Allianz Global Investors

THEMATIC PULSE | DECEMBER 2025

Tech titans go all-in:

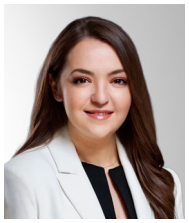
the AI investment boom

Capital expenditure from the world's largest technology companies suggests they remain utterly convinced that artificial intelligence leadership is the prize of a lifetime, a once-in-a-century opportunity that will define the next generation of corporate dominance. Microsoft, Alphabet, Meta, Amazon, and their peers, see AI not as a side project, but as the very foundation of their future business models. To them, failing to lead would mean risking irrelevance.

Yet, while the conviction is absolute, the contours of what that leadership looks like, when it will materialise, and how exactly it will translate into financial returns remain unclear. We know the race is existential, but we still don't know where the finish line lies.

That uncertainty hasn't stopped the money from flowing. In fact, the scale of investment has become staggering. The big tech firms together are expected to spend well over $400 billion[1] in 2025 and just raised their capital expenditure plans for next year with expected $500+[2] billion AI-related infrastructure alone. Microsoft's capital expenditures reached roughly $25 billion in its most recent quarter, driven primarily by AI data-centre buildouts. Alphabet raised its 2025 capex guidance to $91-93 billion, while Meta announced yet another round of datacentre and infrastructure investments for 2026. A few weeks earlier, OpenAI still privately held but now an economic force on its own signed multiple funding agreements reportedly worth tens of billions of dollars to secure its own compute capacity and develop custom chips. McKinsey estimates that global data-centre investment could reach $6.7 trillion by 2030, with roughly $5 trillion of that related to AI workloads[3]. Put simply, we are witnessing one of the largest capital-expenditure cycles in history, rivalling the scale of industrial revolutions or the global railroad build-out of the 19th century.

Nedialko Nedialkov, CFA
Portfolio Manager

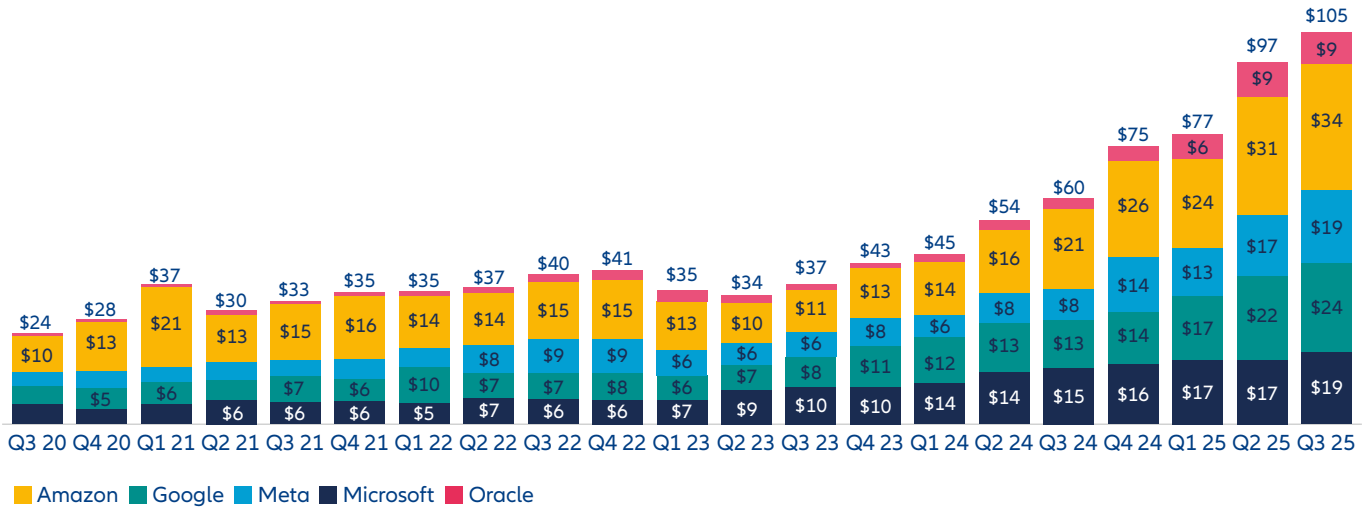


Nezhla Mehmed
Portfolio Manager

Christian Zilien, CFA
Product Specialist
Equity

Exhibit 1: Capital expenditure by major tech companies, over time

Capital expenditure by quarter (in USD billions)



Source: Artificial Analysis (**AI Model & API Providers Analysis | Artificial Analysis**)[4]

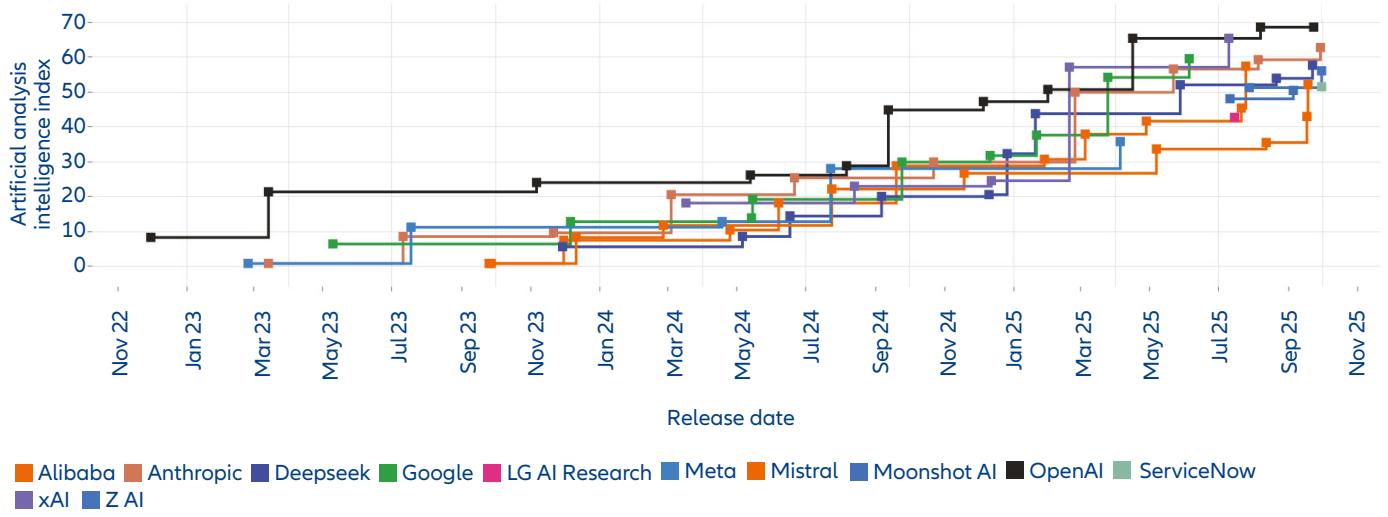
The scale of spending doesn't come without bottlenecks

But that scale of spending doesn't come without stress. The expansion of AI represents a pivotal turning point, intersecting with longstanding challenges in areas like energy, raw materials, grid infrastructure, and computing efficiency. Unlike previous booms in fibre optics or solar energy, this surge is generating real, immediate demand. Supply chains are adapting, but they remain under strain. The most visible constraint is at the silicon level. There simply aren't enough XPUs (advanced AI processors) to meet demand, limiting the number of high-end computing units that can reach data-centre builders. Downstream, the infrastructure around those chips and data centres electrical systems, cooling, and high-speed networking is also proving to be a bottleneck. The construction of next-generation data centres increasingly depends on specialist suppliers of transformers, cables, switchgear, and liquid-cooling systems.

And then there's the elephant in the room: power. The AI boom is driving an unprecedented rise in electricity consumption. Goldman Sachs projects that global data-centre power demand will rise by roughly 50% by 2027 and 165% by 2030 compared to 2023[5]. In regions like Virginia, Ohio, and parts of Texas home to dense clusters of hyperscaler facilities local grids are straining under the load. In some cases, new power-line installations are the gating factor for new data-centre projects. This is no longer just a technology challenge; it's becoming a societal one. Local communities are facing higher energy costs, water-usage conflicts, and pressure on existing infrastructure. The environmental and social footprint of AI infrastructure is rapidly becoming a real political issue.

Exhibit 2: Frontier language model intelligence, over time

Artificial Analysis Intelligence Index v3.0 incorporates 10 evaluations: MMLU-Pro, GPQA Diamond, Humanity's Last Exam, LiveCodeBench, SciCode, AIME 2025, IFBench, AA-LCR, Terminal-Bench Hard, τ^2-Bench Telecom



Source: Artificial Analysis (**AI Model & API Providers Analysis | Artificial Analysis**)[6]

The pace of innovation remains relentless

While these bottlenecks might slow deployment, the pace of innovation remains relentless. The latest generations of large language models (LLMs) are evolving beyond simply producing fluent text – they are now capable of reasoning, systematically working through arguments and weighing options in ways that resemble human thought. In parallel, more lightweight and more cost-effective models are being developed that dramatically reduce the computing power required to run them. This is crucial for "AI democratisation": enabling smaller enterprises, start-ups, and individual users to harness generative AI without requiring the massive budgets of tech giants. In image and video generation, the frontier is moving even faster, with advanced multimodal models now capable of producing realistic images, cinematic-quality videos, and dynamic 3D environments transforming industries like marketing, entertainment, and product design.

Meanwhile, major cloud providers ("hyperscalers") keep scaling their training clusters in pursuit of what they call "AGI artificial general intelligence. Yet innovation is not purely about adding more GPUs or raw computing power. Hardware constraints particularly in places like China are pushing developers to innovate on the software side: such as increasing XPU utilisation, optimising model architectures, pruning parameters, and improving training efficiency. The result is that several recent Chinese models are now performing on par with or even ahead of Western peers, despite using fewer resources. It's a reminder that intelligence doesn't scale linearly with compute.

Monetization still elusive

But for all this technological progress, the economics remain unclear. Marginal returns from simply throwing more computing power at model training appear to be declining. The cost of training frontier models now estimated in the tens of millions of dollars per run is rising faster than their incremental performance improvements[7]. Moreover, the hardware itself depreciates fast. GPUs and networking gear typically have a useful life of only three to five years, as each new generation quickly makes the previous one obsolete. This means that half the cost of a modern AI data centre including chips and networking equipment must be renewed every couple of years. Analysts already warn that depreciation charges will soon start to weigh on earnings across the hyperscaler landscape. Add to that the fact that monetisation remains elusive and slow, and you have a growing mismatch between the scale of investments and the clarity of returns.

A new source of scepticism has also emerged recently around the so-called "circularity deals" spreading through the AI sector. These are arrangements where hyperscalers, chipmakers and model developers act simultaneously as suppliers, customers and investors creating the illusion of booming demand within a closed loop. A leading example is the deal between Nvidia and OpenAI, signed in September 2025, under which Nvidia will invest up to US$100 billion in OpenAI while OpenAI simultaneously commits to purchasing large volumes of Nvidia's chips and systems[8]. Similar patterns exist in Microsoft's relationship with OpenAI and in Anthropic' s dual partnerships with Amazon and Google. On paper, these arrangements create a self-reinforcing flywheel: capacity is monetised, models get trained, and everyone books revenue. In practice, however, they raise some questions around transparency and sustainability of these investments.

Yes, AI is s making waves by powering a new generation of tools and seamlessly integrating into existing products from Copilot subscriptions and AI-powered search assistants to smarter ad-targeting and next-level productivity tools. Yet, despite the hype and the staggering sums being poured into AI development, the revenue uplift is modest relative to the scale of investments. One of the most commercially successful AI applications to date such as AI-enhanced advertising and productivity tools still rely heavily on existing business models. The much-anticipated breakthrough, the true "AI dividend" that could redefine entire industries and unlock new streams of growth, has yet to fully emerge. For now, the promise of AI's transformative economic impact remains just over the horizon, as companies continue to search for the game-changing applications that will justify the billions being spent. This leaves us with a profound question. Are we witnessing the early innings of one of the greatest technological revolutions in human history, one that will reshape productivity, science, and everyday life or are we already deep into the later stages of an investment bubble, akin to the dot-com frenzy or the U.S. railroad boom of the 1880s? There are arguments for both. The momentum of innovation and the long-term potential are undeniable. But the economics the speed of monetisation, the sustainability of returns, and the environmental footprint remain uncertain.

This is why in our forthcoming publications; we will explore in greater depth the emerging monetization pathways of AI technology. Our focus will be on quantifying the scale of potential market opportunities across both consumer and enterprise domains, and on illustrating how customer journeys, business models, and operational processes may evolve in an AI-enabled economy.

Sources:

[1] IEEE ComSoc Technology Blog, AI Spending Boom Accelerates: Big Tech to Invest $400 Billion in 2025.

[2] UBS Research, Invest in transformational innovation.

[3] McKinsey, The Cost of Compute: A $7 Trillion Race to Scale Data Centers.

[4] Artificial Analysis – **https://artificialanalysis.ai/**.

[5] Goldman Sachs Research, Data Centres Set for 165% Electricity Spike.

[6] Artificial Analysis – **https://artificialanalysis.ai/**.

[7] EPOCH AI, How much does it cost to train frontier AI models?

[8] Nvidia Newsroom, OpenAI and NVIDIA Announce Strategic Partnership to Deploy 10 Gigawatts of NVIDIA Systems.

[9] Bloomberg news reporting.

December 2025 LDS-250245

MACS-P-15303024

ADM-5034941