cgam

cgasset.com

The AI Capex Boom

Economic Possibilities for our Silicon Children

Chris Clothier | co-CIO

September 2026

Executive summary

Our most simplistic analysis may offer the greatest insight: **every investment boom of which we are aware has resulted in some combination of severe economic slowdown or financial crisis**

We have no insight into timing except to note the conditions seem suitable for the continuation of boom in the near term

- The scale of investment in AI is nothing short of extraordinary. Hyperscaler CAPEX is forecast to be c. $1 tn per year in coming years
- While the exact level is unknown, a majority of the demand underpinning this CAPEX is believed to come from two companies: Anthropic and OpenAI[1]
- We estimate that end user demand to support such investment and generate a reasonable return on capital in 2030 is of the order of $6 tn per year or 15% of US GDP (expressed in 2030 terms)
- We find this to be incredible as it implies either: a) phenomenal acceleration in the trend US growth rate or b) a completely radical realignment of spending and consumption patterns by consumers, corporates and governments
- Instead, we expect that end demand will disappoint and there will be large writedowns on AI investments, coupled with excess capacity and retrenchment for large scale compute
- The probability of this outcome is further increased by the relative performance of large language models developed at much lower cost by Chinese competitors and/or small language models. These models are significantly cheaper to train and operate than so-called frontier models but their performance is converging with their more expensive cousins
- Historical precedents of investment booms accompanying technological revolutions are supportive of the notion the most likely outcome will be excessive capacity, arising from investor exuberance, followed by a period of financial distress, cut-throat competition, capacity reduction and consolidation[2]
- There are reasons to believe that any downturn associated with a reversal of the AI boom will be more acute than historical investment booms:
    - First, the AI capex boom is, by some measures, the largest in history[3]
    - Second, the circular nature of the financing arrangements (e.g. where chip companies and hyperscalers take equity stakes in the AI companies who then commit to purchasing compute and chips from them) which has, until now, acted as a virtuous circle will turn into a vicious circle should the valuations of the AI companies start to fall
    - In turn, this will materially impact the earnings and (we assume) valuations of the hyperscalers and chip companies. Given the very high concentration of these companies in the S&P500 and in global equity markets, the knock-on impact on equity markets would be very significant
    - Fourth, any fall in equity prices will result in a material slowdown in aggregate demand given the exceptionally low US savings rate and the very high level of household wealth in the US attributable to equities
    - Finally, the current extreme US fiscal position, combined with the bond market nervousness may limit the scope for accommodation by fiscal authorities in the event of a downturn

[1] See for instance https://stevehsu.substack.com/p/ai-revenues-circular-demand-and-the and for a more polemical take https://www.wheresyoured.at/premium-ais-circular-psychosis/

[2] Examples of this include: the canal booms & railways booms (both in the US and the UK), the bicycle boom (UK), the roaring twenties & the TMT boom of the 1990s

[3] https://economist.com/finance-and-economics/2026/07/28/ai-revenues-are-growing-fast-but-not-fast-enough

© Copyright CGAM 2026

The Players in the Ecosystem

The AI Companies

The sums of money raised by OpenAI and Anthropic are extraordinary. OpenAI has raised >$180 bn. The amount raised by Anthropic is a little fuzzier but probably around $100 bn. For comparison, Google raised around $30 m prior to its IPO and then raised $1.7 bn at its IPO of which $1.2 bn was primary capital.[4]

It is not yet clear that Open AI and Anthropic have sustainable business models. It is not unusual for early-stage companies to lose money. The scale of losses, of Open AI in particular, raise eyebrows. In 2025 OpenAI had revenues of $13 bn, expenses of $34 bn and an operating loss of $21 bn.[5] Anthropic appears to be faring better. It has even suggested that it was profitable in the second quarter of 2026, though we should treat this claim with some scepticism until we are able to scrutinise its accounts.

It is perfectly rational for companies to be massively loss making if they are using those losses to obtain a dominant market position which will then generate monopoly like profits. Probably the best recent example of this is Uber which generated cumulative losses of $30 bn from 2014 to 2022. It is forecast to generate a profit of $7 bn this year and cash from operations of around $10 bn.[6] They did this by massively subsidising both drivers and riders to build scale and to drive competition out of business.

There is an implicit assumption that these huge investments will lead to high profits which will be sustainable in the future. Open AI and Anthropic do not appear to have a material advantage in terms either of market share or model performance.[7] This is especially concerning given that the best performing Chinese models appear to be developed at costs which are orders of magnitude lower than the leading US models.[8] The performance of the leading Chinese models e.g. Deepseek V4 and Kimi K3 (developed by Moonshot AI) is similar to, though slightly behind, the leading models by Anthropic and Open AI. Finally, the cost per token[9] of the Chinese models is lower making them cheaper to use.

> *Chinese AI companies have been able to develop high performing models at significantly lower cost and without access to the best performing chips*

In summary it appears that the Chinese AI companies have been able to develop high performing models at significantly lower cost and without access to the best performing chips. The US companies do not appear to have a durable advantage and are having to invest ever greater sums to retain their market leading positions. To make matters worse, it is possible to effectively steal the IP of a model through the process of "distillation": the practice of training a model by making queries and recording the output of more sophisticated models.[10] It may turn out that it is impossible to defend an intellectual property lead against such attacks.

Small Language Models

Just as Chinese models developed with significantly lower budgets are demonstrating similar performance to their more expensive American cousins, so too are Small Language Models (SLMs). Not only are SLMs much cheaper to run (they are designed, for example, to run on a desk-top PC), they also cost 2-3 orders of magnitude less to train.[11] The models are frequently open source.

[4] Wikipedia; [5] https://www.wheresyoured.at/exclusive-openai-financials/; [6] Bloomberg

[7] See, for instance: https://openrouter.ai/rankings#benchmarks ;

[8] https://www.reuters.com/world/china/chinas-deepseek-says-its-hit-ai-model-cost-just-294000-train-2025-09-18/

[9] A "token" is a chunk of text, typically about 4 characters long which is basic unit which LLMs use. In turn, AI compute is typically priced per million tokens;

[10] https://www.bbc.co.uk/news/articles/c5ye2gyz0x4o

[11] E.g. https://aisuperior.com/cost-to-train-large-language-model/ & https://aisuperior.com/llm-token-cost/

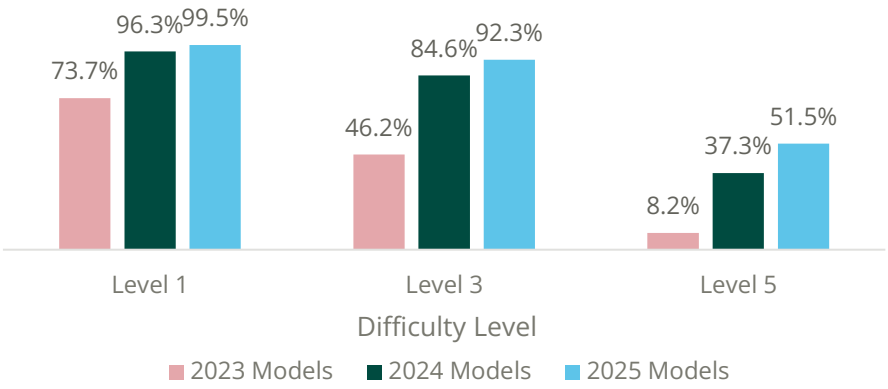
© Copyright CGAM 2026

The Players in the Ecosystem

The absolute performance of SLMs is growing very rapidly and their performance relative to LLMs is narrowing over time. In a recent paper SLMs, were found to perform equally or better than LLMs for 98% of queries of a "single turn chat or reasoning" basis. True, the LLMs continue to perform much better at multi-turn reasoning, typical of the tasks carried out by agentic models. But even in these most demanding applications SLM performance is improving rapidly (see chart below).[12]

While it is extremely challenging to predict how any new technology will develop, it is commonly believed that as AI develops and is more widely adopted, models will become more specialised focussing on narrow tasks (e.g. drug design, code generation, translation, image generation etc.) and trained on bespoke data-sets. Should this trend occur it is likely to reinforce the advantages of SLMs with respect to LLMs.

Chart 1 - Small Language Model performance evolution



Network effects

Nor does it appear that LLMs enjoy any kind of network effects. This should not come as a surprise. Users interact with LLMs on a "one to one" basis. There are no scale advantages from others using the same model. Nor do there appear to be material switching costs: when a better or cheaper model comes along it is straightforward to switch to the new model. Admittedly, switching costs are likely to become greater in the future as AI becomes more integrated in business processes and as applications are stacked upon them.

> *For the same quality model, prices have fallen by roughly 99% over the past 3 years.*

If there are no network effects, and no durable technological leads then are AI companies at least generating high margins on paying users? Again, the answer appears to be "no". The price of tokens (the basic unit of compute) has collapsed. For the same quality model, prices have fallen by roughly 99% over the past 3 years.[13] This is driven in part by algorithmic efficiency, in part by improvements in hardware and in part by better utilisation of hardware (e.g. designing processes to minimise GPUs sitting idle).

[12] *Intelligence per Watt: Measuring Intelligence Efficiency of Local AI* Jon Saad-Falcon, Avanika Narayan *et al* https://arxiv.org/abs/2511.07885

[13] https://valueaddvc.com/blog/how-ai-inference-costs-have-dropped-95-in-two-years-and-what-happens-next

© Copyright CGAM 2026

The Players in the Ecosystem

As importantly, at least for this analysis, is that the AI companies are operating in a very competitive environment. Token costs have collapsed as providers compete for share. This has impacted their margins. OpenAI's "adjusted[14]" gross margin is reported to have fallen to 33% in 2025 from 40% the prior year.

But can they make it up in volume?[15]

Even though the price of tokens is collapsing, overall revenues are growing rapidly. Indeed, as predicted by Jevon's paradox, usage is growing precisely because the cost is falling. It is too early to yet predict how these businesses will scale. There are plenty of sceptics who believe that they don't scale: that they will continue to lose money at greater token volumes[16]. In simple terms, that once you amortise the costs of training new models over the revenue that you can derive from each model, the gross margin falls to a level that does not cover your other overheads.

There have been plenty of instances of businesses which were alleged to have hopeless unit economics which subsequently went on to become very profitable (e.g. Amazon, Uber).

We do not entirely share their scepticism. First, there have been plenty of instances of businesses which were alleged to have hopeless unit economics which subsequently went on to become very profitable (e.g. Amazon, Uber). Second, the industry is in an explosive growth phase. It seems reasonable to assume that the pace of innovation will slow in coming years and therefore the useful life of each model will increase. Third, the cost of training new models may well fall (just as token costs have fallen). Fourth, the economics will *have* to improve. If the AI companies cannot demonstrate a reasonable return they will be starved of capital until the sector's returns improve.

While this may point to an industry that *can* be profitable in the future, it says nothing about whether the historic investments in the sector will turn out to generate reasonable returns nor even if it will be the current incumbents who will enjoy whatever future profit pool is available.

What about the chip companies?

The chip companies exemplified by Nvidia are the clearest beneficiaries of the AI boom. Nvidia's revenues have grown roughly 10x in the space of four years. Its profits have grown more than 20x over the same period.

Nvidia today trades at around 40x current year free cashflow and around 20x next year's free cashflow. Such multiples are not extravagant should those profits prove sustainable.[17] Peter is fond of saying that the semi-conductor industry has been cyclical since God was a small boy. The market does not appear to be share that view.

Chip companies face three main risks.

The first, most simple, is the risk arising from a market slow-down. For example, Cisco's EBITDA fell by 60% from 2000 to 2001 despite retaining its dominant market position.

Second is the risk that Nvidia loses its market leading position. This seems plausible simply because the economics are so attractive that, by definition, new entrants will compete for the profit pool. In the UK alone, two chip companies – Olix and Fractile AI – have collectively raised around $500 m in the last 3 months.

[14] I assume that its *actual* gross margin was rather lower, companies rarely report less-flattering adjusted numbers!

[15] Please forgive wheeling out this dot-com era joke

[16] Notable proponents of this view include Ed Zitron (see e.g. https://www.wheresyoured.at/the-openai-bubble/) and Corey Doctorow (https://www.thersa.org/rsa-journal/issue-1-2026/in-conversation-with-cory-doctorow/)

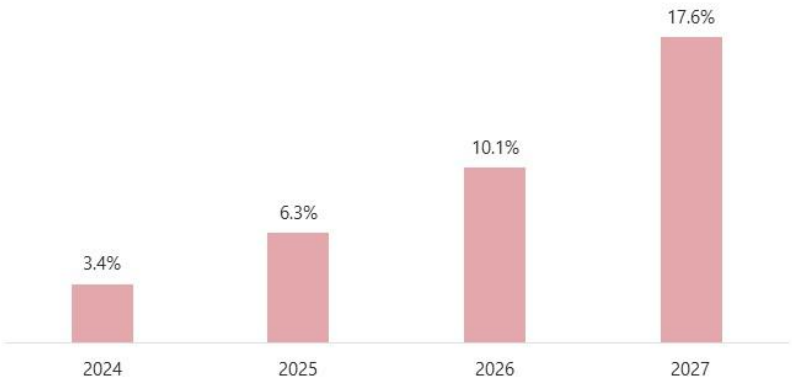
© Copyright CGAM 2026

The Players in the Ecosystem

We have seen reports that $35 bn of Google's tensor chips, manufactured by Broadcom, are being used by Anthropic.[18] Open AI recently announced that its Jalapeno chip, developed in conjunction with Broadcom, outperformed the best available Nvidia chips.[19] Nor is history on Nvidia's side. Today's leaders in microchips can become tomorrow's laggards as demonstrated *inter alia* by Fairchild, Intel and AMD.

The third question is whether the vendor financing and circular financing, that have become so prevalent in the industry, will result in losses for the chip manufacturers.

Chart 2 – Chip Company earnings as % of S&P500



"Chip companies are responsible for 19% of overall earnings growth in 2026 and 77% in 2027. Any set-back to chip companies would be keenly felt by the S&P500.

The proportion of earnings of the S&P 500 attributable to chip companies has risen from 3% in 2024 to 10% in 2026 and is forecast to be 18% in 2027.[20] Chip companies are responsible for 19% of overall earnings growth in 2026 and 77% in 2027. Any set-back to chip companies would be keenly felt by the S&P500.

Can these huge investments generate a reasonable return?

This is a topic we have been writing about for a year[21] and has since become part of mainstream discourse. A year ago the forecast annual capex by the hyperscalers was approximately $600 bn through to 2030. Since then, the forecasts have risen materially. We share detailed forecasts from JP Morgan in the table below. Taken together cumulative spend of $5 tn by the end of 2030 seems likely.

The useful life of these investments is probably greater than 5 years. So it follows that 100% of the physical capital invested between now and 2030 will still be in use. That means total fixed asset investments of $5 tn. We estimate the useful life of these investments will be between 5 and 10 years. It follows that the annual depreciation charge will be between $500 bn and $1 tn.

From the segmental reporting of the cloud divisions of the hyperscalers we are able to make some reasonable assumptions about their business models and, in turn, what revenues are required to generate a reasonable return on invested capital.

[18] https://www.ft.com/content/549f2e23-5aa2-49c7-9ea6-a9784ab7087c?syn-25a6b1a6=1

[19] https://openai.com/index/jalapeno-first-results/

[20] CGAM analysis. Chip company earnings are derived bottom up whereas aggregate index earnings are likely a mixture of bottom-up and top-down which may limit the utility of comparison

[21] https://www.cityam.com/ai-bubble-will-burst-unless-businesses-can-grow-25-times-the-size-of-amazon/

© Copyright CGAM 2026

The Players in the Ecosystem

Table 1 – Hyperscaler Capex Spend($ bn)

$bn	2022	2023	2024	2025	2026F	2027F	2028F
GOOGL	32	32	53	91	185	300	350
AMZN	64	53	83	132	200	300	350
META	31	27	37	70	125	200	225
MSFT	26	36	55	85	190	240	300
ORCL	5	9	7	21	70	85	100
Total	157	157	234	399	770	1125	1325

These point to the hyperscalers needing to generate revenues of around $3 tn per annum. Of course, these revenues are B2B revenues since the compute is provided to AI companies (like Anthropic and Open AI) or other operating divisions of the hyperscalers themselves. We know that Open AI's adjusted gross margin has ranged from 33-40%. We also know that Open AI is wildly unprofitable. We will therefore assume that, to reach profitability, the AI companies will need to generate gross margins of at least 50%. If we assume that 100% of the COGS are attributable to compute spend, it follows that end user revenues from AI will need to be $6 tn per year.[22]

"End user revenues from AI will need to be $6 trillion per year.

To set some context, US GDP is currently around $32.5 tn per year. If we assume nominal growth of 4.5% between now and 2030, those revenues would amount to 15% of GDP. To put this in context this is roughly equivalent to what US consumers currently spend on food and nearly twice the spend on healthcare.[23]

To be successful, these investments require a major rewiring of the global economy

Where might this spend come from? There are three possible sources. First, it displaces spend from other parts of the economy. Were this to happen then we should expect businesses outside tech in the US to suffer no growth or outright contraction as consumer and business spending is diverted.

Second, AI might be used to cut costs by displacing labour. Fretting over losing jobs to innovation is as old as the discipline of economics. Yet historically, such fears have been overblown: the farm and factory workers of yesteryear become baristas, personal trainers and software engineers of today. But a displacement on such a scale in such a short time would be unprecedented and therefore painful. The newly unemployed would doubtless find work eventually, but it would take time.

The third possibility is that the advent of AI results in a step-change in the economic growth rate which, in turn, finances the large expenditures required by the industry. That is what AI proponents hope. We are more sceptical. It reminds us of Robert Solow's quip that "you can see the computer age everywhere but in the productivity statistics".[24] What is more likely is that none of these outcomes happen and instead AI revenues dramatically disappoint. As we discuss below, we don't think this will result in existential risk for the hyperscalers, but it will be significant.

[22] These are clearly highly dependent on our assumptions and should be treated as sitting within a wide plausible range. A helpful sense-check comes from Joachim Klement, a respected analyst, who estimates a range of $2-10 tn with a central estimate of $5 tn

[23] https://www.bls.gov/web/cpi/cpi-relative-importance.xlsx

[24] This may be particularly relevant given that AI appears to introduce productivity dis-benefits: the profusion of "AI slop"; the ability to bring about spurious complaints and legal cases which nevertheless require human responses; the high volume of lower quality code which requires greater levels of engineer maintenance; etc.

© Copyright CGAM 2026

The Provision of Capital

Uses & Sources

Such scale and pace of capital expenditure is placing considerable strain on the free cashflow of these companies. JP Morgan estimates only 18% of the $5.5tn of AI capex expenditure, mostly from hyperscalers, is expected to be met by organic cashflows. In response, hyperscalers have been issuing debt at an extraordinary rate, with greater than $205 bn[25] issued so far this year, and are on track to issue more than in the last five years combined. The size, cadence and currency mixture of issuance is having an impact across global credit markets, increasing borrowing costs and forcing other sectors to plan their financing needs around AI issuance. This has been felt most acutely by longer duration borrowers where hyperscalers have concentrated their issuance. Recent commentary has linked part of the August sell off in the long-end treasuries to AI borrowing, we have yet to see clear evidence of this.

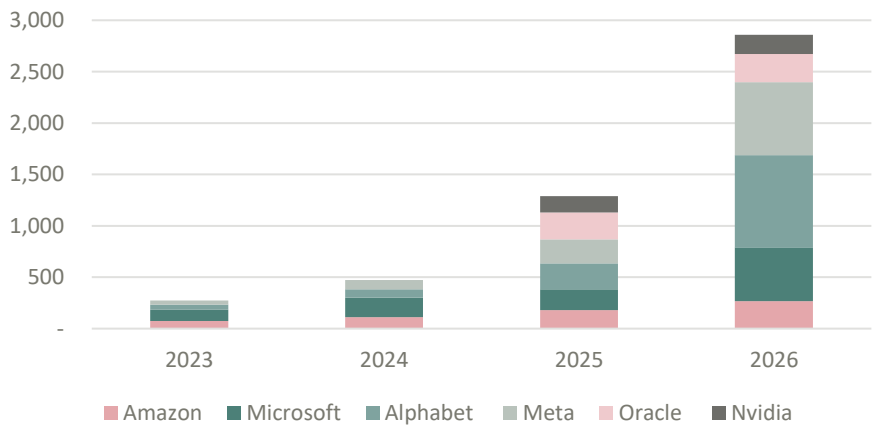
> *Approximately $2.1tn of AI related debt is forecast to be absorbed by the investment grade corporate bond market over the next 4.5 years*

Approximately $2.1tn of AI related debt is forecast to be absorbed by the investment grade corporate bond market over the next 4.5 years with the balance to be funded from equity, structured products and alternative capital. This latter is expected to be the largest single source.

Unsurprisingly, recent attention has turned to the nature of these arrangements and the scope of off-balance liabilities being assumed. For instance, Hyperscalers have entered $1.5tn of financial and operating lease obligations of which $1.2tn are not yet reflected in their financial statements. These are legally binding commitments that are held off balance until the leases commence. We have also found an additional $104bn of credit backstops, financial guarantees and future funding commitments that are included in the footnotes of SEC filings.

Purchase commitments which relate to irrevocable contracts to buy chips, compute and electricity to power data centres are also held off balance. These $1.6tn in commitments form the contractual backbone of the capex plan and are increasing at a remarkable rate. Alphabet alone has agreed $474bn of new purchases in the last quarter.

Chart 2 – Hyperscaler off-balance sheet Commitments ($ bn)



[25] https://www.reuters.com/legal/transactional/us-corporate-ai-debt-surge-tests-investor-limits-fatigue-emerges-2026-08-21/

© Copyright CGAM 2026

The Provision of Capital

The backdrop to assuming these large commitments is the extraordinary growth in revenues for the cloud divisions of the hyperscalers which hit 82% for Google Cloud[26] and 37% for AWS year over year.[27] Their core businesses are also growing strongly.

This creates a horrendous dilemma for the CFO's of the hyperscalers. Given the huge demand for compute, and the profits being generated, it is wholly rational to want to expand capacity. However, the risk is that, at a macro-economic level such demand feels very hard to be sustainable. In addition, the backlogs in demand are believed to largely come from two companies: Anthropic and Open AI. Such revenue concentration is extremely dangerous.

We think it is unlikely that the liabilities that are being assumed by the hyperscalers pose an existential threat to them.

We think it is unlikely that the liabilities that are being assumed by the hyperscalers pose an existential threat to them. They start with credit ratings that range from AA- through to AAA[28] and net debt to EBITDA of less than 0.5x. They all have large and exceptionally high quality core business franchises in sectors which we would expect to withstand a major slowdown in the AI sector.[29]

It is worth noting that they all are subject to additional litigation risk. Meta is perhaps the most exposed to litigation given the concerns of harm of its products to young people.[30]

There are two further risks to highlight. The first is the suggestion that the hyperscalers are already understating their depreciation having extended the useful lives of servers and networking equipment from around 3 years to 5-6 years. A detailed discussion of the merits on either side of this argument is beyond the scope of this paper. Footnote Brief provides a good summary of the issues and suggests that the hyperscalers may be understating depreciation by $75 bn in 2027 and $107 bn in 2028.[31]

The second is the impact of circular financing on the published accounts of the hyperscalers. Their reported earnings have been flattered by revaluations of their stakes in AI companies. For example, in Q1 2026 Alphabet reported gains of $37 bn in unrealised gains in debt and equity securities. This amounted to 47% of total net income in the quarter. We believe that revaluation of Anthropic contributed to substantially all of this gain. Amazon's also experienced large gains.[32]

The Corporate Bond Market

There has been large scale issuance in the corporate bond market and the private credit market by the hyperscalers and data centre SPVs which have some form of hyperscaler backing.

The purpose of this section is to assess whether this extensive debt issuance could cause a credit crisis and, if so, how that crisis might manifest.

[26] https://finance.yahoo.com/markets/stocks/articles/alphabet-q2-2026-earnings-revenue-203058727.html

[27] https://finance.yahoo.com/markets/stocks/articles/amazon-q2-2026-earnings-aws-204411872.html

[28] Bloomberg

[29] We would exclude Oracle from this statement. While Oracle's core data server business is excellent it is become quite small relative to its provision of comput.

[30] https://finance.yahoo.com/technology/articles/meta-u-states-discuss-possible-103911519.html

[31] https://footnotebrief.com/hyperscaler-depreciation-ai-capex-circularity/

[32] Amazon and Alphabet's 10Qs

© Copyright CGAM 2026

The Provision of Capital

Hyperscaler Corporate Bond Issuance

Direct corporate bond issuance by Hyperscalers has gone from annual run rate of $30 bn from 2020-2024 to around $120 bn in 2025 and is forecast to be around $305 bn in 2026[33]. However, as a group their current net debt to EBITDA remains low at 0.5[34]x.

Data Centre Corporate Bond Issuance

Data centre issuance is harder to pin down but we have seen estimates of $250-300 bn in 2026. The split between corporate bonds and private credit is not clear.

Of greatest interest to us is where does the risk sit. Put simply, do bond-holders have recourse to hyperscaler balance sheets or does the risk reside at the data centre SPV?

If the former, then any emerging frailties in the AI eco-system will be borne by the Hyperscalers. Given their prodigious cash generation this is unlikely to result in existential risk for the hyperscalers, though it would likely be extremely damaging to equity markets. If the latter, then a credit crisis is entirely plausible.

> *Our cautious conclusion is that the risks reside primarily with the hyperscalers and with the junior credit tranches (which are primarily backed by private credit) of the SPV structures.*

Answering this question is not straightforward: most of the corporate bond issuance is done privately and therefore the prospectuses are not public.[35] However, our cautious conclusion is that the risks reside primarily with the hyperscalers and with the junior credit tranches (which are primarily backed by private credit) of the SPV structures.

Our rationale for this is two-fold. First, our analysis of the accounts of hyperscalers shows that they are dramatically increasing their off-balance sheet obligations (see Part II, section 1). Second, while we have not been able to read the prospectuses themselves, we have been able to review rating agency reports of one transaction.

The Meta Hyperion transaction[36] is held up as being archetypal. Located in Richland Parish, Louisiana the 4 million sq ft campus has nine data centres with supporting infrastructure including 6 power substations, a water treatment plant and warehouses. Construction costs are expected to be ~$29 bn for the 2.1GW site which demands the equivalent of roughly 17% of Louisiana's total energy generation in 2024.

To finance the construction of Hyperion, Meta partnered with Blue Owl, a private credit firm. They created an SPV to hold the investment which is majority owned by Blue Owl to finance and build the data centre. A subsidiary of Meta entered into 11 triple net leases (one in respect of each individual data centre, an office building and a warehouse) with the SPV. Each lease is 4 years long with 4 renewal options giving a total term of 20 years. Blue Owl and Meta committed $2.5 bn in equity to the SPV. Meta assumes materially all construction risk. Meta also entered into a Residual Value Guarantee (RVG, of which more later) to provide credit enhancement to the SPV. The SPV then issued senior, fully amortising bonds which were rated A+ which is one notch below Meta's AA- credit rating.

[34] Net leverage, just how big is the hidden leverage of AI hyperscalers

[35] E.g. under SEC regulation 144a

[36] The SPV is also known as Beignet named after the famous Louisiana fried snack

© Copyright CGAM 2026

The Provision of Capital

The financial alchemy at the heart of the structure is the combination of the lease structure with the RVG. Should Meta not exercise its option to renew its lease at any of the 4 year points, it becomes liable to pay the RVG. This amount is essentially a make whole for the senior debt holders of the SPV. In the simple case where Meta did not renew any of its leases, the entire SPV would be put up for sale and the RVG is the difference between the third party price for the asset and amount of outstanding senior debt.

Presumably the accounting impact of this is that Meta needs only record a 4 year lease liability on its balance sheet but notes a contingent liability in respect of the RVG. S&P's view is that repayment of the senior debt should be covered by the RVG under nearly all plausible scenarios.[37]

More recently, the c. $14 bn Sopaipilla[38] transaction between Meta and Blackrock has a very similar structure with Meta as tenant and providing a RVG.

Google has also been active in this space providing backstops for multiple data-centres which use Google chips. In a typical transaction the data centre is owned by a specialist data centre operator (often a refugee from the crypto-mining industry), the lessee is an AI company like Anthropic, where Google provides a backstop for the senior debt in addition to providing the chips. In this case the financial alchemy means that while Google is liable for $44 bn in lease liabilities it only records $815 m on its balance sheet (according to FT reporting)[39].

In addition to financing data centres Google is also engaged in more conventional vendor financing, though in this instance appears to have backed the financial liabilities off to Broadcom. The transactions also involve Apollo & Blackstone as financing parties and Anthropic as compute lessee. This is the first tranche – $35 bn – of a financing transaction that could rise to $150 bn when fully deployed[40]. In this structure, Google sells TPU Pods to Broadcom. Broadcom sell them to the SPV which then leases them Anthropic. The TPU Pods are made up of the chips (Tensor Processing Units) which Google has developed with Broadcom which are then used inside Pods which contain thousands of chips per pod. The SPV is financed by Apollo and Blackstone with private credit. The $30 bn of senior notes benefit from a backstop provided by Broadcom. The $5bn in junior notes and equity do not. We understand that the senior notes have been purchased, among others, by Athene, the wholly owned insurance subsidiary of Apollo.

Private Credit

Box 1 – Private Credit

Given the important role that private credit plays in financing the AI buildout, a brief description of our views on private credit is warranted. The challenge with analysing Private Credit is, as the name suggests, its opacity. Our views on Private Credit are therefore lightly held and may change as more information comes available.

[37] https://www.spglobal.com/ratings/en/regulatory/article/-/view/sourceId/101651795

[38] Another fried dough snack!

[39] https://www.ft.com/content/549f2e23-5aa2-49c7-9ea6-a9784ab7087c?syn-25a6b1a6=1

[40] *ibid*

© Copyright CGAM 2026

The Provision of Capital

> *Our instinct is that any asset class which has undergone the explosive growth of Private Credit will have done so at the expense of underwriting standards.*

Our instinct is that any asset class which has undergone the explosive growth[41] of Private Credit will have done so at the expense of underwriting standards. This is further exacerbated by the financial incentives of private credit managers to not realise losses in order to protect their fee income. Nothing that we have read about the sector has contradicted this instinct. We believe that its extensive lending to (by conventional accounting practices) unprofitable software and technology businesses in the 2021-2023 era is an example of undisciplined lending. We suspect that lending to the AI sector will prove to be no different.

Whether Private Credit poses systemic risk to the financial sector is another question. Our view (held lightly!) is that it does not pose systemic risk albeit that the risks associated with it are rising.

Our rationale for this is that in its original form, private credit was structured as private equity style limited partnerships which where the capital was provided by institutions capable of understanding the risks and bearing associated losses. Private Credit managers, keen to grow assets under management, expanded the sources of capital. They launched private closed-ended structures (BDCs) introducing retail investors to the asset class and creating liquidity mis-matches. They acquired life insurers to provide a captive pool of capital to deploy into the sector. As well as increasing the leverage at the investee company level, they increased structural leverage inside their funds.

However, while this may result in poor financial outcomes for many of the parties involved, we don't think it is creating the same systemic risks as built up in the banking sector pre-GFC.

Nevertheless, there are two possible channels for issues in private credit to transmit into the broader economy. The first is a conventional credit squeeze. A dramatic slow down in Private Credit lending could lead to high levels of corporate distress among leveraged borrowers. The second is via the wealth effect. The BIS estimates that 15% of Private Credit AUM comes from retail investors. The Private Credit sector as a whole has c. $2.2 tn of AUM[42]. Write downs in Private Credit could lead to a slow down in consumer spending. However, given that total retail investment assets in the US are c. $37 tn[43], this transmission would only be meaningful if accompanied by a broader equity market sell-off.

Where the risk lies

Our tentative conclusion is that, while AI data centre expansion is largely debt financed the majority of the credit risk resides either with the hyperscalers & chip companies (in the form of lease obligations and guarantees) and with private credit (in the form of junior debt tranches in SPVs). We expect the investment corporate bond market should be insulated from actual losses. However, should there be a general loss of confidence in the AI sector we would expect material spread widening and tightening of financial conditions in bond markets.[44]

[41] The BIS estimates $2.2 tn asset in 2025 up from $100 bn in 2010

[42] https://www.bis.org/publ/bisbull106.pdf

[43] https://www.cerulli.com/press-releases/retail-channel-marketshare-gains-ground-nearing-parity-with-institutional

[44] We exclude from this forecast Neo-cloud operators such as CoreWeave whose debt already trades at distressed levels

© Copyright CGAM 2026

The Wider Economic Impact of the Boom

Profit Pools Arising From The Capex Boom

A common feature of any investment boom is that it leads, in the short term, to hugely increased profits for the corporate sector. The explanation is straightforward: one company's CAPEX is another company's revenue.

We can make some rough estimates as to the overall effect of this. We have read estimates that around 50% of hyper scaler Capex goes to NVIDIA and about 60% goes to technology companies. Of the balancing 40%, 15% goes on acquiring, developing and building sites and 25% goes on power generation, utilities, cooling & network connectivity.[45]

> *One company's CAPEX is another company's revenue.*

If we combine these proportions with the profit margins of each sector we can make reasonable estimates as to the profits generated by this capex by allocating the revenues to each sector and multiplying by the profit margins of those sectors. Of course, this approach is overly simplistic since it is the marginal profitability of each additional dollar of revenue that is relevant rather than the average profit per dollar of revenue. In addition, the boom has resulted in huge price rises among firms like Nvidia who have pricing power[46].

In addition, the major players in the land acquisition / construction part of the value chain are privately owned. In the table below we attempt to make reasonable estimates of the profit pool available assuming $1 tn per year in capex.

Table 2 – Estimated size of AI Capex Boom profit pool

Sector	Companies	Proportion of total Spend	Net Margin	Marginal Profitability	Profit Pool
Chips & Hardware	Nvidia Broadcom AMD	60%	55%	55%	$ 330 bn
Energy & Utilities	Duke Energy Constellation Schneider Electric	25%	11%	20%	$ 50 bn
Land Acqn & Devel	Turner Construction Bechtel	15%	Unknown	20%	$ 30 bn

[45] https://alcapitaladvisory.com/research/intelligence/ai-infrastructure.html . The 50% figure for Nvidia seems incredible but is directionally right. In its last quarter Nvidia had revenues of $90 bn or $360 bn annualised. Forecasts for hyperscaler capex are around $700 bn in 2026.

[46] Nvidia's gross margin has risen from 63% average in the 4 years to 2023 to 73% over the past 3 years.

© Copyright CGAM 2026

The Wider Economic Impact of the Boom

Adding these together gives a total profit pool of around $400 bn.

These numbers are material both in context of US GDP and S&P500 earnings. US GDP is forecast to be $32tn in 2026[47]. S&P500 earnings, on a GAAP basis, are around $2.3 tn[48]. We can now assess the scale of current and projected spend by Hyperscalers both as a % of current GDP and contribution to S&P earnings.[49]

Table 3 – Hyperscaler Capex Contribution to US GDP & S&P500 earnings

Scenario	Annual CAPEX ($ bn)	% US Current GDP	% S&P Current Earnings
Baseline – Pre-AI boom	$150	0.5%	<3%
Current Spend	$650	2.0%	10%
Projected Spend	$1,000	3.1%	17%

Clearly any slow down would have a material first order impact on GDP and S&P earnings. The second order effects could be more consequential. Globally investment banking revenues were $61 bn in H1 2026, an increase of 24% vs. prior year. This is attributable to ECM and DCM fees which, in turn, are driven by AI issuance and AI exuberance.

The Impact Of Circular Financing

There is nothing new in finance. Vendor financing has existed at least since Babylonian times. The presence of vendor financing in an industry is not, by itself, a sign of financial risk or exuberance. A good example of disciplined vendor financing is the auto sector where car companies use their creditworthiness to help their customers finance "large ticket" purchases.

What is unusual about the AI investment boom is the scale and complexity of the different arrangements.

What is unusual about the AI investment boom is the scale and complexity of the different arrangements. As the previous chapter noted, in an investment boom, one companies CAPEX is another's income which leads to higher earnings, higher stock market valuations and so on. Circular financing can be thought of as a turbo-charged version of this process.

In the conventional version, an AI company raises money from venture capital firms. It places an order for compute with the cloud arm of a hyperscaler, the hyperscaler in turn purchases chips from a semi- firm. The hyperscaler generates income from the AI firm which covers (for the time being!) the depreciation of the chips, and sees its earnings increase. The chip manufacturer books income in a straightforward fashion at the high profit margins they currently enjoy.

[47] https://www.imf.org/external/datamapper/NGDPD@WEO/USA

[48] https://www.gurufocus.com/economic_indicators/5749/sp-500-net-income-ttm

[49] There are a number of heroic assumptions here e.g. that 100% of AI capex spend ends up in S&P coffers which is clearly wrong. Conversely, there is plenty of AI CAPEX being done by entities other than the hyperscalers e.g. SpaceEx's run-rate CAPEX is c. $65 bn annually.

[50] Baseline = Approximate levels for 2022 & 2023, current spend $646 bn per Apollo, projected spend based on various forecasts and estimates https://www.apolloacademy.com/wp-content/uploads/2026/02/Hyperscaler-capex-022226_v2.pdf

© Copyright CGAM 2026

The Wider Economic Impact of the Boom

In the circular financing version, the AI company raises equity capital from both the hyperscaler and the chip manufacturer. The money raised then flows back to the hyperscaler and chip company as before. The difference is twofold. First, the investment decision of the hyperscaler and chip manufacturer are conflicted and so capital and valuation discipline may be lost in the rush to increase their revenues. Second, the income statements and balance sheets of the hyperscalers are swelled by the endless cycle of "up rounds" in the AI companies.

We suspect that management teams, equity investors, rating agencies and credit investors attribute too high a multiple to the earnings that arise from the revaluations of these equity investments[51] resulting in balance sheets become extended with ever greater amounts of debt, lease obligations and capex commitments.

How it might unravel

First-Order Effects

If the foregoing analysis is correct then, at some point in the coming years there will be a realisation that: i) the present rate of investment is unsustainable and ii) the investments that have been made will not generate reasonable returns.

We expect that the first step will be a major pull-back in CAPEX and accompanied tighter financial conditions. The most significant impact of this will be on the direct supply chain of data centre boom. The largest beneficiary of the spend is the chip companies who will see a dramatic fall in their earnings. Related sectors such as power and HVAC will see a dramatic slowdown.

"

If the foregoing analysis is correct then, at some point in the coming years there will be a realisation that: i) the present rate of investment is unsustainable and ii) the investments that have been made will not generate reasonable returns.

As the earlier analysis shows, the first order impact on S&P earnings of such a slow-down is highly significant.

The fall in economic activity will flow into the financial sector. Reduced equity and bond issuance will impact investment banking revenues. Equity trading volumes will fall as AI related exuberance dissipates. Property prices will fall impacting, for example, REITs focussed on data centres.

It is possible that secondary impacts will be felt in the prices of raw materials e.g. copper and other metals. Likewise a general slowdown in economic activity seems like to flow from the all the above.

The impact on the hyperscalers is harder to assess. At heart, the hyperscalers have assumed very significant obligations (direct debt, lease obligations, residual value guarantees). The "asset" on the other side of these obligations is effectively the promise of payment by two companies, Anthropic and Open AI. In turn, their creditworthiness rests on two questions. The first is whether they can build sustainably profitable business models, the second is whether they can raise sufficient capital to finance their obligations. In theory, the second is contingent on the first though, in practice, investors frequently make misjudgements about such matters.

© Copyright CGAM 2026

[51] E.g. any multiple greater than 1x!

How it might unravel

Should Anthropic and Open AI be successful in their IPOs and each raise in excess of $100 bn (the rumoured aim) then that would reduce the potential negative financial impact to the hyperscalers of a correction and delay the "day of reckoning".

Nevertheless, it seems likely that a fall in confidence in the prospects of Open AI and Anthropic would have three negative effects on the hyperscalers. First, a direct fall in the revenues of their cloud businesses. Second, falls in the equity valuations of the AI companies would flow into losses in their earnings statements. Third, the increased prospect that the residual value guarantees, and other credit support provided by the hyperscalers to various data centre projects, would be called upon will also reduce earnings as those contingent liabilities rise in value. These latter two impacts would also affect the chip manufacturers given that they have also made equity investments in AI companies and provided credit support of varying kinds.

In turn it seems likely that this could lead to a dramatic fall in the share prices of both chip companies and hyperscalers. Our analysis of the AI related sector of the corporate bond market suggests that material losses are unlikely in the hyperscaler-backed data centre market. Nevertheless, any loss of confidence is likely to be accompanied by a widening in credit spreads and reduction in the availability of credit. The probability of losses in the private credit market seems much higher.

> *Any loss of confidence is likely to be accompanied by a widening in credit spreads and reduction in the availability of credit. The probability of losses in the private credit market seems much higher.*

A feature of the financial economy that has changed since the GFC is that the leverage in the banking system has fallen. However, this has been accompanied by the rapid growth of the Non-bank financial institutions (NBFIs). In turn this means that NBFIs, including private credit funds, play a much greater role in the provision of credit to the broad economy. Any retrenchment resulting from AI related losses could be transmitted into the wider economy through credit tightening.

Probably of greater significance than any credit tightening would be a correction in the equity market. Given the concentration of the Magnificent Seven in the S&P 500, such a correction could be very material.

Second-Order Effects

The very high levels of direct equity ownership by Americans means that, to a greater degree than at any time in the past, the stock market is linked to the real economy and the transmission mechanisms between each are two way. This is particularly true given that the savings rate in the US is exceptionally low[52]. The negative wealth effect of a stock market correction and falling property prices is likely to flow quickly into reduced aggregate demand, potentially triggering a recession.

Under such a scenario the response of fiscal authorities will be interesting. The US is forecast to have a deficit of around 6.3% this year, around $2 tn. The bond market is already showing concerns over the long term fiscal outlook for the US.

© Copyright CGAM 2026

[52] It stood at 2.7% in June compared with an average of 5.2% from the turn of the century until the start of the COVID pandemic

How it might unravel

Historically recessions have been accompanied by falling short & longer term interest rates as central banks attempt to reduce the impact of falling aggregate demand. Presumably a recession is likely to materially increase the US budget deficit; 2-3% would be a reasonable assumption. The ability to use fiscal stabilisers could be limited.

> ***It is possible that a US recession could be the final straw for the bond vigilantes. This could result inter alia in a marked steepening of the yield curve.***

We have been concerned for some time about the risks of a bond market crisis across developed markets. We have identified France, Japan, the UK and the US as being particularly vulnerable. It is possible that a US recession could be the final straw for the bond vigilantes. This could result *inter alia* in a marked steepening of the yield curve. Faced with such an eventuality, the Federal Reserve would likely be forced into renewed quantitative easing on a very large scale. While we believe that the Federal Reserve could ultimately be successful in controlling the yield curve we would expect that it would be more difficult to achieve than in previous eras of quantitative easing and come with greater additional costs (most probably through the currency channel).

A current feature of global bond markets is that, where fiscal and monetary authorities conspire to supress interest rates beyond a "natural" level, the currencies tend to weaken. It seems reasonable that the Federal Reserve will therefore have to choose between controlling the long end of the yield curve and allowing the value of the dollar to fall.

Such an outcome would ripple across other bond markets given that all global markets are largely keyed off US yields. For example, the weekly correlation of moves in the UK 10Y gilt yield to the US 10Y treasury yield is >0.9 over the past 12 months. Nor does it seem likely, given the relative competitive positions, that developed market currencies would appreciate with respect to the dollar. Japan might be a possible exception to this given that the Yen is already very competitive with respect to the US dollar. Presumably the Swiss Franc would also be insulated. It turn, we would expect the knock-on consequences would be elevated inflation and the appreciation of real assets.

To be clear, the scenario we portray is but one of many many possibilities. While each individual step may be reasonably probable, the probability of the final outcome being as described is relatively low given the contingent nature of each step.

To illustrate this, it is also possible to envisage that the economic deceleration arising would be sufficiently severe that it becomes highly deflationary. This deflationary impulse might be sufficient to outweigh the inflationary effects of, what would amount to, yield curve control by the Federal Reserve. We therefore should not be wedded to any particular outcome and ensure that our portfolio is constructed to withstand a wide range of economic consequences.

© Copyright CGAM 2026

How it might unravel

Timing

A feature peculiar to investment booms is that they are self-sustaining. Those supplying the boom enjoy profits today. Those making the investments promise profits in the future. Linking these together is investor belief and provided that the belief endures then the boom can been maintained. In turn this relies upon story telling. Both the language of the investment boom, and its *dramatis personae,* are messianic in character. The world they describe belongs more to the canon of utopias than economic forecasts. The abiding feature of recent economic history is credulity rather than cynicism. That is not to say that there are no AI critics. There are plenty. But in our atomised media landscape each side sits in its own echo-chamber. Furthermore, the AI proponents hold the purse strings at the chip manufacturers, hyperscalers & VC firms. The providers of credit are awash with capital which they need to put to work. A successful IPO of Anthropic and Open AI raising fresh primary capital seems likely. Combined, these factors mean the boom is likely to endure. For how long we cannot say. But we would not confidently predict a reversal withing the next 12 months.

“

The boom is likely to endure. For how long we cannot say. But we would not confidently predict a reversal withing the next 12 months.

Chris Clothier

Chris joined CGAM in 2015 and has co-managed the funds since that date. Prior to that he was a director at IPGL Limited a private investment vehicle with a range of public and private interests predominantly in the financial services industry, where he worked for six years. Before that he worked at MMC Ventures, an early-stage venture capital firm. Chris studied Chemistry at New College, Oxford.

© Copyright CGAM 2026