Series 3

Industry Tailwinds

Firmus AI Infrastructure - Industry Education Series

June 2026

AI creating the fastest-growing demand curve in tech history

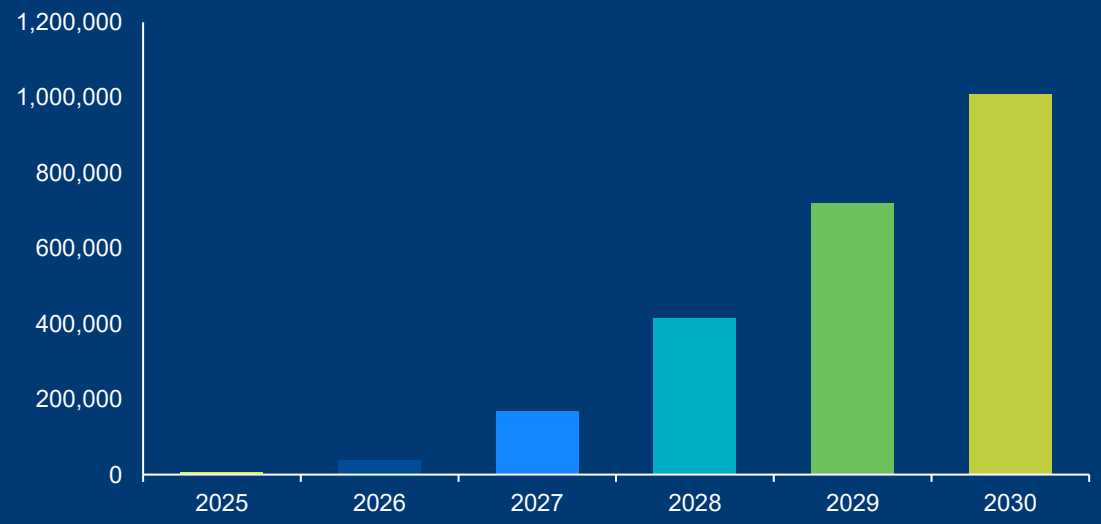
AI demand is growing faster than any prior technology wave

USD 2+ trillion
size of AI data centre market
forecast in 2032[1]

+1m trillion
quadrillion tokens
forecast token consumption by 2030[2]

~176%
CAGR 2025-30
AI token demand growth[2]

Global LLM token demand, 2025–2030 (trillions)[2]



The AI adoption curve is unprecedented

- AI has reached mass adoption faster than any previous technology platform
- Usage is now extending beyond consumers into enterprise software, productivity, search, coding and automation
- Every AI interaction requires compute infrastructure to process and deliver responses
- As adoption broadens, infrastructure demand grows alongside it
- The industry is transitioning from experimentation to large-scale deployment

Source: 1. MarketsandMarkets, AI Data Centre Market, March 2026; 2. Tirias Research June 2026

AI is driven by token consumption

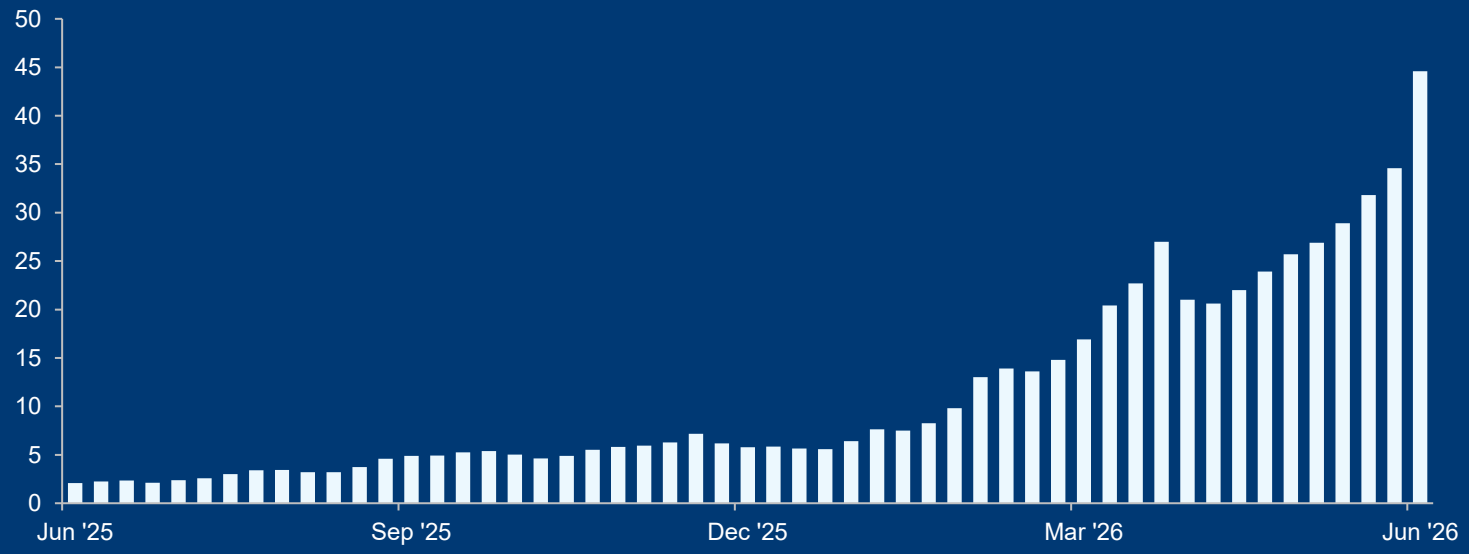
OpenRouter weekly token usage, a real-world proxy for AI inference demand growth

44 trillion
Weekly tokens consumed
at 8 June 2026[1]

1,000x
token demand projected to grow
from 2024 to 2030[2]

+2 quadrillion
~180T/month / ~2 quadrillion tokens
usage expected for 2026[1]

Weekly Token Consumption (Trillions of Tokens)[1]



Why tokens matter

- Tokens are the fundamental unit of AI compute
- Every prompt, response and AI workflow consumes tokens.
- Global token volumes are growing exponentially as AI usage increases.
- More sophisticated models and AI agents require substantially greater token consumption.
- Rising token demand directly drives demand for compute infrastructure.

Source: 1. OpenRouter AI rankings, June 2026 2. Tirias Research June 2026

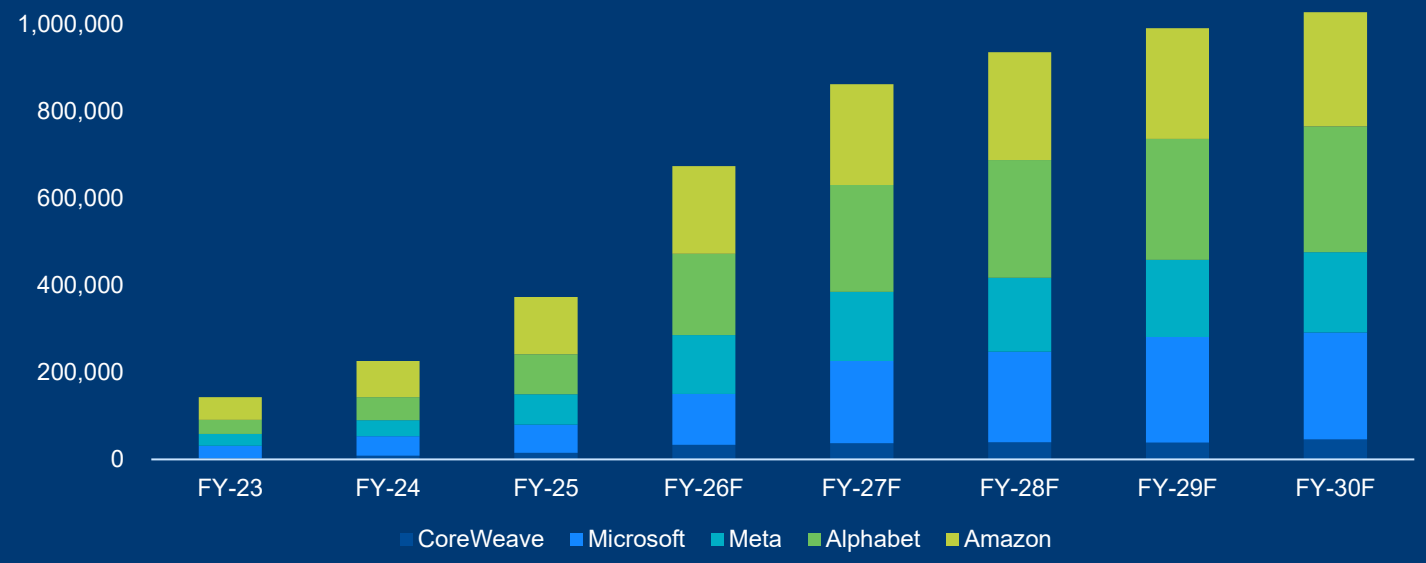
The world's largest companies are investing aggressively in AI infrastructure

USD 670bn+
US hyperscaler AI capex In 2026[1]
~Double the USD 380bn spent in 2025

USD 1T+
forecast Hyperscaler capex on in 2030[2]
Forecast to grow 620% from 2023 to 2030

Capex numbers to be updated right before release – visible alpha

Hyperscaler capital expenditure 2023–2030 (USD billions)[1]



Capital is following demand

- AI has become the highest strategic priority for major technology companies
- Global hyperscalers are investing hundreds of billions of dollars into AI infrastructure
- Infrastructure deployment is occurring at unprecedented speed and scale
- Investment is increasingly being directed toward compute, networking and power capacity
- Spending trends indicate expectations for sustained AI growth over the coming decade

Source: 1. Factset, May 2026 / Visible Alpha May 2026 ; 2. IEA, 2025; 3. MarketsandMarkets, 2026; 4. Morgan Stanley, 2025

Meeting future AI demand requires an unprecedented infrastructure build-out

106M

GPU accelerators required globally by 2030

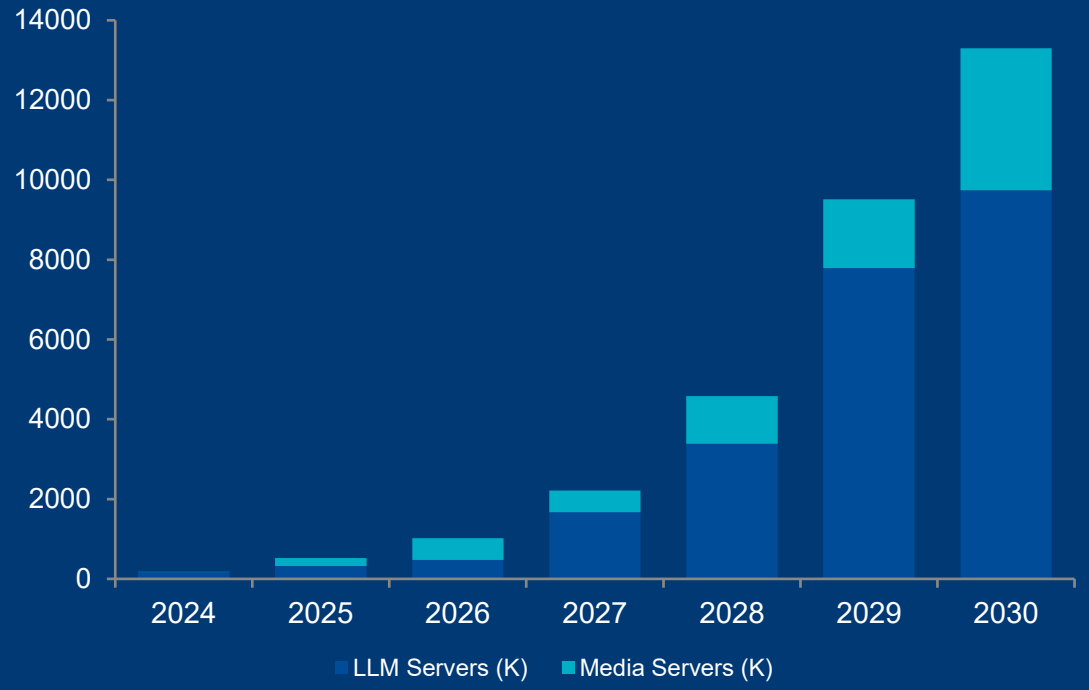
13.3M

AI inference servers required by 2030

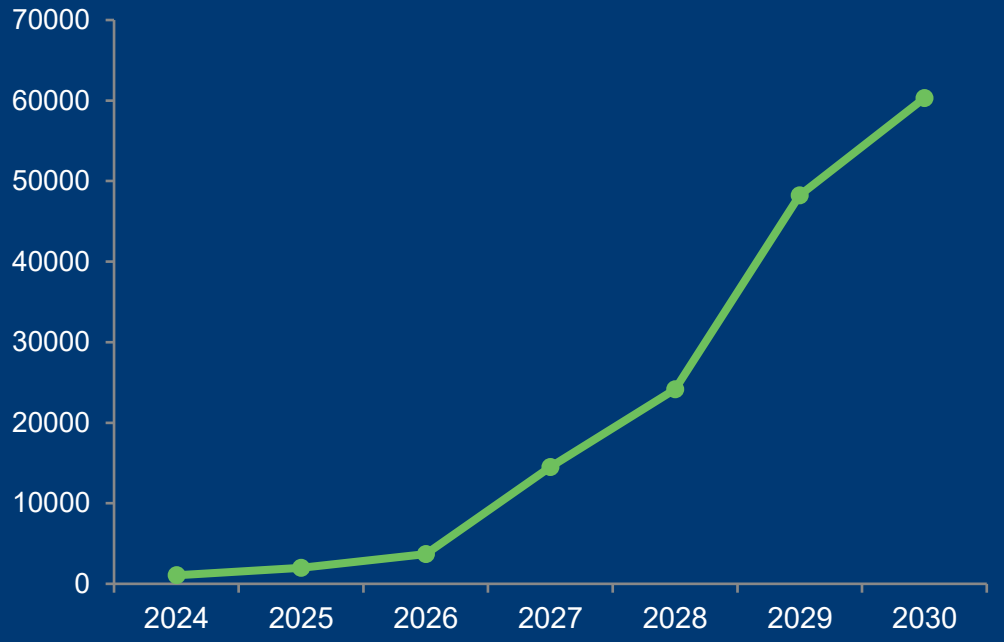
103K

Megawatts of total inference power demand by 2030

Server requirements: Large Language Models (LLM) + Media (thousands), 2024–2030



Total power demand (MW), 2024–2030



Source: Tirias Research, January 2026: Forecasting Generative AI Inference: Demand, Compute, Infrastructure & FTCO 2024–2030

Lower AI costs are expanding AI infrastructure demand

~90%
decline in AI inference costs since 2020 with a further 10-100x reduction projected over 5yrs[1]

1,000x
projected token demand growth 2024–2030[2]

More efficient GPUs lower compute costs

Lower inference costs expand AI affordability

New enterprise and consumer use cases emerge

Token volumes grow exponentially → Infrastructure demand continues to increase

Cost declines are driving adoption, not reducing demand

- Inference costs fell ~90% between 2020 and 2024 as hardware performance improved and model efficiency increased.
- Lower costs have expanded the addressable market, enabling AI adoption across enterprise, consumer and government applications. As with cloud computing and internet bandwidth, lower prices are increasing usage rather than reducing demand.

Usage is growing far faster than cost is falling

- Token demand is forecast to grow 1000x between 2024 and 2030.
- Lower costs are expanding usage faster than efficiency improvements are reducing compute requirements.
- Infrastructure providers are monetised through utilisation and capacity, benefiting from higher workload volumes.

Reasoning models amplify infrastructure demand further

- Reasoning and agentic AI models perform multiple iterative steps to complete a task.
- While the cost per token continues to fall, compute consumption per interaction is increasing.
- More capable models are supporting sustained demand for AI infrastructure.

Lower model costs are making AI accessible to more users, applications & enterprises: driving substantially higher infrastructure demand

1 Mordor Intelligence, GPU as a Service Market, 2026; inference costs fell approximately 90% between 2020 and 2024. 2 Tirias Research, Forecasting Generative AI Inference, June 2026

AI spend is paying off - evidence from the world's largest tech companies

GOOGLE CLOUD

+63%

Cloud revenue of USD 20bn +63% YoY from AI demand (Q1 FY26 revenue USD 110bn)[1]

MICROSOFT

+48%

AI annual revenue run rate, Q1 2026 up 123% YoY[2]

META

+61%

Q1 FY2026 net income increased 61% YoY to US$26.8bn[3]

AI is driving meaningful cloud growth

- AI workloads are becoming a major driver of cloud revenue growth and a key driver of cloud infrastructure demand.
- Google Cloud revenue increased 63% YoY to ~US$20bn in Q1 FY2026.
- Production AI deployments are supporting continued infrastructure investment.

Demand continues to exceed available supply

- AI services have become a meaningful and rapidly growing revenue stream.
- Demand for AI compute continues to exceed available supply.
- Microsoft is expanding its own infrastructure while simultaneously procuring third-party capacity.

AI is generating measurable returns

- The world's largest technology companies are seeing returns from AI to justify the increasing infrastructure investment.
- AI recommendation systems increased engagement and advertising effectiveness across Meta's platforms.
- Meta has increased AI infrastructure guidance to US$125-145bn for 2026, reflecting continued confidence in returns.

The world's largest technology companies are reporting measurable improvements in revenue, profitability and customer demand while continuing to accelerate investment in AI infrastructure and compute capacity

1 Alphabet Inc., Q1 2026 Earnings Release, 29 April 2026: Google Cloud operating margin 32.9% vs 9.4% Q1 2025. 2 Microsoft Corp., Q3 FY2026 Earnings Release, 29 April 2026: AI business annual revenue run rate USD 37bn, up 123% YoY.
3 Meta Platforms, Q1 2025 Earnings Call (internal A/B testing); Q1 2026 Earnings Release, 29 April 2026.

Structural drivers supporting long-term AI infrastructure demand

1	**AI adoption is outpacing infrastructure deployment** AI adoption is growing faster than power, GPU and data centre capacity can be deployed. Supply constraints continue to support pricing, utilisation and long-term infrastructure investment.	**USD 2T+** AI data centre market by 2032[1]
2	**Compute, power and connectivity are becoming strategic assets** The value chain has shifted from software toward physical infrastructure. Access to power, compute and high-performance networking increasingly determines which organisations can deploy and scale AI workloads.	**102,449 MW** global AI inference power demand by 2030[2]
3	**APAC is emerging as a major AI infrastructure market** Governments, enterprises and hyperscalers are investing in domestic AI capacity, supported by data sovereignty requirements and growing regional demand.	**USD 815bn** APAC AI data centre market by 2032[1]
4	**Hyperscaler demand exceeds available capacity** Leading technology companies continue to invest aggressively in AI infrastructure while also procuring third-party compute, creating opportunities for infrastructure providers.	**USD 690B+** US hyperscaler AI capex in 2026 alone[3]
5	**Lower AI costs are expanding the addressable market** Falling inference costs are making AI economic for a broader range of applications, increasing adoption and supporting long-term demand for compute infrastructure.	**1,000,000T+** Projected annual AI tokens consumed by 2030[2]

Demand for AI infrastructure is increasingly being driven by real-world usage, enterprise adoption and hyperscaler investment rather than future expectations

Source: 1. MarketsandMarkets, March 2026; 2. Tirias Research, January 2026; 3. FactSet, May 2026; OECD

Disclaimer

This presentation is provided for general information purposes only and is not intended as an offer to enter into any transaction. The information contained within this document is not necessarily complete and its accuracy can not be guaranteed. We have prepared this presentation without consideration of the investment objectives, financial situation or particular needs of any individual investor.

Before a client makes an investment decision, a client should, with or without Morgans' assistance, consider whether any advice contained in the presentation is appropriate in light of their particular investment needs, objectives and financial circumstances. It is unreasonable to rely on any recommendation without first having spoken to your adviser for a personal recommendation.

The information contained in this presentation has been taken from sources believed to be reliable. Morgans Financial Limited does not represent that the information is accurate or complete and it should not be relied on as such. Any opinions expressed reflect Morgans' judgment at this date and are subject to change. Morgans and/or its affiliated companies may make markets in the securities discussed. Further Morgans and/or its affiliated companies and/or their employees from time to time may hold shares, options, rights and/or warrants on any issue included in this presentation and may, as principal or agent, sell such securities.

The Directors of Morgans advise that they and persons associated with them may have an interest in the above securities and that they may earn brokerage, commissions, fees and other benefits and advantages, whether pecuniary or not and whether direct or indirect, in connection with the making of a recommendation or a dealing by a client in these securities, and which may reasonably be expected to be capable of having an influence in the making of any recommendation, and that some or all of our representatives may be remunerated wholly or partly by way of commission.

The presentation is proprietary to Morgans Financial Limited and may not be disclosed to any third party or used for any other purpose without the prior written consent of Morgans.

Morgans Financial Limited (ABN 49 010 669 726 AFSL 235410)
A Participant of ASX Group

Principal Office: Level 29, Riverside Centre, 123 Eagle Street, Brisbane QLD 4000

This document has been prepared by Morgans Financial Limited in accordance with its Australian Financial Services License (AFSL no. 235410). The views expressed herein are solely the views of Morgans Financial Limited.