizes the evidence hierarchy adopted in this study.

Table 2. Evidence Hierarchy Used in the Review

Evidence Tier	Source Type	Examples	Primary Use
Tier 1	Peer-reviewed	Journal articles,	Theoretical

	literature	academic books	foundations and empirical evidence
Tier 2	Institutional reports	IEA, OECD, IMF, Federal Reserve, NBER	Macroeconomic, infrastructure, and policy evidence
Tier 3	Financial and market reports	Goldman Sachs, company financial disclosures	Capital expenditure, investment trends, and market data
Tier 4	Industry reports	Technology-sector analyses and specialized industry publications	Adoption trends and sector developments
Tier 5	Journalistic sources	Fortune, CNBC, Reuters, Bloomberg	Recent events, announcements, and contextual information

Source: Elaborated by the authors.

Whenever possible, claims supported by Tier 3, Tier 4, or Tier 5 sources were cross-checked against Tier 1 or Tier 2 evidence. This approach was adopted to balance the need for methodological rigor with the necessity of incorporating rapidly evolving developments within the contemporary AI ecosystem.

3. The Dot-Com Bubble: A Structural Retrospective

The dot-com bubble remains one of the most influential examples of a technology-driven speculative cycle in modern economic history. Between the mid-1990s and early 2000, investors poured unprecedented amounts of capital into internet-related ventures, telecommunications infrastructure, and digital business models. The underlying assumption was that the internet would rapidly transform nearly every aspect of economic and social activity, creating extraordinary opportunities for growth and profitability. While the long-term technological prediction ultimately proved correct, financial markets significantly overestimated both the speed and the economic returns associated with the transition (Shiller, 2000; Ofek & Richardson, 2003).

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

A common misconception is that the dot-com collapse represented the failure of the internet itself. In reality, the internet continued expanding rapidly after the market crash and eventually became one of the most economically significant technologies in human history. The failure occurred primarily in the relationship between investor expectations and the practical constraints governing technology adoption. Understanding these constraints is essential because many contemporary comparisons between artificial intelligence and the dot-com era focus on superficial similarities while overlooking the structural conditions that ultimately determined the outcome of the earlier cycle (Lansing, 2008).

3.1 Infrastructure Overinvestment and Dark Fiber

One of the defining characteristics of the dot-com era was the enormous wave of infrastructure investment directed toward telecommunications networks. Anticipating explosive growth in internet traffic, firms invested heavily in fiber-optic networks, switching equipment, data transmission systems, and internet backbone capacity. Investors correctly identified that digital communication would become increasingly important; however, they frequently assumed that demand would materialize at a pace sufficient to justify immediate large-scale infrastructure deployment (Couper et al., 2003).

The most visible manifestation of this phenomenon was the construction of vast quantities of fiber-optic infrastructure. Telecommunications companies competed aggressively to expand network capacity, often financing projects through substantial debt issuance and speculative capital. By the early 2000s, large portions of installed fiber remained unused, leading to the emergence of what became known as “dark fiber”—optical cables that had been deployed but were not carrying traffic because demand had not yet reached projected levels (Couper et al., 2003; Lansing, 2008).

Importantly, the existence of dark fiber did not imply that the investment was entirely wasted. Much of the infrastructure built during the bubble eventually became essential for supporting later phases of internet growth. Nevertheless, the timing mismatch between infrastructure deployment and commercial adoption proved financially devastating for many firms. Investors had correctly anticipated the long-term direction of technological change but had incorrectly estimated the speed at which economic returns would materialize (Couper et al., 2003).

The speculative dynamics surrounding internet-related firms amplified these problems. Ofek and Richardson (2003) demonstrated that many internet company valuations became increasingly disconnected from conventional financial fundamentals. Investors frequently justified extreme valuations using assumptions regarding future market dominance, network effects, and exponential revenue growth. In many cases, firms with limited revenue, uncertain business models, and persistent losses achieved market capitalizations comparable to those of established corporations. The resulting divergence between market prices and observable fundamentals became a defining feature of the bubble.

3.2 The Last-Mile Problem

Although excessive valuation and speculative behavior contributed to the collapse, infrastructure constraints played an equally important role. The telecommunications sector successfully expanded backbone capacity, but the physical infrastructure connecting households and businesses to the internet developed more slowly. This challenge became known as the “last-mile problem” and represented one of the most significant barriers to widespread internet adoption during the late 1990s (Couper et al., 2003).

At the time, many consumers relied on dial-up connections operating through existing telephone networks. Broadband technologies such as cable internet and DSL remained geographically limited and required substantial investment to achieve broad deployment. As a result, internet backbone infrastructure often expanded faster than the capacity of end users to access and utilize it. The mismatch created a structural bottleneck: while network operators prepared for massive increases in traffic, the number of users capable of generating that traffic grew more slowly than anticipated.

The importance of the last-mile problem extends beyond its historical significance. It demonstrates that technological adoption depends not only on the existence of a core innovation but also on the availability of complementary infrastructure capable of supporting widespread use. Investors during the dot-com era frequently focused on the transformative potential of the internet while underestimating the practical constraints that could delay realization of that potential (Lansing, 2008).

This distinction between technological capability and deployment capacity would ultimately become one of the most important lessons of the dot-com collapse. The internet was not overvalued because it lacked utility. Rather, expectations regarding adoption speed, infrastructure readiness, and commercial monetization proved excessively optimistic relative to prevailing economic conditions.

3.3 The Collapse and Its Structural Lessons

The NASDAQ Composite Index reached its peak in March 2000 before entering a prolonged decline. Over the following years, internet-related equities suffered

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

dramatic losses, numerous startups ceased operations, and several telecommunications firms entered bankruptcy or underwent major restructurings. The collapse became a defining case study in speculative finance and remains one of the most extensively analyzed market events in modern economic literature (Shiller, 2000; Lansing, 2008).

However, the long-term outcome of the dot-com era was more nuanced than a simple story of technological failure. The internet continued to expand after the crash, e-commerce adoption accelerated, broadband deployment increased, and many of the underlying technologies eventually generated enormous economic value. Companies such as Amazon survived the correction and later became among the most valuable firms in the world. In this sense, the technological thesis proved substantially more accurate than the financial expectations embedded within market valuations at the height of the bubble (Ofek & Richardson, 2003).

The primary lesson of the dot-com experience is therefore not that investors were wrong about the importance of the internet. Rather, they were wrong about the relationship between technological potential, infrastructure constraints, and the timing of economic returns. Financial markets frequently price transformative technologies as though adoption, monetization, and supporting infrastructure will develop simultaneously. In practice, these processes often occur at different speeds, generating periods of both extraordinary optimism and significant disappointment (Minsky, 1992; Shiller, 2000).

This lesson is particularly relevant when evaluating contemporary claims regarding artificial intelligence. The key analytical question is not whether AI possesses transformative potential, but whether the supporting conditions required to realize that potential differ materially from those that characterized the internet economy of the late 1990s. Answering this question requires a broader theoretical framework for understanding speculative technology cycles, which is developed in the following section.

4.The Anatomy of a Speculative Bubble: Theoretical Framework

Understanding whether the current AI investment cycle constitutes a speculative bubble requires a clear theoretical framework. Although the term “bubble” is widely used in public discourse, economists have long debated its precise definition. At its broadest level, a bubble occurs when asset prices become disconnected from underlying economic fundamentals and are sustained primarily by expectations of future price appreciation rather than by observable cash flows, productivity gains, or intrinsic value (Shiller, 2000). However, distinguishing between justified optimism and irrational speculation is often difficult, particularly when transformative technologies are involved.

Technological revolutions frequently create environments in which investors must evaluate opportunities whose future economic value is highly uncertain. Under such conditions, financial markets may simultaneously reflect rational expectations regarding long-term technological change and irrational expectations regarding the timing or magnitude of future returns. This duality explains why many of history's most important innovations—from railroads and electrification to the internet—have been accompanied by episodes of speculative excess (Shiller, 2000; Lansing, 2008).

4.1 Defining a Bubble

The concept of a speculative bubble remains contested within the economic literature. Different theoretical traditions emphasize distinct mechanisms, including valuation disconnection, speculative financing, narrative contagion, and unsustainable investment behavior. To avoid conceptual ambiguity, this study adopts a multidimensional definition that incorporates insights from both behavioral-finance and financial-instability perspectives.

For the purposes of this review, a speculative bubble is defined as a market condition characterized by one or more of the following elements: (1) a significant divergence between asset valuations and underlying economic fundamentals; (2) self-reinforcing speculative behavior driven by expectations of future price appreciation; (3) investment decisions based primarily on optimistic growth narratives rather than demonstrated economic performance; and (4) capital-allocation patterns that become increasingly dependent upon expectations of future expansion rather than current cash flows or productivity outcomes (Shiller, 2000; Minsky, 1992).

One of the most influential approaches to understanding speculative episodes is provided by Shiller (2000), who argues that bubbles are often driven by feedback mechanisms in which rising prices reinforce investor confidence, attracting additional participants and further increasing prices. In such environments, narratives become increasingly important. Investors may begin purchasing assets not because of current earnings or measurable fundamentals, but because they expect future buyers to value those assets even more highly.

This dynamic does not necessarily imply that the underlying technology lacks merit. Indeed, many bubbles emerge around innovations that later prove transformative.

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

The challenge lies in determining whether market valuations remain plausibly connected to future economic outcomes. Ofek and Richardson (2003), examining internet-related firms during the late 1990s, found that investors frequently assigned valuations based on assumptions of future dominance that were difficult to justify using conventional financial metrics. Their findings suggest that speculative behavior often emerges when expectations regarding future growth substantially outpace observable economic performance.

Lansing (2008) further argues that technological breakthroughs create a particularly difficult environment for valuation because traditional models struggle to estimate the future impact of disruptive innovations. As a result, investors may simultaneously underestimate long-term technological significance while overestimating short-term financial returns. This distinction is important because it helps explain how the internet could be both one of the most important innovations of the modern era and the center of one of history's largest speculative bubbles.

Table 3 summarizes the principal theoretical frameworks employed in this review and their specific relevance to the contemporary AI investment cycle.

Table 3. Theoretical Frameworks and Their Relevance to AI

Theory	Main Idea	Relevance to AI
Shiller (2000)	Bubbles emerge through narratives and feedback loops that amplify investor optimism.	AI enthusiasm may elevate valuations beyond current fundamentals.
Ofek & Richardson (2003)	Investors often value firms based on expectations of future dominance rather than present performance.	AI firms may receive valuations based on anticipated future market positions.
Lansing (2008)	Technological revolutions create uncertainty that complicates asset valuation.	AI's long-term economic impact remains difficult to estimate accurately.
Minsky (1992)	Periods of stability encourage increasing speculative risk-taking.	Strong confidence in AI may contribute to excessive investment behavior.

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

Source: Elaborated by the authors based on Shiller (2000), Minsky (1992), Ofek and Richardson (2003), and Lansing (2008).

4.2 Minsky's Financial Instability Hypothesis

A complementary perspective is provided by Minsky's Financial Instability Hypothesis. According to Minsky (1992), periods of economic stability often encourage increasingly aggressive risk-taking behavior. As confidence grows, investors become more willing to finance projects whose success depends on optimistic assumptions regarding future conditions. Over time, speculative financing expands, leverage increases, and financial markets become progressively more vulnerable to adverse shocks.

Minsky identifies a progression from conservative financing toward increasingly speculative forms of investment. During the early stages of a technological revolution, investment is often directed toward productive activities with identifiable economic value. As optimism intensifies, however, market participants may begin allocating capital on the assumption that favorable conditions will persist indefinitely. Under these circumstances, valuations can become increasingly sensitive to changes in expectations rather than changes in underlying fundamentals (Minsky, 1992).

The dot-com bubble exhibits many characteristics consistent with this framework. Investors correctly anticipated that internet technologies would become economically significant, but the widespread belief that nearly every internet-related venture would generate extraordinary returns proved unsustainable. The resulting correction demonstrated how rapidly confidence can deteriorate when expectations become detached from realistic assessments of adoption, infrastructure constraints, and profitability (Shiller, 2000; Lansing, 2008).

4.3 The Gartner Hype Cycle and Technology Diffusion

While financial theories explain speculative behavior, technology-adoption frameworks help explain why transformative innovations often generate exaggerated expectations. One of the most widely cited models is the Gartner Hype Cycle, which proposes that emerging technologies typically progress through a sequence of stages beginning with a technological breakthrough, followed by a period of inflated expectations, subsequent disillusionment, gradual maturation, and eventual productive adoption (Linden & Fenn, 2003).

The model is particularly useful because it distinguishes between short-term expectations and long-term technological value. Technologies frequently fail to meet the most optimistic projections made during the peak of hype, yet still achieve substantial economic significance over time. The internet itself followed a trajectory broadly consistent with this pattern. Excessive expectations during the late 1990s were followed by market correction, but the underlying technology continued evolving and ultimately transformed the global economy.

Research on technology diffusion reinforces this interpretation. Brynjolfsson and McElheran (2016) demonstrate that digital technologies capable of generating measurable productivity improvements often experience accelerating adoption once firms begin integrating them into operational processes. Adoption therefore tends to occur in waves rather than through smooth linear progression. Early enthusiasm may exceed immediate economic reality, but genuine productivity gains can eventually sustain widespread deployment and long-term value creation.

4.4 Applying the Framework to Artificial Intelligence

Taken together, these theoretical perspectives suggest that the presence of speculation alone is insufficient evidence that a technology lacks economic value. Shiller (2000) demonstrates that narratives and feedback loops can inflate valuations. Minsky (1992) explains how periods of optimism encourage increasingly speculative investment behavior. The Gartner framework highlights the tendency for expectations to exceed short-term capabilities, while technology-diffusion research shows that genuinely useful innovations often require time to achieve broad economic impact.

Consequently, evaluating the AI investment cycle requires more than identifying elevated valuations or rapid capital inflows. The critical question is whether the assumptions supporting current investment levels are fundamentally different from those that characterized the dot-com era. If AI adoption is already generating measurable economic value, if infrastructure constraints differ materially from those of the late 1990s, and if investment is being driven primarily by profitable incumbent firms rather than speculative startups, then comparisons with the internet bubble may be incomplete or misleading.

The following section examines this question directly by evaluating the three principal pillars supporting the contemporary AI investment thesis: continued model improvement, the strategic necessity of data center expansion, and the durability of current adoption trends.

5. The Three Pillars of the AI Investment Thesis

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

The contemporary AI investment cycle rests upon a set of assumptions regarding technological progress, infrastructure requirements, and market adoption. While investors, firms, and policymakers often disagree about the magnitude of future economic impacts, much of the current wave of investment can be traced to three interconnected propositions. First, AI models will continue to improve in capability and usefulness. Second, expanding computational infrastructure is necessary to support that improvement. Third, adoption will remain sufficiently strong to justify the capital being deployed throughout the ecosystem.

If any of these assumptions proves fundamentally incorrect, current investment levels may become difficult to justify. Conversely, if all three remain broadly valid, comparisons between the present cycle and the dot-com bubble become substantially weaker. Each proposition is therefore examined separately.

Table 4 summarizes the three pillars and the principal evidence associated with each.

Table 4. The Three Pillars of the AI Investment Thesis

Pillar	Core Assumption	Supporting Evidence
Continued Model Improvement	AI capabilities will continue advancing through scaling and innovation.	Scaling laws, benchmark improvements, frontier-model development.
Infrastructure Expansion	Data-center investment is necessary to support AI growth.	Growing demand for training, inference, cloud services, and GPUs.
Widespread Adoption	Users and organizations will continue adopting AI technologies.	Consumer adoption, enterprise deployment, and workflow integration.

Source: Elaborated by the authors based on Kaplan et al. (2020), Hoffmann et al. (2022), McElheran et al. (2023), Goldman Sachs (2025), and IEA (2025).

5.1 AI Will Continue to Improve

The first pillar of the contemporary AI investment thesis is the belief that artificial intelligence systems will continue to improve as additional computational resources, data, and algorithmic innovations are applied. This assumption is particularly important because many current investments are justified not only by the capabilities

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

of existing models but also by expectations regarding future performance and economic utility (Kaplan et al., 2020; Hoffmann et al., 2022).

Historically, improvements in machine learning systems were often viewed as uncertain and difficult to predict. However, research conducted over the last decade increasingly suggests that model performance follows identifiable scaling relationships. Kaplan et al. (2020) demonstrated that performance improvements in large language models could be predicted with remarkable consistency as model size, dataset size, and computational expenditure increased. Their findings suggested that capabilities previously considered emergent or unpredictable might instead result from relatively stable scaling dynamics (Kaplan et al., 2020).

Subsequent work by Hoffmann et al. (2022) refined this understanding by demonstrating that optimal performance does not depend solely on increasing parameter counts. Rather, model quality improves most efficiently when computational resources, training data, and model size are scaled in a balanced manner. The resulting “Chinchilla scaling laws” significantly influenced industry practice and contributed to a broader consensus that continued performance improvements remained achievable without requiring unlimited growth in model size (Hoffmann et al., 2022).

The practical implications of these findings extend beyond academic research. If model capabilities improve predictably as computational investment increases, expenditures on semiconductors, data centers, and training infrastructure become more than speculative bets. They become investments in a process that has repeatedly demonstrated measurable returns in model performance. This relationship helps explain why major technology firms continue allocating substantial resources toward AI development despite the enormous associated costs (Kaplan et al., 2020; Hoffmann et al., 2022).

Empirical developments since the release of ChatGPT have generally reinforced this perspective. Successive generations of frontier models have demonstrated improvements across a wide range of benchmarks, including reasoning, coding, multimodal understanding, scientific problem solving, and agentic task execution. Although benchmark methodologies remain imperfect, the overall trend has been one of consistent capability growth across successive model generations (OpenAI, 2023; Kaplan et al., 2020; Hoffmann et al., 2022).

The rapid diffusion of generative AI technologies also provides indirect evidence supporting the continued improvement hypothesis. Technologies that fail to deliver practical utility rarely achieve sustained adoption at scale. By contrast, generative AI

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

systems have been integrated into software development, content creation, education, customer support, scientific research, and enterprise productivity workflows within a relatively short period of time. Research on digital technology adoption suggests that such deployment patterns are more likely when users perceive measurable value from the technology being adopted (Brynjolfsson & McElheran, 2016; McElheran et al., 2023).

At the same time, the assumption of indefinite improvement should not be treated as certain. Several researchers have raised concerns regarding diminishing returns, data availability, training costs, and benchmark saturation as model scales continue to expand. As models become increasingly capable, achieving additional gains may require disproportionately larger investments in computation, infrastructure, and specialized hardware. The existence of scaling laws therefore does not guarantee that progress will continue indefinitely at current rates, nor does it eliminate the possibility of technological plateaus or slower future advancement (Hoffmann et al., 2022; Sinha et al., 2025).

Recent developments in model efficiency further complicate the picture. The emergence of highly efficient open-source models and Chinese-developed systems, including DeepSeek, generated substantial discussion regarding the future economics of AI development. Some commentators interpreted these advances as evidence that demand for computational infrastructure might decline because comparable performance could potentially be achieved using fewer resources. However, this interpretation may misunderstand the historical relationship between efficiency improvements and resource consumption (Sinha et al., 2025).

The phenomenon known as the Jevons Paradox suggests that improvements in efficiency frequently increase rather than decrease total resource usage. Originally formulated in the context of coal consumption during the nineteenth century, the principle proposes that reductions in the cost of using a resource often expand demand sufficiently to increase overall consumption. In the context of artificial intelligence, more efficient models reduce deployment costs, enabling new applications, expanding accessibility, and encouraging broader adoption. Consequently, efficiency improvements may increase aggregate demand for computation even if individual tasks become less expensive to perform (Jevons, 1865; Alcott, 2005).

Historical evidence from the computing industry provides support for this interpretation. Advances in semiconductor efficiency repeatedly reduced the cost of computation over several decades, yet total computational demand continued growing rapidly. The availability of cheaper computing resources enabled entirely new

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

categories of software, internet services, cloud platforms, mobile applications, and digital products. Artificial intelligence may follow a similar trajectory, where efficiency improvements accelerate adoption rather than suppress infrastructure requirements (Alcott, 2005; IEA, 2025).

The distinction between capability improvement and economic value is nevertheless important. Continued technical progress does not automatically imply that all firms participating in the AI ecosystem will generate attractive financial returns. The dot-com era demonstrated that investors can correctly identify transformative technologies while simultaneously overestimating the profitability of individual participants. Therefore, evidence supporting continued model improvement should be interpreted as strengthening the technological foundation of the AI cycle rather than guaranteeing favorable outcomes for every company involved (Ofek & Richardson, 2003; Lansing, 2008).

Taken together, the available evidence suggests that the first pillar of the AI investment thesis remains substantially intact. Research on scaling laws, empirical improvements in frontier models, technology-adoption patterns, and the historical relationship between efficiency and demand all indicate that continued progress remains plausible. Although uncertainty persists regarding the pace and ultimate limits of advancement, current expectations are supported by a stronger empirical foundation than many of the assumptions that characterized internet-related investment during the late 1990s (Kaplan et al., 2020; Hoffmann et al., 2022; Brynjolfsson & McElheran, 2016).

5.2 Data Center Investment May Be Structurally Justified Under Current Conditions

The second pillar of the contemporary AI investment thesis is the belief that large-scale investment in computational infrastructure may be structurally justified under current conditions of growing AI adoption, rising inference demand, and expanding enterprise deployment. While future demand remains uncertain, supporters of the current investment cycle argue that infrastructure expansion is supported by observable market activity rather than purely speculative expectations.. Unlike many technology booms in which capital is directed primarily toward anticipated future markets, a significant portion of current AI-related spending is being driven by existing demand for cloud computing services, AI inference workloads, model training, and enterprise deployment (IEA, 2025; Goldman Sachs, 2025).

Since the public emergence of generative AI, major technology firms have announced unprecedented capital expenditure programs focused on expanding data-center capacity. Microsoft, Amazon, Alphabet, and Meta have collectively committed hundreds of billions of dollars toward infrastructure development, including advanced semiconductor procurement, power generation agreements, networking equipment, cooling systems, and specialized AI facilities (Goldman Sachs, 2025). These investments represent one of the largest infrastructure buildouts in the history of the digital economy.

Critics often compare this spending wave to the telecommunications boom of the late 1990s, arguing that excessive infrastructure investment may once again create substantial overcapacity. At first glance, the comparison appears reasonable. Both periods involve large capital commitments, optimistic growth projections, and expectations regarding transformative technological change. However, important structural differences exist between the two cases (Couper et al., 2003; Ofek & Richardson, 2003).

During the dot-com era, telecommunications firms frequently expanded network infrastructure based on projected future demand that had not yet materialized. In many cases, companies financed these projects through debt and speculative capital while generating limited operating cash flow. The result was a substantial mismatch between infrastructure deployment and realized utilization, contributing to the phenomenon of dark fiber and significant financial losses throughout the sector (Couper et al., 2003; Lansing, 2008).

The contemporary AI ecosystem differs in several important respects. First, many of the largest infrastructure investments are being undertaken by highly profitable incumbent firms with established revenue streams. Cloud-computing platforms already serve millions of customers and generate substantial recurring revenue independent of future AI growth. As a result, current infrastructure expansion is occurring within firms that possess significantly stronger financial foundations than many of the telecommunications companies that dominated the late 1990s investment cycle (Goldman Sachs, 2025).

Second, demand for computational resources is not purely hypothetical. AI model training, inference services, cloud-hosted applications, recommendation systems, software development tools, and enterprise automation platforms are already consuming large quantities of computational capacity. Industry reports consistently indicate that demand for advanced graphics processing units (GPUs) and AI accelerators has exceeded available supply across multiple periods since 2023, suggesting that infrastructure constraints are already influencing deployment decisions (Goldman Sachs, 2025; IEA, 2025).

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

The economics of AI infrastructure also differ from those of earlier internet investments because computational demand scales directly with usage. Each user interaction with a large language model requires real-time processing resources, unlike many traditional software products whose marginal costs approach zero after deployment. As AI applications become increasingly integrated into enterprise workflows and consumer products, demand for inference capacity may continue growing alongside adoption rates (Kaplan et al., 2020; McElheran et al., 2023).

Another important consideration involves the distinction between training and inference workloads. Public discussions frequently focus on the enormous costs associated with training frontier models. However, long-term infrastructure demand may be driven more heavily by inference—the process of serving AI-generated outputs to users. As adoption expands, aggregate inference demand could eventually exceed training demand by a substantial margin, requiring continued investment in computational infrastructure even if frontier-model development slows (IEA, 2025; Goldman Sachs, 2025).

Recent advances in model efficiency have prompted renewed debate regarding future infrastructure requirements. Some observers argue that improvements in algorithmic efficiency could reduce the need for additional data-center construction. While efficiency gains undoubtedly lower the computational requirements associated with individual tasks, historical evidence suggests that reductions in cost often stimulate additional demand. This dynamic is consistent with the Jevons Paradox and has repeatedly appeared throughout the history of computing, where lower computational costs enabled entirely new categories of applications and services (Jevons, 1865; Alcott, 2005).

The emergence of highly efficient models such as DeepSeek illustrates this complexity. Rather than eliminating demand for infrastructure, more efficient systems may expand the range of economically viable AI applications. Tasks that were previously too expensive to automate may become profitable, encouraging broader deployment across industries. Under such conditions, efficiency improvements can increase total infrastructure utilization even while reducing the computational cost of individual operations (Sinha et al., 2025).

Nevertheless, infrastructure investment is not without risk. Forecasts regarding future demand remain uncertain, and the possibility of temporary overcapacity cannot be dismissed. History demonstrates that infrastructure buildouts frequently overshoot short-term requirements before eventually becoming fully utilized. The existence of future demand therefore does not guarantee that every infrastructure investment will generate attractive returns within expected time horizons (Couper et al., 2003; Lansing, 2008).

The most important difference from the dot-com era may therefore be the nature of the uncertainty itself. During the late 1990s, investors questioned whether sufficient internet demand would emerge to justify infrastructure deployment. In the AI era, the more pressing question increasingly concerns whether physical infrastructure can expand rapidly enough to accommodate growing demand. Evidence from cloud markets, enterprise adoption, and semiconductor supply chains suggests that computational demand is already substantial and continues to grow. The central uncertainty is no longer whether demand exists, but whether supporting infrastructure can scale at the required pace (IEA, 2025; Muro et al., 2025; Goldman Sachs, 2025).

Taken together, the available evidence suggests that large-scale infrastructure investment is supported by stronger economic foundations than those that characterized many telecommunications projects during the dot-com era. Although risks of overinvestment remain, current spending is being driven by observable demand, existing revenue streams, and measurable computational requirements rather than by purely speculative expectations regarding future internet adoption (Goldman Sachs, 2025; IEA, 2025; McElheran et al., 2023).

5.3 User Adoption Is Widespread and Potentially Durable

The third pillar of the contemporary AI investment thesis is the belief that adoption is widespread and may prove durable over time. While current adoption levels are observable and measurable, the long-term persistence of usage patterns, retention rates, and monetization outcomes remains subject to ongoing evaluation.. Unlike speculative expectations regarding future demand, adoption can be observed directly through user behavior, enterprise deployment, software integration, and market usage patterns. Consequently, the strength of this pillar depends less on forecasts and more on empirical evidence concerning how individuals and organizations are currently utilizing AI technologies.

The rapid adoption of generative AI has few historical precedents. Following its public release in November 2022, ChatGPT became one of the fastest-growing consumer applications ever recorded, reaching approximately 100 million monthly active users within a matter of months. This rate of adoption substantially exceeded that observed for many earlier digital platforms, including major social media networks and internet services (Omnibound, 2026). While rapid adoption alone does not guarantee long-term economic value, it provides evidence that users perceive immediate utility from the technology.

Consumer adoption represents only part of the broader picture. More significant from an economic perspective is the growing integration of AI systems into enterprise operations. Organizations increasingly employ AI tools for software development, customer service, content generation, data analysis, document processing, research assistance, and workflow automation. The expansion of enterprise deployments suggests that AI adoption is not confined to experimentation but is gradually becoming embedded within routine business processes (McElheran et al., 2023; McElheran et al., 2026).

Research on digital transformation provides additional context for interpreting these developments. Brynjolfsson and McElheran (2016) argue that technologies capable of generating measurable productivity improvements often diffuse rapidly once firms identify practical applications and implementation pathways. Historically, adoption tends to accelerate when technologies move beyond novelty and begin delivering operational benefits. The increasing integration of AI into professional workflows is broadly consistent with this pattern of technology diffusion (Brynjolfsson & McElheran, 2016).

McElheran et al. (2023) found that adoption rates are influenced by factors such as organizational readiness, digital infrastructure, workforce capabilities, and access to complementary technologies. This finding is significant because it indicates that AI adoption depends not only on the availability of models but also on the capacity of firms to integrate them effectively into existing operations.

5.3.1 General-Purpose Technologies and the Productivity Paradox

The pattern of adoption observed in the contemporary AI cycle is broadly consistent with the behavior of what economists have termed general-purpose technologies — innovations characterized by pervasiveness across sectors, continuous improvement over time, and the capacity to generate complementary innovations that amplify their economic effects (Bresnahan & Trajtenberg, 1995). Electricity, the internal combustion engine, and information technology are canonical examples. Each generated transformative long-run economic effects while simultaneously producing extended periods during which aggregate productivity gains remained difficult to measure.

This temporal disconnect between deployment and measurable economic impact is central to what Brynjolfsson (1993) identified as the productivity paradox: the observation that periods of intensive investment in information technology have frequently coincided with stagnant or declining measured productivity growth at the macroeconomic level. Brynjolfsson and Hitt (2000) later demonstrated that this

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

paradox was not evidence of technological failure but rather reflected the time required for firms to reorganize processes, retrain workers, and develop the complementary organizational capabilities needed to extract value from new technologies. Productivity gains, in this account, are not instantaneous — they are contingent on a broader ecosystem of adaptation that unfolds over years or decades.

The implications for evaluating current AI adoption are significant. The absence of large AI-driven productivity gains in aggregate economic statistics at this stage of the cycle is neither surprising nor necessarily indicative of limited economic potential. Rogers (2003) similarly notes that the diffusion of innovations follows an S-shaped adoption curve in which early adoption is often concentrated among technologically ready organizations, with broader diffusion — and its associated productivity effects — occurring only as complementary infrastructure, skills, and institutional arrangements develop more widely.

What distinguishes the AI case from prior GPT episodes is the speed of initial adoption, which has exceeded that observed for personal computing and internet technologies (Omnibound, 2026; McElheran et al., 2023). Whether this accelerated diffusion will compress the lag between deployment and aggregate productivity impact, or whether organizational adaptation will remain the binding constraint regardless of adoption speed, is one of the most consequential open questions in the current debate.

5.3.2 Empirical Evidence on AI Productivity

While widespread adoption provides evidence that users perceive value in AI systems, adoption alone does not demonstrate productivity gains, profitability, revenue generation, or long-term economic sustainability. These dimensions must be evaluated separately.

Recent empirical research nevertheless suggests that AI systems can produce measurable productivity gains across a range of professional activities. In customer-service environments, Brynjolfsson, Li, and Raymond (2023) found that generative AI assistance increased worker productivity, with particularly strong benefits among less experienced employees. Their findings indicate that AI can improve both efficiency and service quality under specific organizational conditions.

Evidence from knowledge-intensive occupations has produced similar results. Experimental and field studies have reported productivity improvements in writing, content generation, administrative support, and software development tasks, particularly when AI tools are used as complements rather than substitutes for human expertise (Brynjolfsson et al., 2023; Noy & Zhang, 2023).

Research in scientific and technical work has also produced promising results. Toner-Rodgers (2024) found that AI-assisted researchers achieved substantial increases in research output and innovation performance within certain scientific environments. However, the benefits were not distributed uniformly across all workers, highlighting the importance of organizational context, task characteristics, and complementary human capabilities.

Despite these encouraging findings, significant uncertainty remains regarding the translation of localized productivity gains into broader macroeconomic outcomes. Historically, many transformative technologies generated measurable benefits at the firm level long before their effects became visible in aggregate productivity statistics. Consequently, current evidence supports cautious optimism regarding AI's economic potential, while also recognizing that the long-term magnitude of productivity improvements remains an open empirical question.

For this reason, widespread adoption should not be treated as synonymous with economic success. Instead, adoption metrics should be interpreted alongside evidence regarding productivity, retention, willingness to pay, revenue generation, and profitability. Evaluating these dimensions together provides a more complete basis for assessing whether the current AI investment cycle is supported by durable economic fundamentals.

Table 5. Empirical Studies on AI Productivity Effects

Study	Context	Main Finding
Brynjolfsson, Li & Raymond (2023)	Customer-service operations	AI assistance increased worker productivity, particularly among less experienced employees.
Noy & Zhang (2023)	Professional writing tasks	Generative AI significantly improved task completion speed and quality.
Toner-Rodgers (2024)	Scientific research	AI-assisted researchers achieved higher research output and innovation performance.

McElheran et al. (2023)	Enterprise adoption	Adoption depends on organizational readiness and complementary investments.

Source: Elaborated by the authors based on Brynjolfsson et al. (2023), Noy and Zhang (2023), Toner-Rodgers (2024), and McElheran et al. (2023).

Productivity gains, however, should not be interpreted as equivalent to revenue generation, profitability, or sustainable monetization. Organizations may experience measurable efficiency improvements while still facing challenges related to pricing, competitive pressures, infrastructure costs, and uncertain returns on investment. Similarly, user adoption does not necessarily imply willingness to pay for AI services at levels sufficient to support long-term business sustainability. Consequently, adoption, productivity, revenue generation, and profitability should be treated as analytically distinct dimensions when evaluating the economic foundations of the current AI investment cycle.

At the same time, adoption should not be confused with full economic transformation. Many organizations remain in the early stages of implementation, and substantial variation exists across industries, regions, and firm sizes. Historical experience demonstrates that productivity gains from general-purpose technologies often emerge gradually rather than immediately. Electrification, personal computing, and internet technologies all required extended periods of organizational adaptation before their full economic benefits became visible in aggregate productivity statistics (Brynjolfsson & McElheran, 2016; Lansing, 2008).

This observation has important implications for evaluating current AI investments. Critics frequently argue that measurable productivity gains remain insufficient to justify prevailing market valuations. While this concern deserves consideration, the absence of immediate economy-wide productivity acceleration does not necessarily imply that adoption lacks substance. Technology diffusion literature consistently shows that organizational learning, process redesign, workforce adaptation, and complementary investments often precede the realization of large-scale productivity gains (Brynjolfsson & McElheran, 2016; McElheran et al., 2026).

Another relevant indicator is user retention. Many technologies experience rapid initial adoption followed by equally rapid abandonment once novelty fades. Available evidence suggests that generative AI differs from this pattern because usage increasingly reflects recurring utility rather than temporary curiosity. Consumers continue using AI systems for information retrieval, writing assistance, learning, programming, and content generation, while enterprises are incorporating AI capabilities directly into products and services. The persistence of these use cases suggests that adoption is being driven by practical value rather than purely speculative enthusiasm (McElheran et al., 2023).

The comparison with the dot-com era is particularly instructive. During the late 1990s, many internet-related firms were valued based on assumptions regarding future user adoption that had not yet materialized. In contrast, contemporary AI investment is occurring in an environment where hundreds of millions of users are already interacting with AI-powered systems and where large enterprises are actively integrating these technologies into operational workflows. The existence of observable adoption does not eliminate investment risk, but it substantially reduces uncertainty regarding whether demand exists at all (Ofek & Richardson, 2003; Lansing, 2008).

Nevertheless, risks remain. Adoption growth may slow, implementation costs may prove higher than expected, and some organizations may struggle to convert experimentation into measurable productivity improvements. Furthermore, competition among AI providers could compress margins even if overall adoption continues expanding. These risks highlight the distinction between widespread usage and profitable monetization. A technology can achieve broad adoption while still generating uneven financial outcomes across participating firms (Ofek & Richardson, 2003; Sinha et al., 2025).

Despite these uncertainties, the available evidence suggests that the third pillar of the AI investment thesis remains substantially supported. Consumer adoption has occurred at unprecedented speed, enterprise deployment continues expanding, and technology-diffusion research indicates that meaningful productivity gains often emerge gradually as organizations adapt to new capabilities. Unlike the speculative expectations that characterized many internet investments during the late 1990s, current AI adoption is observable, measurable, and already integrated into a growing number of economic activities (Brynjolfsson & McElheran, 2016; McElheran et al., 2023; McElheran et al., 2026).

Taken together, the evidence examined in this section suggests that all three pillars of the contemporary AI investment thesis—continued model improvement, the necessity of infrastructure expansion, and widespread adoption—possess stronger empirical foundations than many of the assumptions that supported internet-related investment during the dot-com era. This conclusion does not imply the absence of risk, but it does suggest that comparisons between the two cycles must account for

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

important structural differences. These differences are examined in the following section.

6. Key Structural Differences Between the AI Cycle and the Dot-Com Bubble

The analysis developed throughout the previous sections suggests that comparisons between the contemporary AI investment cycle and the dot-com bubble must be approached with caution. Both periods exhibit strong investor enthusiasm, substantial infrastructure investment, elevated market valuations, and widespread expectations regarding transformative technological change. These similarities explain why comparisons between the two eras remain attractive to commentators and investors.

However, superficial similarities do not necessarily imply structural equivalence. The evidence examined thus far indicates that several of the most important characteristics underlying the current AI cycle differ materially from those that defined the internet bubble of the late 1990s. Understanding these differences is essential for evaluating whether contemporary market conditions represent a repetition of the earlier speculative episode or a fundamentally different form of technological expansion (Shiller, 2000; Lansing, 2008).

Table 6. Structural Comparison Between the Dot-Com Bubble and the AI Investment Cycle

Dimension	Dot-Com Bubble (1995–2000)	AI Investment Cycle (2022–Present)
Infrastructure Maturity	Nascent and incomplete	Mature and globally deployed
Primary Investors	Startups and speculative ventures	Large profitable incumbents
Revenue Base	Often limited or nonexistent	Existing and substantial

Adoption Status	Expected future adoption	Observable large-scale adoption
Infrastructure Bottleneck	Last-mile internet access	Energy and power availability
Capital Structure	High reliance on debt and speculation	Strong operating cash flows
Efficiency Improvements	Reduced infrastructure utilization	May increase demand through Jevons effects
Core Uncertainty	Whether demand would emerge	Whether infrastructure can scale

Source: Elaborated by the authors based on Couper et al. (2003), Ofek and Richardson (2003), Lansing (2008), McElheran et al. (2023), Goldman Sachs (2025), IEA (2025), and Muro et al. (2025).

As Table 6 illustrates, the central differences between the two periods involve infrastructure readiness, adoption dynamics, revenue generation, and the nature of the primary constraints facing continued growth.

6.1 Mature Infrastructure vs. Nascent Infrastructure

Perhaps the most important distinction concerns infrastructure maturity. During the late 1990s, investors attempted to build businesses on top of an internet ecosystem that remained incomplete in critical respects. Although backbone networks expanded rapidly, residential broadband penetration remained limited, e-commerce systems were still developing, online payment systems lacked maturity, and many consumers had only intermittent access to internet services (Couper et al., 2003; Lansing, 2008).

The contemporary AI ecosystem operates under fundamentally different conditions. Modern AI systems are deployed on top of a globally integrated digital infrastructure

JIRRS
Journal International Review of Research Studies

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

consisting of hyperscale cloud platforms, high-speed broadband networks, mature software ecosystems, digital payment systems, and billions of connected devices. The technological foundations required for widespread deployment already exist and have been continuously refined over more than two decades of internet-driven development (Kaplan et al., 2020; McElheran et al., 2023).

This distinction does not eliminate risk, but it substantially alters its nature. During the dot-com era, uncertainty centered on whether supporting infrastructure would eventually become sufficient to enable mass adoption. In the AI era, adoption is already occurring within a mature digital environment. The challenge is therefore less about creating infrastructure from scratch and more about expanding existing infrastructure to accommodate growing computational demand (IEA, 2025; Muro et al., 2025).

6.2 Incumbent Firms vs. Startup-Dominated Speculation

A second structural difference concerns the identity of the principal investors. The dot-com boom was characterized by the proliferation of newly created internet companies, many of which possessed limited operating histories, uncertain business models, and minimal revenue generation. Investors frequently valued these firms based on projected future growth rather than demonstrated economic performance (Ofek & Richardson, 2003).

The contemporary AI cycle is largely driven by established technology firms possessing substantial financial resources and diversified revenue streams. Microsoft, Alphabet, Amazon, and Meta generate hundreds of billions of dollars in annual revenue and maintain significant operating cash flows independent of future AI-related success. Consequently, much of the current infrastructure expansion is financed by firms capable of absorbing substantial investment costs while continuing to operate profitably (Goldman Sachs, 2025).

This difference does not imply that valuations cannot become excessive. However, it does mean that the financial foundations supporting current investment are generally stronger than those supporting many internet ventures during the late 1990s (Shiller, 2000; Ofek & Richardson, 2003).

6.3 Demonstrated Revenue vs. Speculative Promise

A third distinction involves revenue generation. During the internet bubble, many firms achieved extraordinary market valuations despite generating little revenue and possessing no clear path to profitability. Investors often justified these valuations through assumptions regarding future market dominance and eventual monetization (Ofek & Richardson, 2003).

The AI cycle presents a different picture. Cloud-computing services, enterprise software platforms, digital advertising networks, and subscription-based products already generate substantial revenue for many of the firms investing most heavily in AI. Although future AI-related profits remain uncertain, investment decisions are frequently being made by organizations with existing cash-generating businesses rather than by firms dependent entirely upon future expectations (Goldman Sachs, 2025; McElheran et al., 2023).

As a result, the economic viability of many current AI investments depends less on proving the existence of demand and more on determining how much value can ultimately be captured from demand that already exists.

6.4 The DeepSeek Debate and the Jevons Misinterpretation

One of the most common contemporary arguments for viewing AI as a bubble centers on recent improvements in model efficiency. The emergence of highly efficient models such as DeepSeek led some observers to argue that future demand for computational infrastructure may be significantly lower than previously expected. If comparable performance can be achieved using fewer resources, then large-scale investments in data centers and advanced hardware could potentially become difficult to justify (Sinha et al., 2025).

While superficially plausible, this argument may overlook an important historical pattern. Efficiency improvements have repeatedly expanded rather than reduced aggregate demand for computational resources. Lower costs enable new applications, increase accessibility, and encourage broader deployment. This dynamic is consistent with the Jevons Paradox, which predicts that greater efficiency often increases overall resource consumption by reducing the cost of utilization (Jevons, 1865; Alcott, 2005).

Historical developments in computing, telecommunications, and cloud services broadly support this interpretation. The cost of computation has declined dramatically over several decades, yet total computational demand has continued increasing. Artificial intelligence may follow a similar trajectory in which efficiency gains stimulate adoption and expand infrastructure requirements rather than rendering them obsolete (Alcott, 2005; IEA, 2025).

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

However, the Jevons Paradox should be understood as a plausible demand-side mechanism rather than a deterministic law. The magnitude of rebound effects varies considerably depending on the elasticity of demand for the resource in question, the availability of substitutes, prevailing price levels, regulatory constraints, and the degree of market saturation (Alcott, 2005; Sorrell, 2007). The literature distinguishes among at least three categories of rebound intensity. A weak rebound occurs when efficiency improvements increase total consumption, but by less than the reduction achieved through the efficiency gain itself — overall resource use still declines relative to a no-efficiency baseline. A strong rebound occurs when the demand response is large enough to fully offset the efficiency saving, leaving total consumption unchanged or increased. A backfire scenario — the most extreme case — occurs when demand expansion so substantially exceeds the efficiency gain that aggregate resource consumption grows faster than it would have without the improvement (Sorrell, 2007; Alcott, 2005).

In the context of artificial intelligence, the relevant question is therefore not whether some rebound effect will occur, but rather how large that effect is likely to be. If AI inference costs fall substantially, the range of economically viable applications will expand, supporting at least a weak or moderate rebound in computational demand. Whether this rebound will be strong enough to drive aggregate infrastructure requirements higher than current projections — or whether market saturation, enterprise budget constraints, or regulatory intervention will dampen demand expansion — remains an open empirical question. Consequently, the argument advanced in this article is not that Jevons effects will necessarily sustain infrastructure demand at current projected levels, but that efficiency improvements alone cannot be treated as a reliable mechanism for reducing aggregate computational requirements without accounting for demand-side responses (Alcott, 2005; Sorrell, 2007; IEA, 2025).

6.5 Counterarguments and Remaining Risks

The structural differences identified throughout this article do not imply that current AI valuations are necessarily justified, nor do they eliminate the possibility of significant capital destruction within specific segments of the ecosystem. A balanced analysis must therefore consider the strongest counterarguments advanced by critics of the current investment cycle.

The first concerns valuation concentration. A substantial share of AI-related market gains has been concentrated in a relatively small number of firms, particularly semiconductor manufacturers and hyperscale cloud providers. Critics argue that current valuations implicitly assume sustained growth in AI-related spending over many years. If adoption, monetization, or productivity gains fail to meet expectations, significant valuation corrections may occur even if the underlying technology remains economically valuable (Ofek & Richardson, 2003; Sinha et al., 2025).

A second concern involves the relationship between infrastructure investment and realized economic returns. Although major technology firms currently generate substantial revenue, capital expenditures are increasing rapidly. History demonstrates that even correct technological forecasts can produce poor investment outcomes when infrastructure deployment significantly outpaces realized demand (Couper et al., 2003; Lansing, 2008).

A third risk concerns commoditization. Open-source models, declining inference costs, and increasingly accessible AI capabilities may reduce barriers to entry over time. If model performance converges while costs continue falling, economic value may migrate toward firms possessing proprietary data, distribution channels, established customer relationships, or specialized industry expertise rather than toward model developers themselves (McElheran et al., 2023; Sinha et al., 2025).

Finally, adoption metrics require careful interpretation. While AI usage has expanded rapidly, widespread experimentation does not automatically translate into deep organizational transformation. Research suggests that successful implementation depends heavily upon complementary investments in infrastructure, workforce training, organizational processes, and digital capabilities (Brynjolfsson & McElheran, 2016; McElheran et al., 2023).

6.6 Within-Cycle Heterogeneity: Avoiding Asymmetric Comparison

A recurring limitation in comparisons between the dot-com era and the contemporary AI cycle is the tendency to contrast the weakest participants of the earlier period with the strongest participants of the present one. Such asymmetry overstates the structural differences between the two cycles and understates the risks present within the current AI ecosystem.

The dot-com era was not composed exclusively of undercapitalized startups with implausible business models. Established incumbents such as Cisco, Intel, IBM, and Microsoft were also active participants in the investment cycle of the late 1990s, possessed substantial revenues and balance sheets, and were not immune to valuation overextension or subsequent correction (Ofek & Richardson, 2003; Lansing, 2008). Similarly, telecommunications firms such as AT&T and WorldCom were large, established organizations whose infrastructure investments nonetheless

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

proved financially unsustainable. The existence of profitable incumbents during the dot-com era did not prevent significant capital destruction across the sector.

Conversely, the contemporary AI cycle is not composed exclusively of well-capitalized hyperscale incumbents. A large and growing population of AI startups currently operates with limited revenue, uncertain monetization pathways, and valuations that depend heavily upon expectations of future market capture. The concentration of investment in early-stage AI application companies, foundation model challengers, and infrastructure startups reproduces structural features more analogous to the speculative segment of the dot-com cycle than to its incumbent-led component (Sinha et al., 2025; Ofek & Richardson, 2003).

A more analytically rigorous comparison therefore requires distinguishing among participant categories within each cycle rather than treating either period as internally uniform. Table 7 presents this decomposition.

Table 7. Within-Cycle Participant Comparison: Dot-Com Era vs. AI Cycle

Participant Category	Dot-Com Era (1995–2000)	AI Cycle (2022–Present)
Large profitable incumbents	Cisco, Intel, Microsoft, IBM — substantial revenues but exposed to valuation excess	Microsoft, Alphabet, Amazon, Meta — large revenues; AI investments financed by diversified cash flows
Telecommunications / infrastructure firms	AT&T, WorldCom, Global Crossing — large firms whose infrastructure overbuild proved unsustainable	Hyperscale data-center operators and energy infrastructure developers — exposed to execution risk and permitting delays
Speculative startups	Pets.com, Webvan, eToys — minimal revenue, no path to profitability, reliant on narrative-driven capital	AI application startups and foundation model challengers — limited revenue, uncertain monetization, high dependence on continued investor confidence
Platform survivors	Amazon, eBay — survived correction and became dominant long-	Uncertain — no equivalent correction has yet occurred to identify

JIRRS
Journal International Review of Research Studies

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

	term platforms	which firms will prove durable

Source: Elaborated by the authors based on Ofek and Richardson (2003), Lansing (2008), and Goldman Sachs (2025).

This decomposition reveals that the structural argument advanced in this article applies most strongly to the incumbent segment of the AI ecosystem — firms with established revenues, diversified cash flows, and the financial capacity to absorb substantial infrastructure investment. It applies considerably less to the speculative startup segment, where conditions more closely resemble those that characterized the most vulnerable participants of the dot-com era. Evaluating the AI cycle as a whole therefore requires recognizing that it simultaneously contains segments with materially different risk profiles, and that aggregate assessments necessarily obscure this internal variation (Ofek & Richardson, 2003; Shiller, 2000).

Taken together, these considerations suggest that the strongest argument against the view that AI represents an imminent repetition of the dot-com collapse is not that risks are absent, but that the risks differ fundamentally from those that characterized the late 1990s. The central uncertainty facing the AI ecosystem is no longer whether digital demand exists. Instead, it concerns whether economic value creation, infrastructure expansion, and energy availability can evolve rapidly enough to support continued growth. This final issue—the emerging energy bottleneck—is examined in the following section (IEA, 2025; Muro et al., 2025).

6.7 Summary: Evidence For and Against the Dot-Com Analogy

Table 8 synthesizes the principal evidence for and against the dot-com analogy across the dimensions examined throughout this review. Rather than resolving the comparison definitively, the table is intended to map areas of relative consensus, identify dimensions where competing interpretations remain plausible, and highlight the indicators most likely to clarify outstanding uncertainties as the AI investment cycle continues to evolve.

Table 8. Synthesis of Evidence For and Against the Dot-Com Analogy

Dimension	Evidence Supporting the Dot-Com Analogy	Evidence Weakening the Dot-Com Analogy	Unresolved Questions	Indicators to Monitor

Valuation dynamics	Elevated market valuations concentrated in a small number of firms; expectations of sustained growth over many years (Ofek & Richardson, 2003; Sinha et al., 2025)	Primary investors are profitable incumbents with diversified revenue streams, not speculative startups (Goldman Sachs, 2025)	Whether current valuations are justified by future cash flows or reflect narrative-driven excess	Price-to-earnings ratios; AI-specific revenue growth among leading firms
Infrastructure investment	Large-scale capital expenditure with uncertain long-term returns; historical precedent of infrastructure overbuild (Couper et al., 2003; Lansing, 2008)	Spending driven by observable demand rather than speculative future adoption; GPU supply constraints already influencing deployment (IEA, 2025; Goldman Sachs, 2025)	Whether current capex will generate returns within expected time horizons; risk of temporary overcapacity	Data-center utilization rates; inference cost trends; capex-to-revenue ratios
User adoption	Rapid initial adoption may not translate into durable retention or monetization; some usage subsidized by platform-level investment	Adoption is observable and measurable at scale; enterprise integration is deepening across multiple industries (McElheran et	Whether adoption will persist after novelty effects; whether willingness to pay is sufficient for long-term sustainability	User retention metrics; enterprise renewal rates; AI-specific subscription revenue

		al., 2023; McElheran et al., 2026)		
Productivity and monetization	Aggregate productivity gains not yet visible in macroeconomic statistics; localized gains may not translate into economy-wide effects (Brynjolfsson & McElheran, 2016)	Experimental and firm-level studies show measurable productivity improvements in specific tasks and occupations (Brynjolfsson et al., 2023; Noy & Zhang, 2023; Toner-Rodgers, 2024)	Long-run magnitude of AI productivity effects remains an open empirical question	Firm-level productivity surveys; AI-related output per worker; enterprise ROI disclosures
Business model sustainability	Commoditization of models may compress margins; open-source competition may reduce pricing power (Sinha et al., 2025)	Firms with proprietary data, distribution, and customer relationships retain structural advantages independent of model access	Whether competitive equilibrium will favor model developers or application-layer firms	Operating margins of AI firms; market share concentration; open-source adoption rates
Energy and physical infrastructure	Growing electricity demand may outpace grid expansion, creating deployment bottlenecks analogous to the last-mile problem (IEA, 2025; Muro et	Unlike dot-com infrastructure, energy constraints reflect demand pressure rather than demand absence — a structurally different form	Whether grid expansion and energy investment will keep pace with computational demand	Electricity prices in data-center clusters; permitting timelines; power purchase agreement terms

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

	al., 2025)	of risk		
Technological progress	Scaling laws may face diminishing returns; benchmark saturation may reduce the informativeness of capability claims (Hoffmann et al., 2022)	Empirical improvements across successive model generations are observable and consistent with predicted scaling behavior (Kaplan et al., 2020)	Whether current scaling trajectories will continue, plateau, or require architectural breakthroughs	Benchmark performance across model generations; training cost per capability unit

Source: Elaborated by the authors based on Shiller (2000), Ofek and Richardson (2003), Lansing (2008), Couper et al. (2003), Brynjolfsson et al. (2023), Noy and Zhang (2023), Toner-Rodgers (2024), McElheran et al. (2023), Goldman Sachs (2025), IEA (2025) and Muro et al. (2025).

7. The Critical Bottleneck: Energy Infrastructure

The preceding sections argued that the contemporary AI investment cycle differs from the dot-com bubble in several important respects. Infrastructure is more mature, adoption is already occurring at scale, and many of the largest investments are being undertaken by profitable incumbent firms rather than speculative startups. However, these differences do not imply the absence of constraints. Instead, they suggest that the primary bottleneck facing AI has shifted from digital connectivity to physical infrastructure.

During the late 1990s, one of the most significant limitations on internet adoption was the so-called “last-mile problem.” Telecommunications companies expanded backbone capacity rapidly, yet many households and businesses lacked the broadband connectivity required to fully utilize the growing network. As a result, infrastructure deployment frequently outpaced realized demand, contributing to the emergence of excess capacity and financial losses throughout the sector (Couper et al., 2003; Lansing, 2008).

The AI ecosystem faces a comparable challenge today, but the bottleneck is located within the energy system rather than telecommunications networks. Modern AI models operate on highly specialized computational infrastructure housed within data centers that require substantial and continuously available electricity supplies. Consequently, the ability of the AI sector to continue expanding increasingly depends upon the capacity of electrical grids, generation facilities, transmission systems, and supporting energy infrastructure rather than on internet connectivity itself (IEA, 2025).

7.1 AI and the Growth of Electricity Demand

Artificial intelligence is fundamentally a physical technology. Although often discussed in terms of software and algorithms, every AI-generated response ultimately depends upon computational hardware consuming electricity within a data center. As model sizes increase and AI applications become more widespread, electricity demand grows alongside computational demand (IEA, 2025).

According to the International Energy Agency, global data-center electricity consumption reached approximately 415 TWh in 2024, representing about 1.5% of global electricity consumption. Under the agency’s base-case scenario, this figure is expected to increase to roughly 945 TWh by 2030, more than doubling within six years. This growth corresponds to approximately 15% annually, significantly exceeding the expected growth rate of overall global electricity demand (IEA, 2025).

AI is expected to be the primary driver of this expansion. The IEA estimates that accelerated servers used for AI workloads will account for nearly half of the projected increase in global data-center electricity consumption through 2030. In the United States, data centers are projected to account for nearly half of total electricity-demand growth during the remainder of the decade, highlighting the increasingly important relationship between AI development and energy infrastructure planning (IEA, 2025).

Importantly, these figures represent global averages. The impact is often far more pronounced at regional and local levels because data centers tend to cluster geographically. Large AI facilities can require power consumption comparable to that of energy-intensive industrial operations, placing substantial pressure on local grids, transmission networks, and permitting systems (IEA, 2025).

7.2 Why Efficiency Alone May Not Solve the Problem

A common argument advanced by AI skeptics is that improvements in model efficiency will eventually eliminate the need for large-scale infrastructure expansion. The emergence of increasingly efficient models has strengthened this view by

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

demonstrating that similar levels of performance can sometimes be achieved using fewer computational resources.

While efficiency improvements are unquestionably important, historical evidence suggests that their long-term effects are not always intuitive. William Stanley Jevons observed in the nineteenth century that improvements in the efficiency of coal usage often increased rather than decreased total coal consumption because lower costs encouraged broader adoption. This phenomenon became known as the Jevons Paradox and has since been observed in numerous technological and industrial contexts (Jevons, 1865; Alcott, 2005).

Artificial intelligence may exhibit similar dynamics. As models become more efficient, deployment costs decline, making AI applications economically viable in a larger number of industries and use cases. Tasks that were previously too expensive to automate may become profitable, encouraging broader adoption and increasing total computational demand. Under such conditions, efficiency improvements may reduce the cost of individual operations while simultaneously increasing aggregate infrastructure requirements (Alcott, 2005; IEA, 2025).The strength of this effect, however, is not predetermined. Whether efficiency-driven demand expansion will be sufficient to increase aggregate infrastructure requirements depends on factors including price elasticity of AI services, the emergence of new use cases, enterprise adoption capacity, and the pace of regulatory development governing AI deployment (Alcott, 2005; Sorrell, 2007).

Historical developments in computing support this interpretation. Over several decades, the cost of computation declined dramatically due to advances in semiconductor technology, yet total computational demand continued expanding. Rather than reducing the need for computing infrastructure, lower costs enabled entirely new categories of software, cloud services, digital platforms, and internet applications. AI may follow a similar trajectory in which efficiency accelerates adoption instead of reducing overall resource consumption (Alcott, 2005).

7.3 Industry Responses to Emerging Energy Constraints

The growing importance of electricity has not gone unnoticed by industry participants. Major technology firms have increasingly pursued long-term strategies designed to secure reliable energy supplies for future AI expansion. These strategies include long-term power-purchase agreements, direct investments in renewable-energy projects, support for grid modernization, and renewed interest in advanced nuclear technologies (IEA, 2025).

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

The International Energy Agency projects that renewable-energy sources will provide nearly half of the additional electricity generation required to meet growing data-center demand through 2035. Natural gas is also expected to play an important role because of its ability to provide dispatchable generation. Meanwhile, nuclear energy is projected to become increasingly significant later in the decade, particularly as governments and technology firms explore new reactor designs and small modular reactor technologies (IEA, 2025).

These developments illustrate an important distinction between the AI cycle and many previous technology booms. The challenge is no longer simply producing better software. Instead, continued expansion increasingly depends upon the construction of physical infrastructure, including generation facilities, transmission lines, substations, cooling systems, and energy-storage technologies. In this sense, AI is becoming as much an infrastructure story as a software story (IEA, 2025).

7.4 The Real Constraint Facing the AI Economy

The evidence reviewed throughout this section suggests that the most significant risk facing the AI ecosystem differs fundamentally from the risks that characterized the dot-com era. During the late 1990s, investors questioned whether internet adoption would become sufficiently widespread to justify infrastructure investments. Today, user adoption, enterprise deployment, and computational demand are already observable realities. The primary uncertainty increasingly concerns whether supporting infrastructure can expand rapidly enough to accommodate continued growth (McElheran et al., 2023; IEA, 2025).

Consequently, the central bottleneck facing AI is not a shortage of users, algorithms, or investor capital. It is the ability of energy systems to deliver reliable and affordable electricity at the scale required by future computational workloads. Just as the last-mile problem constrained the pace of internet adoption during the dot-com era, energy infrastructure may become the defining constraint shaping the future trajectory of artificial intelligence. The difference is that today's bottleneck emerges not from insufficient connectivity, but from the physical realities of power generation, transmission, and distribution (Couper et al., 2003; IEA, 2025; Muro et al., 2025).

8. Investment Implications and Risk Assessment: Bubble, Buildout, or Both?

The preceding sections suggest that the contemporary AI investment cycle differs in important ways from the dot-com bubble. Infrastructure is more mature, adoption is already occurring at scale, and many of the largest investments are being undertaken

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

by highly profitable incumbent firms. At the same time, the presence of stronger structural foundations does not imply the absence of financial risk. History demonstrates that transformative technologies can generate both substantial economic value and significant investor losses simultaneously. Railroads, electrification, telecommunications, and the internet all produced periods in which technological success coexisted with financial overexuberance (Shiller, 2000; Lansing, 2008).

Consequently, the central question facing investors, policymakers, and researchers is not whether artificial intelligence will be economically important, but rather where risks are concentrated, which segments possess the strongest structural foundations, and which indicators are most useful for evaluating the sustainability of current investment levels.

8.1 Where the Risks Are Concentrated

Although AI adoption continues to expand, risks are not distributed evenly throughout the ecosystem. The most vulnerable participants may be firms that lack meaningful technological differentiation or sustainable competitive advantages. During periods of rapid technological expansion, capital often flows toward a large number of companies pursuing similar business models. As competition intensifies, many of these firms struggle to maintain pricing power, attract customers, or establish defensible market positions (Ofek & Richardson, 2003; Shiller, 2000).

This risk is particularly relevant for startups whose products depend primarily upon access to foundation models developed by larger organizations. As AI capabilities become increasingly commoditized, differentiation may shift toward proprietary data, industry-specific expertise, customer relationships, distribution networks, and workflow integration rather than model access alone (McElheran et al., 2023; Sinha et al., 2025).

Another area of potential vulnerability involves hardware-related expectations. Semiconductor manufacturers and AI infrastructure providers have benefited substantially from the rapid expansion of AI-related capital expenditures. However, current valuations often reflect assumptions regarding sustained growth in computational demand. If adoption slows, infrastructure utilization declines, or future efficiency gains reduce hardware requirements more rapidly than expected, portions of the hardware ecosystem may face significant valuation pressure despite continued growth in AI adoption itself (Goldman Sachs, 2025; Sinha et al., 2025).

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

Infrastructure investment also carries execution risk. As discussed in the previous section, data-center expansion depends heavily upon electricity availability, transmission capacity, permitting processes, and energy infrastructure development. Delays in any of these areas may affect expected returns on infrastructure investments even if long-term demand remains strong (IEA, 2025; Muro et al., 2025).

Importantly, these risks resemble neither the complete absence of demand observed in many speculative bubbles nor the technological failure of the underlying innovation. Rather, they reflect uncertainty regarding market structure, competitive dynamics, and the distribution of future economic value among participants in the ecosystem (Ofek & Richardson, 2003; Lansing, 2008).

8.2 Where the Structural Thesis Remains Strong

While risks remain significant, some areas of the AI ecosystem appear supported by stronger structural foundations. One such area is hyperscale cloud infrastructure. Firms such as Microsoft, Amazon, Alphabet, and Meta possess substantial financial resources, existing customer bases, and diversified revenue streams that reduce dependence upon any single AI-related outcome. Their investments are frequently supported by established businesses that generate recurring cash flows independent of future AI monetization (Goldman Sachs, 2025).

A second area involves data-center and energy infrastructure. Regardless of which firms ultimately dominate the AI software market, computational workloads require physical infrastructure. The growth of AI deployment therefore creates demand for data centers, networking equipment, power generation, transmission systems, cooling technologies, and related infrastructure. While individual projects may underperform expectations, the broader requirement for physical capacity expansion appears supported by observable growth in computational demand (IEA, 2025; Muro et al., 2025).

Strategic semiconductor technologies also occupy a relatively strong position within the value chain. Although competitive dynamics may evolve, advanced AI systems remain dependent upon specialized hardware capable of supporting training and inference workloads. Consequently, semiconductor development continues to represent a critical enabling technology for the broader AI ecosystem (Kaplan et al., 2020; Goldman Sachs, 2025).

Enterprise software may represent another comparatively resilient segment. Historically, technologies that improve productivity often create lasting value when integrated directly into business processes. Organizations increasingly deploy AI

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

within customer-service platforms, software-development environments, data-analysis systems, and workflow-automation tools. Such applications derive value from recurring operational usage rather than purely speculative expectations regarding future adoption (Brynjolfsson & McElheran, 2016; McElheran et al., 2023).

8.3 Metrics Worth Monitoring

Because the AI investment cycle remains in an early stage of development, evaluating its sustainability requires attention to a range of economic and operational indicators. One of the most important metrics is the cost of inference. Long-term adoption depends not only on model capability but also on the affordability of deploying AI at scale. Declining inference costs may support broader adoption, while persistently high costs could limit deployment across cost-sensitive industries (Kaplan et al., 2020; Hoffmann et al., 2022).

Data-center utilization rates provide another important signal. If infrastructure expansion continues while utilization remains weak, concerns regarding overcapacity may become increasingly justified. Conversely, sustained capacity constraints would support the argument that infrastructure demand remains structurally strong (Goldman Sachs, 2025; IEA, 2025).

Energy-related indicators also deserve close attention. Electricity prices, transmission constraints, grid reliability, and permitting timelines increasingly influence the economics of AI deployment. Because energy infrastructure has emerged as a primary bottleneck, developments within electricity markets may become as important to the future of AI as advances in algorithms or hardware (IEA, 2025; Muro et al., 2025).

On the demand side, enterprise adoption rates, user-retention metrics, and AI-related revenue growth provide valuable evidence regarding the durability of current usage patterns. Rapid adoption followed by declining engagement would support more skeptical interpretations, whereas sustained utilization and expanding commercial integration would strengthen the case for continued investment (McElheran et al., 2023; McElheran et al., 2026).

Finally, investors should monitor the operating margins of major AI participants. Technological adoption alone does not guarantee profitability. The long-term sustainability of the ecosystem depends upon whether firms can convert growing usage into durable economic returns while managing the substantial costs associated with iSinha et al., 2025).

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

8.4 Bubble, Buildout, or Both?

The evidence reviewed throughout this article suggests that framing the AI cycle as either a pure speculative bubble or a purely rational investment boom presents a false dichotomy. History indicates that transformative technologies frequently generate both speculative excess and genuine infrastructure development simultaneously. Railroads, electrification, and the internet all experienced periods in which investor enthusiasm exceeded short-term economic reality, yet the infrastructure constructed during those periods ultimately became foundational to future economic growth (Shiller, 2000; Lansing, 2008).

Artificial intelligence may follow a similar pattern. Elevated valuations, concentrated market expectations, and substantial capital expenditures create conditions in which financial corrections remain entirely possible. At the same time, observable adoption, measurable technological progress, and growing infrastructure demand suggest that the underlying technological thesis possesses stronger foundations than many of the assumptions that characterized the late stages of the dot-com bubble (Kaplan et al., 2020; McElheran et al., 2023).

Consequently, the most plausible interpretation may be that the current cycle contains elements of both a buildout and a speculative episode. Some firms may ultimately prove overvalued, some investments may fail to generate expected returns, and portions of the ecosystem may experience significant corrections. Yet these outcomes would not necessarily invalidate the broader economic importance of artificial intelligence. As occurred after the dot-com collapse, financial disappointment and technological success can coexist within the same historical process (Ofek & Richardson, 2003; Shiller, 2000).

The critical distinction is that the central uncertainty no longer concerns whether AI will be adopted. Adoption is already occurring. Instead, the key questions involve the pace of value creation, the distribution of economic returns, and the ability of infrastructure—particularly energy infrastructure—to support continued expansion. These considerations form the basis for the final conclusions presented in the following section (IEA, 2025; Muro et al., 2025).

8.5 Boundary Conditions: When the Thesis Would Weaken

The argument advanced in this article — that the AI investment cycle rests upon stronger structural foundations than the dot-com bubble — is conditional rather than unconditional. Several developments could materially weaken this thesis and bring the AI cycle closer to the dot-com analogy than current evidence suggests.

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

The thesis would weaken significantly if AI-related revenue growth among major incumbents decelerates sharply while capital expenditures continue rising, producing a sustained divergence between investment and realized returns analogous to the capex-to-revenue mismatches observed in telecommunications during the late 1990s (Couper et al., 2003). It would also weaken if enterprise adoption rates plateau before generating measurable productivity gains at scale, suggesting that current usage reflects experimentation rather than durable organizational integration (Brynjolfsson & McElheran, 2016; McElheran et al., 2023).

On the infrastructure side, severe and prolonged energy constraints — including grid bottlenecks, permitting failures, or sharp increases in electricity costs — could suppress deployment below projected levels, creating underutilization of installed data-center capacity and reproducing the dark-fiber dynamic of the dot-com era (IEA, 2025; Muro et al., 2025). Finally, rapid commoditization of model capabilities combined with persistent negative margins among AI-focused firms would challenge the monetization assumptions underlying current valuations (Sinha et al., 2025; Ofek & Richardson, 2003).

Monitoring the indicators identified in Section 8.3 provides the most direct means of assessing whether these boundary conditions are being approached.

9. Conclusion

The rapid expansion of artificial intelligence has generated intense debate regarding whether current market conditions represent a technological revolution, a speculative bubble, or some combination of both. Comparisons with the dot-com bubble have become increasingly common as capital expenditures rise, market valuations expand, and investor enthusiasm reaches historically elevated levels. Given the profound economic implications of this question, distinguishing between superficial similarities and structural realities is essential.

This article examined the relationship between the contemporary AI investment cycle and the dot-com bubble through a narrative review of academic literature, institutional reports, and historical evidence. The analysis began with a retrospective examination of the internet boom of the late 1990s and the theoretical characteristics commonly associated with speculative bubbles. These historical and conceptual foundations provided a framework for evaluating the contemporary AI ecosystem.

The findings suggest that the current AI cycle shares several characteristics with previous periods of technological exuberance. Elevated expectations, substantial infrastructure investment, concentrated market enthusiasm, and ambitious growth projections are all present. These conditions create the possibility of valuation excesses, capital misallocation, and future market corrections, particularly among firms lacking sustainable competitive advantages or clear monetization pathways (Shiller, 2000; Ofek & Richardson, 2003).

However, the analysis also identified important structural differences between the AI cycle and the dot-com bubble. Unlike many internet firms of the late 1990s, the primary investors in AI infrastructure are largely established technology companies possessing substantial revenue streams, strong balance sheets, and significant operating cash flows. Furthermore, AI adoption is already observable across consumer, enterprise, and industrial applications rather than remaining dependent upon uncertain future demand. These distinctions suggest that the economic foundations supporting current investment are generally stronger than those that characterized many internet-related ventures during the dot-com era (Goldman Sachs, 2025; McElheran et al., 2023).

The evaluation of the three pillars underlying the contemporary AI investment thesis further supports this conclusion. Evidence from scaling-law research indicates that continued improvements in model performance remain plausible, although future progress may become increasingly expensive and technically challenging (Kaplan et al., 2020; Hoffmann et al., 2022). Infrastructure investment appears supported by observable computational demand rather than purely speculative expectations (IEA, 2025; Goldman Sachs, 2025). Finally, adoption patterns demonstrate that AI technologies are already being integrated into a growing range of economic activities, providing stronger empirical evidence of utility than was available during many stages of the internet bubble (Brynjolfsson & McElheran, 2016; McElheran et al., 2023).

One of the most significant findings emerging from this analysis concerns the changing nature of infrastructure constraints. During the dot-com era, the principal bottleneck involved internet connectivity and last-mile access. In contrast, the contemporary AI ecosystem increasingly faces constraints associated with electricity generation, transmission capacity, and energy infrastructure. The evidence reviewed from the International Energy Agency and Brookings Institution suggests that energy availability may become one of the defining factors shaping the future trajectory of artificial intelligence over the coming decade (IEA, 2025; Muro et al., 2025).

At the same time, the existence of stronger structural foundations should not be interpreted as evidence that current valuations are universally justified. History demonstrates that investors can correctly identify transformative technologies while simultaneously overestimating the profitability of individual firms. The railroad boom, electrification, the internet, and other technological revolutions all produced periods in which genuine innovation coexisted with speculative excess. Artificial intelligence may follow a similar path in which long-term technological success occurs alongside episodes of financial overvaluation and periodic market corrections (Lansing, 2008; Shiller, 2000).

Consequently, the findings of this study do not support the view that artificial intelligence represents a simple repetition of the dot-com bubble. The available evidence suggests that the AI cycle is supported by stronger infrastructure, broader adoption, and more established revenue foundations than those that characterized many internet investments during the late 1990s. Nevertheless, significant uncertainties remain regarding future monetization, competitive dynamics, energy constraints, and the distribution of economic value across participants in the ecosystem (Ofek & Richardson, 2003; Sinha et al., 2025).

The most plausible interpretation is therefore that the contemporary AI cycle contains elements of both a speculative episode and a genuine infrastructure buildout. Some valuations may ultimately prove excessive, some firms may fail, and portions of the market may experience substantial corrections. Yet such outcomes would not necessarily invalidate the broader technological transformation currently underway. As occurred following the collapse of the dot-com bubble, financial disappointment and technological progress are not mutually exclusive outcomes.

Future research should continue examining the long-term relationship between AI adoption, productivity growth, infrastructure investment, and energy demand. Particular attention should be given to the economic effects of declining inference costs, the evolution of enterprise adoption patterns, and the capacity of energy systems to support continued computational expansion. These factors are likely to play a decisive role in determining whether the current AI cycle ultimately resembles a temporary speculative episode or the early stage of a broader technological transformation.

In conclusion, the evidence reviewed in this article does not support a simple one-to-one analogy between the contemporary AI investment cycle and the dot-com bubble. The two periods share surface-level similarities — elevated valuations, large infrastructure commitments, and optimistic growth narratives — but differ materially in infrastructure maturity, investor composition, revenue generation, and adoption dynamics. At the same time, the AI cycle is not free of speculative elements, and the structural advantages identified in this review apply most clearly to incumbent-led segments of the ecosystem rather than to the cycle as a whole. The most accurate characterization is therefore that the current AI cycle contains both structurally grounded and speculative components, and that the balance between them will depend on developments in monetization, energy infrastructure, and competitive dynamics that remain unresolved at the time of writing.

REFERENCES

Alcott, B. (2005). Jevons' paradox. Ecological Economics, 54(1), 9–21. https://doi.org/10.1016/j.ecolecon.2005.03.020

Brynjolfsson, E., & McElheran, K. (2016). The rapid adoption of data-driven decision-making. American Economic Review, 106(5), 133–139. https://doi.org/10.1257/aer.p20161016

CNBC. (2025, October 31). How much Google, Meta, Amazon and Microsoft are spending on AI. CNBC Technology. https://www.cnbc.com/2025/10/31/tech-ai-google-meta-amazon-microsoft-spend.html

Couper, E. A., Hejkal, J. P., & Wolman, A. L. (2003). Boom and bust in telecommunications. Economic Quarterly, Federal Reserve Bank of Richmond, 89(4), 1–24. https://www.richmondfed.org/publications/research/economic_quarterly/2003/fall/couperhejkalwolman

Dedehayir, O., & Steinert, M. (2016). The hype cycle model: A review and future directions. Technological Forecasting and Social Change, 108, 28–41. https://doi.org/10.1016/j.techfore.2016.04.005

DeepSeek-AI. (2025). DeepSeek-R1: Incentivizing reasoning capability in LLMs via reinforcement learning. arXiv preprint arXiv:2501.12948. https://arxiv.org/abs/2501.12948

Fenn, J., & Raskino, M. (2008). Mastering the hype cycle: How to choose the right innovation at the right time. Harvard Business Press.

Fortune. (2025, September 28). Everyone's wondering if, and when, the AI bubble will pop: Here's what went down 25 years ago. Fortune Magazine. https://fortune.com/2025/09/28/ai-dot-com-bubble-parallels-history-explained-companies-revenue-infrastructure/

Global Electricity. (2025). Data centers and AI energy consumption: The surge in electricity demand. https://www.globalelectricity.org/data-centers-energy-consumption/

Goldman Sachs. (2019). The late 1990s dot-com bubble implodes in 2000. Goldman Sachs Historical Moments. https://www.goldmansachs.com/our-firm/history/moments/2000-dot-com-bubble

Hoffmann, J., Borgeaud, S., Mensch, A., Buchatskaya, E., Cai, T., Rutherford, E., de Las Casas, D., Hendricks, L. A., Welbl, J., Clark, A., Hennigan, T., Noland, E., Millican, K., van den Driessche, G., Damoc, B., Guy, A., Osindero, S., Simonyan, K., Elsen, E., Rae, J. W., Vinyals, O., & Sifre, L. (2022). Training compute-optimal large language models. In Proceedings of the 36th Conference on Neural Information Processing Systems (NeurIPS 2022). https://doi.org/10.48550/arXiv.2203.15556

International Energy Agency (IEA). (2025). Energy and AI. International Energy Agency. https://www.iea.org/reports/energy-and-ai

International Energy Agency (IEA). (2025). Energy demand from AI. In Energy and AI. International Energy Agency. https://www.iea.org/reports/energy-and-ai/energy-demand-from-ai

International Energy Agency (IEA). (2025). Energy supply for AI. In Energy and AI. International Energy Agency. https://www.iea.org/reports/energy-and-ai/energy-supply-for-ai

International Energy Agency (IEA). (2026). Key questions on energy and AI: Executive summary. International Energy Agency. https://www.iea.org/reports/key-questions-on-energy-and-ai/executive-summary

ISE Magazine. (2025, July 24). The perils of irrational exuberance — the 25th anniversary of the dot-com boom. ISE Magazine. https://www.isemag.com/professional-development-leadership/article/55293922

Kaplan, J., McCandlish, S., Henighan, T., Brown, T. B., Chess, B., Child, R., Gray, S., Radford, A., Wu, J., & Amodei, D. (2020). Scaling laws for neural language models. arXiv preprint arXiv:2001.08361. https://arxiv.org/abs/2001.08361

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

Lansing, K. J. (2008). Speculative bubbles and overreaction to technological innovation. FRBSF Economic Letter, 2008(18). Federal Reserve Bank of San Francisco. https://www.frbsf.org/research-and-insights/publications/economic-letter/2008/06/speculative-bubbles-technology-innovation/

McElheran, K., Li, J. F., Brynjolfsson, E., Kroff, Z., Dinlersoz, E., Foster, L., & Zolas, N. (2023). AI adoption in America: Who, what, and where. National Bureau of Economic Research Working Paper. https://www.nber.org/papers/w31514

Minsky, H. P. (1992). The financial instability hypothesis (Working Paper No. 74). Levy Economics Institute of Bard College. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=161024

Muro, M., Maxim, R., & Whiton, J. (2025). Global energy demands within the AI regulatory landscape. Brookings Institution. https://www.brookings.edu/articles/global-energy-demands-within-the-ai-regulatory-landscape/

Ofek, E., & Richardson, M. (2003). DotCom mania: The rise and fall of internet stock prices. The Journal of Finance, 58(3), 1113–1137. https://doi.org/10.1111/1540-6261.00561

Omnibound. (2026). ChatGPT user statistics 2026: 56+ data points on users, revenue and growth. https://www.omnibound.ai/blog/chatgpt-user-statistics

Shiller, R. J. (2000). Irrational exuberance. Princeton University Press.

Shin, J. K., & Subramanian, C. (2019). Asset price bubbles and technological innovation. Economic Inquiry, 57(1), 482–497. https://doi.org/10.1111/ecin.12719

Sinha, R., Gupta, A., & Mehta, S. (2025). Brief analysis of DeepSeek R1 and its implications for generative AI. arXiv preprint arXiv:2502.02523. https://arxiv.org/abs/2502.02523

Statista. (2025). Google, Amazon, Meta, Apple, and Microsoft (GAMAM) — statistics and facts. https://www.statista.com/topics/4213/google-apple-facebook-amazon-and-microsoft-gafam/

Received: 01/04/2026 - **Accepted:** 16/06/2026
Vol: 01.08
DOI: 10.66104/x4ve3t91
Pages: 1-54

Baumeister, R. F., & Leary, M. R. (1997). Writing narrative literature reviews. Review of General Psychology, 1(3), 311–320.

Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. Health Information and Libraries Journal, 26(2), 91–108.

Mahoney, J., & Rueschemeyer, D. (Eds.). (2003). Comparative Historical Analysis in the Social Sciences. Cambridge University Press.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. Journal of Business Research, 104, 333–339.

Bresnahan, T. F., & Trajtenberg, M. (1995). General Purpose Technologies: 'Engines of Growth'? Journal of Econometrics, 65(1), 83–108.

Rogers, E. M. (2003). Diffusion of Innovations (5th ed.). Free Press.

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

Noy, S., & Zhang, W. (2023). Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence. Science, 381(6654), 187–192.

Toner-Rodgers, A. (2024). Artificial Intelligence, Scientific Discovery, and Product Innovation. Working Paper, Massachusetts Institute of Technology.

Sorrell, S. (2007). The Rebound Effect: An Assessment of the Evidence for Economy-Wide Energy Savings from Improved Energy Efficiency. UK Energy Research Centre. https://ukerc.rl.ac.uk/UCAT/PUBLICATIONS/The_Rebound_Effect_An_Assessment_of_the_Evidence_for_Economy-wide_Energy_Savings_from_Improved_Energy_Efficiency.pdf

Brynjolfsson, E. (1993). The productivity paradox of information technology. Communications of the ACM, 36(12), 66–77. https://doi.org/10.1145/163298.163309

Brynjolfsson, E., & Hitt, L. M. (2000). Beyond computation: Information technology, organizational transformation and business performance. Journal of Economic Perspectives, 14(4), 23–48. https://doi.org/10.1257/jep.14.4.23

McElheran, K., Yang, M.-J., Kroff, Z., & Brynjolfsson, E. (2026). The Adoption of Industrial AI in America. AEA Papers and Proceedings, 116, 20–25. https://doi.org/10.1257/pandp.20261033

OpenAI. (2023). GPT-4 Technical Report. arXiv preprint, arXiv:2303.08774. https://arxiv.org/abs/2303.08774

Linden, A., & Fenn, J. (2003). Understanding Gartner's Hype Cycles. Gartner Research Note. https://www.bus.umich.edu/KresgePublic/Journals/Gartner/research/115200/115274/115274.pdf

Jevons, W. S. (1865). The Coal Question: An Inquiry Concerning the Progress of the Nation, and the Probable Exhaustion of Our Coal-Mines. London: Macmillan.