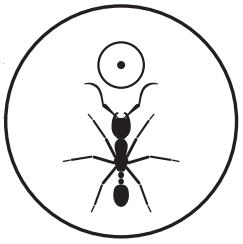


Myrmikan Research

August 14, 2026

Daniel Oliver
Myrmikan Capital, LLC
doliver@myrmikan.com
(646) 797-3134

AI Debt Failure Will Prompt Another Wave of Fed Bailouts

The ever-present desire to reduce costs entices the market to use as its monetary medium the most liquid good, that is, the one with the lowest trading costs. For millennia the market has chosen gold and silver to fulfill the monetary role. Because gold and silver can be inconvenient to carry, especially across long and dangerous distances, the market developed banking, the ability to deposit gold in one place and withdraw it somewhere else.

Banks soon developed a more important function: Banks would lend out deposited gold at interest, keep a margin for itself, and pay the rest to the depositor. If gold began to enter the bank faster than the bank could lend it out, it meant that the interest rate was too high, and the bank would lower it in order to reduce its primary cost. If depositors began to withdraw their gold, it meant that rates were too low, and the bank would raise rates to retain or attract gold.

In this way, before the age of the bureaucratic state, the market itself, operating through the distributed knowledge of depositors, standardized and broadcast the interest rate.

World War I abolished this market system. Belligerent countries forced their banks to buy war bonds at well-below-market rates forcing their banks in turn to drop the interest rate paid to depositors. Depositors would have withdrawn their gold, so the government outlawed the right of redemption to prevent bank runs (the U.S. followed in 1933), putting the world on a fiat standard. From then on, bureaucrats set the cost of money and had to guess, based on economic activity, banking statistics, and the stock market, whether they had it right.

The problem with this design was readily apparent to Republican senator Elihu Root, who voted against the Federal Reserve Act in 1913: "With the exhaustless reservoir of the government of the United States furnishing easy money, the sales increase, the businesses enlarge, more new enterprises are started, the spirit of optimism pervades the community. Bankers are not free from it. They are human. The members of the Federal Reserve Board will not be free of it. They are human."

Everyone wants economic growth, and the more growth there is, the more it looks like the Fed is doing a great job. This is the story of the 1920s (and again in the 1960s and the 1990s and again today). The credit structure collapsed in the 1930s because the Fed had set rates too low, which had encouraged the financing and construction of industrial overcapacity. The crash was the market's attempt to clear out the malinvestments. But then Roosevelt used the power of government to prevent liquidation, turning a correction into a depression.

The intelligentsia interpreted the Great Depression as a market failure, so when the 1960s bubble began to pop in its turn, Fed chairman Arthur Burns began

NOTE: This material is for discussion purposes only. This is not an offer to buy or sell or subscribe or invest in securities. The information contained herein has been prepared for informational purposes using sources considered reliable and accurate, however, it is subject to change and we cannot guarantee the accurateness of the information.

extraordinary interventions to keep the bubble alive. Every time a major bank or company threatened to fail—Penn Central in 1970 or Franklin National Bank in 1974—the Fed would take it as a signal that rates were too high and would therefore lower them. Rates set low enough to save impaired businesses were far too low for the general economy, and inflation became an increasing problem.

Volcker arrived at his post in 1979 with a mandate to slay inflation. He announced a change in policy in that he would raise rates as required to reduce the growth in money supply. He jacked the fed funds rate above 19%, which did help end the inflation amidst a terrible recession, though there were other structural reasons, including overcapacity constructed during the inflationary 1970s, the Soviet Union dumping commodities into the market in an effort to survive, technology causing the measured prices of some goods to plummet, women and then Chinese and then third-world immigrants entering the global economy reducing labor rates.

In early 1984, M2 money supply was still growing above 9% per year, but then Continental Illinois National Bank failed. Several Federal Reserve governors had wanted to raise rates despite the Reagan team's vigorous objection, but Volcker warned the FOMC: "Continental is probably manageable with difficulty; $40 billion institutions are difficult to manage. Having two or three $40 billion institutions is a horse of a different color. If we have two or three, I don't think we're going to stop at two or three."[1] Continental's equity holders were wiped out, but its uninsured general creditors, who had benefited from well-above-market returns for a decade, were bailed out completely, and the Fed dropped rates to help other troubled banks.

Alan Greenspan became chairman of the Federal Reserve in August 1987 and raised the federal funds rate just three weeks into the job. Greenspan had built a career as an economic forecaster, and his personal examination of production and trade statistics convinced him that business conditions risked pushing prices higher. By late September, he had raised the federal funds rate from 6.5% on the date of his appointment to 8%.

Greenspan may well have been correct about his inflation assessment, but on October 19, 1987 the Dow Jones Industrial Average fell by 22.6 percent, its largest one-day plunge ever. Within two days, Greenspan had lowered the federal funds rate back to 6.5%.

Greenspan's attempt to control the market had failed. The Fed's core mission was to protect the banking system no matter how reckless its loans, and Greenspan took that mandate a step further: if the banks were functioning properly, there should be economic growth, and if there is economic growth, then the stock market should be rising, so the best way to gauge the credit system was by looking at the Dow Jones Industrial Average. This is what 1987 taught him. "I'm always worried about the stock market," Greenspan admitted on the television with David Brinkley. "Every day, every hour?" "Yeah."[2]

And thus was born the Greenspan put: the Fed would allow the stock market indexes to fall only so far before interpreting the fall as a sign that rates were too high and injecting newly printed money into the banking system to lower rates. From the end of 1994 to the end of 1999, the Dow Jones Industrial Average gained an average of 24.6 percent *per year.* The press called Greenspan "The Maestro" for his ability to levitate stocks. *The spirit of optimism pervades the community. Bankers are not free from it. They are human.*

1 Federal Open Market Committee Meeting Transcript from May 21–22, 1984.

2 Mallaby, Sebastian. *The Man Who Knew: The Life and Times of Alan Greenspan*. New York: Bloomsbury Press, 2016: 399.

NOTE: This material is for discussion purposes only. This is not an offer to buy or sell or subscribe or invest in securities. The information contained herein has been prepared for informational purposes using sources considered reliable and accurate, however, it is subject to change and we cannot guarantee the accurateness of the information.

Greenspan Grossly Overestimates the Productivity Gains of Computers

In February 2000, Greenspan became worried that inflation was too low. "The most recent data suggest that high-tech prices are dropping at an annual rate somewhere in the area of 17 to 18 percent. Thus, even though that sector's share of GDP is only a few percent, these price declines are having an appreciable influence on the overall inflation rate."[1] Therefore: "If we hold [M2 money supply] to where we are now in the context of rising productivity and rising real growth—that is, keep nominal the same—we will then get declining long-term prices." In other words, the same amount of money chasing more goods would make prices fall, falling prices would make it difficult (or impossible) to repay loans, and the banking system would become impaired.

The obvious solution was to increase the rate of money supply growth: "We can say [to the public] that we are trying to find the range for price stability and that we envisage that potential growth has risen. And, therefore, other things equal, the ranges [of money supply growth] should go up if our objective remains where it was several years ago.… [U]nless there is a change in income velocity … it's an arithmetical truism."[2] Thus, one month before the peak of one of the largest stock market manias the world had ever seen, the Fed was scheming how to accelerate its money printing. *The members of the Federal Reserve Board will not be free of it. They are human.*

Greenspan turned out to be completely wrong. In 1987, economist Robert Solow quipped: "You can see the computer age everywhere but in the productivity statistics." From 1955 to 1979, total factor productivity at constant national prices increased at an average of 0.8% per year; whereas, from 1980 to 1997, the peak years of computer development, total factor productivity crashed by nearly half, increasing by only 0.45% per year.[3]

Some economists denied Solow's Paradox by blaming shortcomings in statistical measurements. Others argued that productivity would simply take longer to appear. Offered as an analogy, economist Paul David pointed out that: "Factory electrification did not … have an impact on productivity growth in manufacturing before the early 1920's. At that time only slightly more than half of factory mechanical drive capacity had been electrified.… This was four decades after the first central power station opened for business."[4]

David's argument failed because the effort in electrifying a factory, including all of the municipal infrastructure, is radically greater than installing a few computers and redesigning a work flow. With the advent of the internet, productivity would, in fact, increase to 0.62 percent annually from 1998 to 2023, yet growth was still materially worse than in the time before computers.

There were, instead, two more reasonable explanations for Solow's apparent paradox. First, those who focused only on the computational efficiencies enabled by computers ignored the limited benefits to actual productivity as well as the risks and costs of implementation. In 1998, the *Wall Street Journal* reported: "The waste is staggering. A 1996 survey of 360 companies … found that 42% of corporate information-technology projects were abandoned before completion."[5] Some projects failed spectacularly "as executives of Fox-Meyer Drug Co. discovered in

1 Federal Open Market Committee Meeting Transcript from February 3, 1999.

2 Federal Open Market Committee Meeting Transcript from February 2, 2000.

3 University of Groningen and University of California, Davis, Total Factor Productivity at Constant National Prices for United States [RTFPNAUSA632NRUG], and U.S. Bureau of Labor Statistics, Manufacturing Sector: Labor Productivity (Output per Hour) for All Workers [PRS30006091].

4 Triplett, Jack. "The Solow Productivity Paradox: What Do Computers Do to Productivity?" 7 May 1998. Brookings Institution: 19.

5 Wysocki, Bernard. "Companies Let Down by Computers Opt to 'De-Engineer' After Clashes." 30 Apr 1998. *The Wall Street Journal*.

NOTE: This material is for discussion purposes only. This is not an offer to buy or sell or subscribe or invest in securities. The information contained herein has been prepared for informational purposes using sources considered reliable and accurate, however, it is subject to change and we cannot guarantee the accurateness of the information.

1996, when the company's brand-new $65 million computer project failed. New software couldn't handle the huge daily volume of orders from pharmacies, among other problems. The breakdown played a significant role in the … company's decision to seek bankruptcy-court protection."

Second, to the extent that computers were effective, they benefited primarily the paper industries, the ones that relied most on computations and word processing, such as finance and law. Moving capital from savers to entrepreneurs and businesses is a necessary and beneficial component of the economy, but it is a cost, the same way that going to the doctor is necessary and beneficial, though medicine is a cost that patients and society should seek to minimize through healthier living and more effective treatments (as opposed to government rationing).

Technology, therefore, should have *reduced* the percentage of GDP spent on finance-related activities. Instead, finance, insurance, real estate, rental, and leasing, which had represented 10.4% of GDP in 1947, grew to 18.9% during Greenspan's tenure and to 21.7% by 2025.[1]

The entertainment industry was the other economic sector to see a massive cost reduction due to technology, which allowed an ever greater part of the population to spend ever more time playing video games and watching movies—activities that may increase GDP but reduce productivity.

Greenspan's effort to increase the pace of monetary growth, therefore, did not operate to balance productivity gains, but instead fueled the internet bubble, and his reaction to the popping of that bubble led to the real estate bubble.

Even assuming Greenspan had been correct about productivity causing consumer prices to fall, preventing that decline repeated the precise error that the Fed had committed in the 1920s, the error Hayek had explained in 1932, words that bear repeating every time the stabilizers gain control of a central bank:

> Instead of prices being allowed to fall slowly [because of productivity gains], to the full extent that would have been possible without inflicting damage on production, such volumes of additional credit were pumped into circulation that the level of prices was roughly stabilized.… Experience has now confirmed what theory was already aware of; that such inflation [i.e., monetary expansion] can also lead to production being misdirected to such an extent that, in the end, a breakdown in the form of a crisis becomes inevitable.[2]

Let us note also that the way the Fed increases the money supply in order to increase prices is by buying Treasury bonds, or funding the growth of the socialist state. And what could undermine productivity more than expanding the size of government, which is another explanation for Solow's Paradox (one that professional economists miss because professional economists are mostly Keynesian and, therefore, socialist (Keynes himself was, more precisely, a fascist)).

AI is also Following Solow's Paradox

Trump appointed Kevin Warsh as chairman of the Fed as of May 22: "I think we're going to do something that's really spectacular.… You know, every [percentage] point

1 Lyndake, Amanda S. "Integrated Historical Input-Output and GDP by Industry Accounts, 1947–1996." Feb 2016 BEA Briefing. U.S. Department of Commerce and U.S. Bureau of Economic Analysis, Gross Domestic Product: Finance, Insurance, Real Estate, Rental, and Leasing (52, 53) in the United States [USFININSREALNGSP].

2 Hayek, F. A. "The Fate of the Gold Standard," in Money, Capital, and Fluctuations: Early Essays, ed. Roy McCloughry (Chicago: University of Chicago Press, 1984): 129.

NOTE: This material is for discussion purposes only. This is not an offer to buy or sell or subscribe or invest in securities. The information contained herein has been prepared for informational purposes using sources considered reliable and accurate, however, it is subject to change and we cannot guarantee the accurateness of the information.

is six hundred billion dollars; all he has to do—if we went down two points, we don't have a deficit any more. And that's without cutting, and it's just a paper charge. We should be the lowest interest rate in the world."[1]

At his swearing in ceremony, Warsh proclaimed: "Chairman Greenspan was the first to tell me and show me what this role demands." One thing Greenspan did was shift the Fed's preferred measure of inflation from the CPI to the PCE, which generally runs well below the CPI, another reason why Greenspan could argue that inflation was too low and, therefore, rates too high. Recently, the PCE has been running *above* the CPI, so the *Wall Street Journal* reports that Warsh will: "'evaluate new information sources and consider methodological changes' to give policymakers more 'actionable' readings on the economy."[2]

Second, just as Greenspan kept rates low because he erroneously believed that technology was causing productivity gains that required printing more money to maintain stable prices, Warsh wrote in the *Wall Street Journal*: "AI will be a significant disinflationary force, increasing productivity and bolstering American competitiveness."[3] He has been using this argument in presentations to support Trump's assertions that interest rates need to be lower now.

Meanwhile, Apollo Global Management's chief economist, Torsten Slok, published a report in June showing no signs of rising profit margins outside of the seven large technology stocks: "this is ultimately what we are waiting for, because the value of AI companies today rests entirely on the promise that margins in the S&P 493 will eventually climb."[4] Slok does not argue that AI cannot increase margins, merely that for the large, capital-intense, heavily regulated industries, the time to implement AI solutions may delay productivity gains far beyond that which the market requires to sustain current valuations—the same argument that Paul David used for the computer.

Warsh disagrees. On July 1, he told the ECB Forum on Central Banking: "The rate of change of improvement in these [AI] models is moving at an exponential level—this is hyper Moore's Law stuff. And so while we might see business surveys that say no big deal, my speculation is six months from now the survey will be saying quite the opposite.… The U.S. is not afraid of productivity-led growth … who knew when the internet was born that the internet was going to create a million and a half jobs as Uber drivers."[5] The average Uber driver makes $20/hr in gross pay, which sinks to $15/hr after expenses, or roughly equal to the minimum wage in twenty-seven states.[6] A 2018 study of the fifty largest metropolitan areas in the U.S. found that the earnings of taxi drivers declined by 10% as Uber was introduced.[7] The total value of taxi medallions in New City York alone plunged from $16.6 billion to $2 billion.

Notwithstanding dramatic improvements in AI models, anecdotal evidence is mounting that AI is following a similar path as computers. Uber recently announced that it burned through its annual AI budget in just four months. Regarding the relationship between the AI spend and innovations, "That link is not there yet,"

1 "Trump to Kudlow: 'This country should have the lowest interest rates in the world.'" 10 Feb 2026. Fox Business.

2 Timiraos, Nick. "Kevin Warsh Invokes Alan Greenspan to Shrink the Fed—and Strengthen Its Chair." 27 Jun 2026. *The Wall Street Journal*.

3 Warsh, Kevin. "The Federal Reserve's Broken Leadership." 16 Nov 2025. *The Wall Street Journal*.

4 Slok, Torsten. "AI: The ROI Runway Could Be Long Outside the Tech Sector." 30 Jun 2026. Apollo Global Management.

5 https://www.ecb.europa.eu/press/conferences/html/20260629_ecb_forum_on_central_banking.en.html and https://www.youtube.com/watch?v=KWKvOruDjd4

6 See Hall, Jonathon and Alan Krueger. "An Analysis of the Labor Market for Uber's Driver-Partners in the United States.." Working Paper 22843. Nov 2016. National Bureay of Economic Research; and "How Much Do Uber Drivers Make in 2026? (Data from 500k+ Drivers)." 1 Apr 2026. Gridwise.

7 Thor Berger, Chinchih Chen, Carl Benedikt Frey. "Drivers of disruption? Estimating the Uber effect." European Economic Review, Volume 110, 2018: 197-210.

NOTE: This material is for discussion purposes only. This is not an offer to buy or sell or subscribe or invest in securities. The information contained herein has been prepared for informational purposes using sources considered reliable and accurate, however, it is subject to change and we cannot guarantee the accurateness of the information.

admitted Uber's COO.[1] Microsoft pushed its employees into Anthropic's Claude Code AI environment in December only to cut access six months later.[2] Even at Nvidia, the microchip-manufacturer at the center of the AI bubble, the vice president of Applied Deep Learning admitted: "For my team, the cost of compute is far beyond the costs of the employees."[3]

Part of the problem is in the nature of AI itself, which uses statistical models to predict the incremental output. As discussed in Myrmikan's September letter, randomness is embedded in the structure of AI, which creates the famous hallucinations that make AI results unreliable. That said, the growth in AI is not in chat functions but agentic AI, creating agents that *do* things. For example, not: "what does a plane ticket to New York cost," but: "research the best flight to New York and book it."

Agentic AI massively increases costs and risks. According to an April 2026 academic paper: "agentic tasks are uniquely expensive, consuming 1000X more tokens than code reasoning and code chat.…" Even worse for businesses: "Task difficulty rated by human experts only weakly aligns with actual token costs, revealing a fundamental gap between human-perceived complexity and the computational effort agents actually expend.… Frontier models fail to accurately predict their own token usage … and systematically underestimate real token costs."[4]

In other words, asking the AI to book the flight is a thousand times more computationally intense than simply inquiring about it (or maybe 500 times more or maybe 2,000 times more; the bill arrives after the activity and is due whether the agent succeeds or not). And, whereas an incorrect response on an inquiry is an annoyance, a mistake booking that flight could be catastrophic: maybe it books by mistake multiple nonrefundable business-class round-the-world tickets to New York. *Axios* reported in May that a mystery company mistakenly blew through $500 million in a single month using Anthropic's agentic AI.[5] *Fortune* reported in March: "Engineer Alexey Grigorev was using Claude Code—a popular Anthropic tool that helps developers write and run code—to update a new website. At first everything seemed normal, until he realized the system had begun destroying the site's live environment: the network, services and, most critically, the database holding years of course data," even worse than what happened to Fox-Meyer Drug Co.[6]

Given the state of the technology, it should not surprise that an MIT study found that "95% of organizations are getting zero return" on their GenAI investment.[7] In a survey of 318 corporate executives, 96% said that GenAI costs were higher then expected, with 71% percent saying they had little to no control over costs.[8] Another survey reported that "84% of companies see 6%+ gross margin erosion due to AI infrastructure costs; 26% report margin impact of 16% or higher." Because of those challenging results, "67% are actively planning to repatriate AI workloads; another 19% are evaluating." That same survey found that 80% of companies miss AI forecasted costs by more than 25%, and 24% miss costs by 50% of more.[9]

1 Angelo, Jake. "Uber burned through its entire 2026 AI budget in four months. Now its COO is questioning whether it's worth it." 26 May 2016. *Fortune*.

2 Warren, Tom. "Microsoft starts canceling Claude Code licenses." 14 May 2026. *The Verge*.

3 Rogelberg, Sasha Rogelberg. "'The cost of compute is far beyond the costs of the employee': Nvidia executive says right now AI is more expensive than paying human workers." 14 Jun 2026. *Fortune*.

4 Bai, Longju et al. "How Do Coding Agents Spend Your Money? Analyzing and Predicting Token Consumptions in Agentic Coding Tasks." Apr 2026. University of Michigan et al.

5 Mills, Madison. "AI sticker shock hits corporate America." 28 May 2026. *Axios*.

6 Nolan, Beatrice. "An AI agent destroyed this coder's entire database. He's not the only one with a horror story." 18 Mar 2026. *Fortune*.

7 Challapally, Aditya. "The GenAI Divide." Jul 2025. MIT NANDA.

8 Gohring, Nancy. "Insight from early enterprise deployments: A Strategic Approach to Scaling GenAI and Agentic AI." Nov 2025. IDC Research, Inc.

9 PR Newswire. "2025 State of AI Cost Management Research Finds 85% of Companies Miss AI Forecasts by >10%." Mavvrik.

NOTE: This material is for discussion purposes only. This is not an offer to buy or sell or subscribe or invest in securities. The information contained herein has been prepared for informational purposes using sources considered reliable and accurate, however, it is subject to change and we cannot guarantee the accurateness of the information.

In other words, implementation of AI is following exactly the same trajectory as computers: intellectuals and paper-pushers see the benefit, but implementation, at least for now, is not measurably adding to *productivity*.

The AI Labs Valued at $2.5 Trillion Despite Flawed Business Models

Business implementation is the only demand that has any potential to justify the capital being directed toward the AI build out. So far, it seems the main net-cost successes are primarily in specific fields, such as analyzing CT scan images or target acquisition for military drones—but these applications, by their nature, do not require *general* intelligence, the kind that requires the ongoing massive capital investment.

AI has had spectacular success in terms of retail usage: AI search is far superior to Google results (especially now that Google is cluttered with ads), AI offers individualized "Dear Abbey" advice on demand, AI can write essays for lazy students, do basic research, can out-compete astrologers. These applications have value that exceed the current pricing of $0 (or $20/month); none of them can support revenue that even dents the cost to provide the services.

OpenAI itself inadvertently revealed the problem in a January 2026 paper:[1] The company boasted that ChatGPT had 800 million users, an impressive figure, yet only 50 million (6.3%) were willing to pay for the service. Of those paid users, the top 5th percentile used the service seven times more intensely than the median paid user. In other words, of the 800 million users, only 5% of the 6.3%, or 2.5 million users, demonstrate that they are getting any real value out of the $20 per month service.

Then there is the threat from Chinese competitors developing AI models that may not be quite as good but are vastly less computationally intense.[2] Last year the scare was DeepSeek, a Chinese company that replicated much of ChatGPT's model capability on a $6 million budget. Deepseek currently charges $0.28 per million output tokens (the unit of AI data) versus $50 for Anthropic's latest model.[3] Last month another Chinese company, Moonshot, released a model that rivals the best that the AI labs offer at a fraction of the cost.

The funding market seems oblivious to the threats to the big three AI labs: OpenAI, Anthropic, and xAI. In March, OpenAI raised $122 billion at an $822 billion valuation. Lead investors included Amazon at $50 billion, AI microchip-maker Nvidia at $30 billion, and SoftBank at $30 billion.[4] Anthropic raised $65 billion in May, valuing the company at $965 billion. xAI was valued at $250 billion prior to being acquired by SpaceX in early 2026.

In 2025, xAI had $3.2 billion in revenue and had an operating loss (i.e., before capital costs) of $6.4 billion. OpenAI had $13 billion in revenue and suffered $21 billion in operational losses.[5] OpenAI also has commitments to spend $600 billion through 2030 (down from a previous plan to spend $1.4 trillion), whereas it expects total revenue to be only $280 billion.[6] Anthropic's annualized revenue doubled from Q1 2026 to Q2, to reach $40 billion, and it reportedly broke even operationally.[7] But it

1 "Ending the Capability Overhang." Jan 2026. OpenAI.

2 Mims, Christopher. "AI's Wider Availability Is Good for China, Not Great for OpenAI and Anthropic." 17 Jul 2026. *The Wall Street Journal*.

3 Rai, Saritha. "Why China's Affordable AI Is a Worry for Silicon Valley." 27 Apr 2026. *Bloomberg News* and Ford, Brody. "Microsoft's AI Sales Mostly Come From OpenAI, Disclosures Show." 5 Aug 2026. *Bloomberg News*.

4 Capoot, Ashley. "OpenAI closes record-breaking $122 billion funding round as anticipation builds for IPO." 31 Mar 2026. CNBC.

5 Edwards, Jim. "OpenAI's financials have leaked, showing $21 billion in losses against $13 billion in revenue." 16 Jun 2026. *Fortune*.

6 Capoot, Ashley and Rooney, Kate. "OpenAI resets spending expectations, tells investors compute target is around $600 billion by 2030." 20 Feb 2026. CNBC.

7 Jin, Berber. "Mind-Blowing Growth Is About to Propel Anthropic Into Its First Profitable Quarter." 20 May 2026. *The Wall Street Journal*.

NOTE: This material is for discussion purposes only. This is not an offer to buy or sell or subscribe or invest in securities. The information contained herein has been prepared for informational purposes using sources considered reliable and accurate, however, it is subject to change and we cannot guarantee the accurateness of the information.

has committed to spend $200 billion on Google Cloud over the next five years, among other multi-billion dollar spend commitments (including $15 billion per year to xAI for compute capacity) that operational cash flow figures do not reflect. In short, the roughly two-and-a-half trillion dollar valuation for the three main AI labs is roughly forty times *revenue* and ignores steep losses.

In April, the *Wall Street Journal* reported that OpenAI "missed its own targets for new users and revenue" as it heads towards a public listing.[1] Then in July, the paper reported that "AI Giants Are Handing Out Tons of Free Computing Power to Grab Startup Share."[2] The AI labs are subsidizing users to get them to integrate their systems to increase the cost of switching to a competitor and to increase market perception of AI demand as the labs head toward public listings. The AI labs are also exchanging credits for equity: the lab gets to count the credits as revenue when they are used plus it gets to write up the value of its equity investment as earnings. In other words, the AI providers are not just granting steep discounts to drive sales, they are actually funding their own customers.

If that were as far as the AI bubble reached, there would be little risk to the general economy nor would there be any public policy implications. The overbuilding of computational infrastructure would elevate electricity and microchip prices temporarily, the bubble would pop, and the rest of the economy would have at least the residual benefit of practically free computing power and lower electricity prices for a time afterward. This is more or less the story of the internet bubble, in which equity fortunes rose and fell, broadband communication become effectively free in the aftermath, and there were no major bank failures.

But the AI bubble does not stop with the three absurdly overvalued AI labs. The tech incumbents have no intention of letting startups control the next technological frontier. Whereas the AI labs currently have the best models that make AI work, they are only just beginning to build out the computational infrastructure required to run them. For now, they must make deals with the so-called hyperscalers: Amazon, Microsoft, Google, and Oracle, which control roughly 70% of global cloud computing capacity, plus there is Meta (Facebook), which uses its infrastructure for its own businesses and AI offerings.

Hyperscalers are Betting More than All of their Cash Flow on AI

The four hyperscalers have announced they have over $2 trillion in firm, non-cancelable orders from AI providers, half of which come just from OpenAI and Anthropic, and are radically changing their business models to meet this demand.

Microsoft (market cap $3 trillion) reports $627 billion in back orders, half of which is just from OpenAI.[3] The company has not yet hit the bond market for financing its foray into AI, but it has engaged in questionable accounting. For example, Microsoft invested $13 billion into OpenAI in 2023, but reportedly $10 billion of that figure was in the form of credits for OpenAI to use Microsoft's cloud services. Microsoft then counts the use of those credits as revenue.

From Q1 2025 to Q1 2026, Microsoft's revenue grew 10.7%. Revenue then jumped 8.6% just from Q1 to Q2 2026 to $90 billion. Cash flow from operations rose an incredible 26% year over year from Q1 2025 to Q1 2026.[4] Then it jumped by another 11.1% in Q2 2026 to $55.4 billion, stunning growth for such a large company. Microsoft

1 Jin, Berber. "OpenAI Misses Key Revenue, User Targets in High-Stakes Sprint Toward IPO." 28 Apr 2026. *The Wall Street Journal*.
2 Clark, Kate et al. "AI Giants Are Handing Out Tons of Free Computing Power to Grab Startup Share." 6 Jul 2026. *The Wall Street Journal*.
3 Third Quarter Fiscal Year 2026 Results. 29 Apr 2026. Microsoft: 6.
4 Q1 refers to the calendar year, which is Microsoft's Q3 in terms of its fiscal year.

NOTE: This material is for discussion purposes only. This is not an offer to buy or sell or subscribe or invest in securities. The information contained herein has been prepared for informational purposes using sources considered reliable and accurate, however, it is subject to change and we cannot guarantee the accurateness of the information.

shares leapt 15% (increasing the market cap by $435 billion) the morning Q2 results were announced.

But cash flows show a different story. Official capex jumped 84% from Q1 2025 to Q1 2026 to $31 billion, then soared to $36 billion in Q2 2026. Dividends and buybacks remained steady at around $11 billion per quarter. Investments into privately held companies (which, as discussed below, is mostly laundered capex) ran from $5.5 billion in Q1 2025 to $7.8 billion in Q1 2026 to $11 billion in Q2 2026. Backing those expenditures out sent adjusted free cash flow from positive $3.9 billion in Q1 2025 to negative $7.8 billion in Q1 2026 to negative $11 billion in Q2.

For normal companies, capex and investments are not counted against operating cash flows because of the assumption that they will produce revenue for many years. However, Microsoft revealed that "roughly two thirds of our capex was for short-lived assets, primarily CPUs and GPUs," meaning the company will begin to suffer soaring depreciation charges (put another way, operating cash flows are well overstated). The company also purports that the huge growth in its cloud services, which represent two-thirds of revenue, is somehow immune from the AI bubble because "nearly 90% [of revenue is] from customers outside of frontier model companies." But why are companies manically paying to put their data on the cloud if not for the intention of using AI to extract value? The most telling comment from Microsoft's earnings call was from CEO Satya Nadella: "All of us are reading this *1873* as the book to be read." Liaquat Ahamed's latest book, *1873*, tells the story of the burst of the railroad investment bubble and the ensuing "first great depression."

Amazon's CEO boasted in the company's most recent earnings release that the company's web services division grew 36.7% year over year while revenue grew 20% to $200 billion, with $496 billion in backlog orders. Net income increased from $18.2 billion in Q2 2025 to $62.6 billion a year later, though $53.4 billion of that figure is Amazon's write-up in the value of its equity investment in Anthropic. As with Microsoft, the market sent the stock up 21% after the earnings release, adding half a trillion dollars in market cap to make the company worth $3 trillion.

But, as with Microsoft, cash flows reveal a different reality. Capex increased 68% to $54 billion in the quarter, an "increase [that] primarily reflects investments in artificial intelligence," while official free cash flow sank from flat to negative $8.8 billion. In Q2 2025, Amazon invested $1.7 billion in corporate acquisitions, whereas that figure jumped to $24.4 billion in Q2 2026—net of other items, the company consumed $23.7 billion of cash in the recent quarter. The company has raised $77 billion in debt over the past twelve months to sustain its spending spree.

Google (market cap $4 trillion) reports over $460 billion in backlog orders, and its Q2 2026 net income increased by a stunning 298%, from $28.2 billion in Q2 2025 to $112 billion. Backing out the $99 billion in net income attributable to a write-up in Google's equity stake in Anthropic, there was still an impressive $10.8 billion gain. Yet to achieve this increase, capital investment nearly doubled to $44.9 billion and investments into private companies jumped to $21 billion.

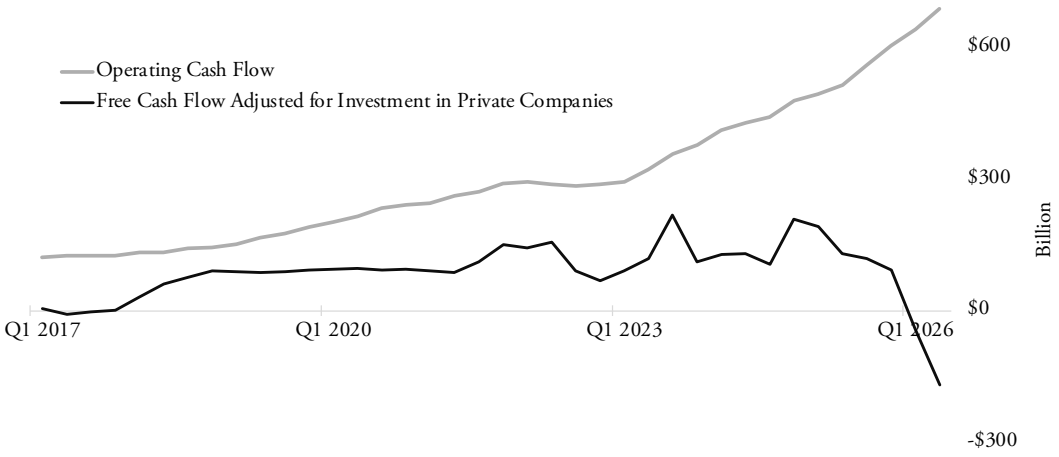
Google's manic spending sent adjusted free cash flow down sharply, to negative $26 billion, and the company had to issue $30.5 billion in common shares, $19 billion in convertible debt, and $24.8 billion in straight debt during the quarter to continue its spendings. In Q1, the company had raised $32 billion in debt, including a $1.4 billion century bond with a fixed nominal yield of 6.125%—the U.S. dollar is unlikely to survive a hundred years, much less a technology company.

NOTE: This material is for discussion purposes only. This is not an offer to buy or sell or subscribe or invest in securities. The information contained herein has been prepared for informational purposes using sources considered reliable and accurate, however, it is subject to change and we cannot guarantee the accurateness of the information.

Half of Google's revenues come from the now ad-cluttered search that produces archaic results compared to AI queries—the company has little choice but to bet everything on AI, and the market currently values this highly uncertain bet at $4 trillion.

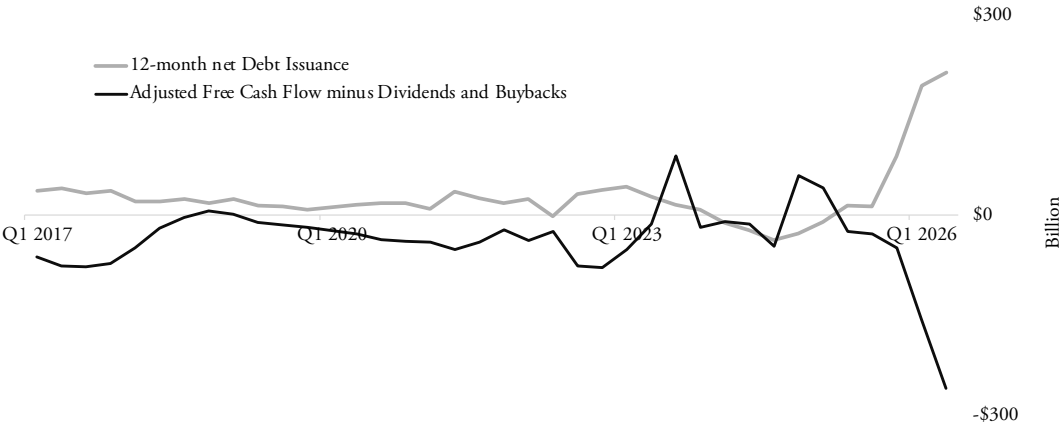
Oracle is the smallest of the hyperscalers with a market cap of only $420 billion, but it has the largest reported backlog at $638 billion. Revenue rose 21% year over year in the last quarter to $19.2 billion, and operating income jumped 137% to $8.4 billion. But, like the others, free cash flow was negative, at minus $1.9 billion. The company has raised $46 billion in long-term debt over the past three quarters, and it expects to raise another $40 billion of capital over the next year. Oracle has ended its share buy back program but continues to pay $1.5 billion per quarter in dividends, which means that the company is, at essence, replacing equity with debt.

HYPERSCALERS ARE SPENDING MORE THAN ALL OF THEIR CASH FLOW
(GOOGLE, MICROSOFT, AMAZON, META, ORACLE, 12-MONTH TRAILING)



The numbers are actually worse than the chart above portrays for two reasons. First, free cash flow is calculated before dividend payments and share buybacks. Including those outflows drives real cash flow even further negative. In other words, these massive technology companies, which are supposed to gush cash, which are supposed to have near infinite gross margins on the marginal sale, need to turn to the debt markets to fund AI capex and continued payments to shareholders.

HYPERSCALERS FUND THEIR SHARE BUYBACKS AND DIVIDENDS THROUGH DEBT



NOTE: This material is for discussion purposes only. This is not an offer to buy or sell or subscribe or invest in securities. The information contained herein has been prepared for informational purposes using sources considered reliable and accurate, however, it is subject to change and we cannot guarantee the accurateness of the information.

The second reason why the finances of the hyperscalers are worse than they appear is because they are laundering capital expense through the neocloud companies, the largest being CoreWeave and Nebius. Unlike the hyperscalers, the cloud offerings of which are powered by microchips designed for generalized use—for businesses running a virtual server to send out marketing emails and to process sales—the neoclouds are building cloud computer data centers with the specialized Nvidia microchips specifically designed for AI.

Hyperscalers Launder Capex Through the Neoclouds

Just as the incumbent hyperscalers are funding the AI labs to ensure that they maintain an equity ownership in their own customers/competitors, they are funding the neoclouds to establish equity ownership in the suppliers of the computing power that makes AI possible.

Microsoft, for example, recently signed a contract with Nebius to purchase $17 billion of computing power. As of November 2025, Microsoft had signed over $60 billion of such take-or-pay contracts. Note that Microsoft's commitments to pay do not appear as liabilities on its balance sheet. If Microsoft were to build these data centers itself, the expenditure would show up as a large, upfront capital expense. By laundering the capital expense through neoclouds, Microsoft and the other hyperscalers can hide investments in the AI infrastructure as an operational expense spread over the next several years.

According to Goldman Sachs, the hyperscalers have $1.5 trillion in lease commitments, of which $1 trillion does not appear on their balance sheets: "From a credit perspective, this treatment can understate leverage and future liquidity needs as these obligations are eventually recognized and contractual payments come due."[1] Separately, Morgan Stanley calculates that the hyperscalers have also contracted $982 billion of purchase commitments for computing power, which also does not appear on their balance sheets.[2]

Meta recently signed a $21 billion purchase contract with CoreWeave. The day that deal was announced, CoreWeave launched an offering of $1.25 billion in senior notes as well as $3 billion in convertible notes. Demand was so robust that the company upsized its senior note offering to $1.75 billion that same day and raised the convertible debt offering to $3.5 billion the following day. Again, Meta's promise, which backs the newly-issued neocloud debt, does not appear as a liability on Meta's balance sheet.

Five days later, CoreWeave announced that the secretive trading firm Jane Street had signed a $6 billion computing contract. The following day, CoreWeave launched another $1 billion bond offering.

Under classical economic theory, savers surrender their current purchasing power when they lend their funds to entrepreneurs; they gain an increase when the investment produces a liquid return. In the case of the neocloud buildout, no one is surrendering purchasing power: the debt providers hold an asset that does not appear on the payer's balance sheet as a liability—the funds to purchase the raw materials are conjured from thin air by Wall Street debt purveyors the moment that a large company indicates demand, and that debt can be pledged as collateral to borrow more funds. And that is why we have inflation.

1 Wigglesworth, Robin. "Just how big is the hidden leverage of AI hyperscalers?" 10 Aug 2026. *Financial Times*.
2 Ibid.

NOTE: This material is for discussion purposes only. This is not an offer to buy or sell or subscribe or invest in securities. The information contained herein has been prepared for informational purposes using sources considered reliable and accurate, however, it is subject to change and we cannot guarantee the accurateness of the information.

Warsh acknowledges this inflation, which is hitting especially the electricity bills of retail customers: "In the near term we can observe it on the demand side, but it's up to the central bank to decide whether it's inflationary." Not the market, not the mass of individual savers withdrawing their gold from banks because interest rates are too low (as during the gold standard), no, the mandarins at the central bank will decide. And as long as the stock market is going up, their decisions must be good. *The sales increase, the businesses enlarge, more new enterprises are started, the spirit of optimism pervades the community*.

Like the AI labs, the neoclouds are market darlings. CoreWeave announced a $104 billion backlog, and the market has assigned the company a value of $62 billion. CoreWeave reported that its revenues increased 24% *quarter-over-quarter* in Q2 to $1.7 billion, and it managed to cut its operating loss by 65% to $49 million. Capex remained constant at $7 billion, and total liabilities jumped from $51 billion to $72 billion. CoreWeave trades at nine times annual revenue—it is not easy to imagine the combination of revenue growth, profit margin, and debt repayment that makes that valuation make sense.

Nebius is worse, but the market likes it more. It has a $50 billion backlog, and revenues grew 51% quarter-over-quarter to $449 million in Q2. But the operating loss increased by 37% to $176 million. The company more than doubled its capex to $5.6 billion, and total liabilities grew from $7.8 billion to $15 billion. But the company's share price surged by 37% on the earnings announcement to give the company a $73 billion market cap, making the company worth *forty times* revenue.

Nscale reported a backlog of $51 billion and raised $2 billion in March at a $15 billion valuation on Q1 revenues of $37 million, a valuation of 101 times revenue.[1] By comparison, during Greenspan's internet bubble, Webvan.com traded at 45 times peak revenue, TheGlobe.com at 28 times peak revenue, and Pets.com at 12 times peak revenue.

The take-or-pay backlog contracts that back neocloud valuations have durations of five to seven years. By then, the hyperscalers will have built out their own infrastructure, and the microchips the neoclouds are currently buying will be obsolete. Moreover, there are few barriers to entry—more than one hundred neocloud companies have entered the market, which, as with any commodity business, will drive pricing down to the marginal cost of production, leaving the debt to drag down the first movers.[2]

The are two bottlenecks, however: electricity and microchips. Navigating the first of these requires real estate savvy and political connections. The latter requires making a deal with Nvidia.

Nvidia Controls the AI Microchip Bottleneck and Funds Both Demand and Supply

Nvidia was founded in 1993 to develop microchips to process graphics for the video game industry. Whereas CPUs (Central Processing Units, the heart of every computer) are designed to execute complex tasks in serial order with speed and efficiency, GPUs (Graphics Processing Units) are designed to executed millions of tiny tasks simultaneously, like updating every pixel on a computer screen during a first-person shooting video game.

In 2006, Nvidia developed CUDA (Compute Unified Device Architecture), a programming platform that allowed developers to use GPUs for general-purpose

1 Lipschultz, Bailey. "Nscale Touts $51 Billion in Contracts, Targets September US IPO." 6 Aug 2026. *Bloomberg*.
2 Mazza, Massimo et al. "The evolution of neoclouds and their next moves." 19 Nov 2025. McKinsey & Company.

NOTE: This material is for discussion purposes only. This is not an offer to buy or sell or subscribe or invest in securities. The information contained herein has been prepared for informational purposes using sources considered reliable and accurate, however, it is subject to change and we cannot guarantee the accurateness of the information.

computing. The initial market was mainly scientists playing with early versions of neural networks and deep learning. Nvidia ensured that new versions of CUDA would exploit continuing advances in the company's microchip design while being backwards compatible, extending the user base, similar to how Microsoft would update its applications to match advances in the Windows operating system long before competitors could react.

When AI mania broke out with the release of ChatGPT in November of 2022, Nvidia found itself with a monopoly over the AI architecture, and the company now sells roughly 90% of AI microchips (nearly all of which are fabricated in Taiwan). Rivals would have to design similar microchips, then recreate the organic software architecture that developed around CUDA over the past two decades, then peel off customers from Nvidia's installed base.

Nvidia's share price ran from a split-adjusted $0.50 the month CUDA was released to $15 when ChatGPT came out, then to $206 currently, granting the company a $5 trillion market capitalization, tying it with Apple as the most valuable company in the world.

Other companies further down the supply chain have also benefited from the AI bubble: SK Hynix is the dominant player in the high-bandwith memory (HBM) that goes into Nvidia's GPUs. The stock rocketed higher by thirtyfold between November 2022 and June of 2026 to a market cap of $1.5 trillion before being cut in half in the past six weeks. Samsung and Micron Technology are also critical suppliers of HBM—their stocks have also soared and their combined market cap exceeds $2.3 trillion. Taiwan Semiconductor manufactures Nvidia's GPUs and has seen its share price jump six times since ChatGPT, giving the company a market cap of over $2.2 trillion. Broadcom makes the high-speed networking equipment that connects Nvidia's GPUs in the datacluster. Its stock has moved sevenfold since late 2022, valuing the company at more than $2 trillion. Adding Intel (which makes CPUs tailored to coordinate Nvidia's GPUs), ASML (the dominant supplier of the photolithography systems needed to manufacture advanced microchips), and AMD (the only credible future competitor to Nvidia) adds another $2 trillion in market cap. These AI suppliers alone represent $9.5 trillion in market cap, up 550% since ChatGPT, 900% since 2019. Part of the reason these stocks have soared so much is that gamblers can speculate on them using single-name leveraged exchange-traded funds.[1]

At least the AI suppliers are currently making money. According to Apollo's chief economist, the silicon suppliers are reporting operating margins of 41%, the cloud companies have operating margins (i.e., before capex) of just 11%, and the AI lab companies have a negative 59% operating margin.[2] These relationships are opposite to every other industry, in which the consumer-facing company has the largest margins, and margins shrink as the supply chain reaches back toward commodity producers. If the AI labs can't make money, neither can the companies elsewhere in the value chain, at least not at the current scale.

Just as the hyperscalers are funding their neocloud suppliers, Nvidia signed a $500 billion strategic partnership with SK Hynix to codevelop HBM optimized for CUDA, keeping the development market on the Nvidia standard.[3] Nvidia is also funding an SK Hynix affiliate to build a data center that will use its microchips.

---

1 Cha , Sangmi and Haram Lim. "Korea Markets Chief Regrets Allowing Single-Stock Leveraged ETFs." 22 June 2026. *Bloomberg*.

2 Slok, Torsten. "In AI, the 41% Depends on the -59%." 7 Aug 2026. Apollo Global Management.

3 Press Release. "SK Group and NVIDIA Expand Strategic Partnership Across AI Factories and Next-Generation Memory." 24 Jul 2026. NVIDIA Corporation.

NOTE: This material is for discussion purposes only. This is not an offer to buy or sell or subscribe or invest in securities. The information contained herein has been prepared for informational purposes using sources considered reliable and accurate, however, it is subject to change and we cannot guarantee the accurateness of the information.

Just as the hyperscalers are funding AI labs to establish equity ownership of and to fund their customers, Nvidia is also aligning with its customers to subsidize its own revenue and ensure that its microchip design remain the standard in the user market. Nvidia has invested $30 billion into OpenAI and $10 billion into Anthropic. On July 26, the *Wall Street Journal* reported that Nvidia is in talks with OpenAI to guarantee a $250 billion financing for a data center in Ohio: "Nvidia's backing would allow the data-center developer … to raise debt at more favorable terms than it could if OpenAI had no financial backer, since OpenAI has no investment-grade credit rating as an unprofitable private company."[1] Much of that $250 billion would return to Nvidia as revenue when OpenAI uses the funds to buy the microchips.

Nvidia also funded Nebius with $2 billion in equity, taking a 9.3% stake, and placed $2 billion of equity into CoreWeave along with a $6.3 billion agreement to repurchase unused computing capacity: "Under the terms of agreement, in instances where the Company's datacenter capacity is not fully utilized by its own customers, NVIDIA is obligated to purchase the residual unsold capacity through April 13, 2032."[2] In other words, CoreWeave gets to buy the specialized microchips, Nvidia books the sale as revenue, records the profit, but then Nvidia guarantees CoreWeave's revenue, a guarantee that CoreWeave can then use to raise debt it needs to issue in order to pay Nvidia for the microchips.

AI Debt, Not the Overvalued Equity, Threatens Systemic Instability Through PE-Captive Life Insurance Companies

CoreWeave has raised more than $18 billion in debt, that, according to the *The Financial Times*, is "secured against CoreWeave's stock of Nvidia GPUs and the capital is used to buy thousands more."[3] This should sound familiar, as it mimics the housing bubble: people borrowed money using their house as collateral, used the loan proceeds to buy more house (either an expansion, or a larger, or a second), which pushed housing prices higher, which created more collateral, which allowed the borrower to borrow more. In this case, that leveraged money shows up as revenue on Nvidia's income statement.

As with the housing bubble, it is the debt that pushes the AI bubble into the Fed's mandate. The most recent CoreWeave GPU debt, which has additional backing by Meta contracts, was rated A3 and placed through a group of banks, including Morgan Stanley, Goldman Sachs, and JP Morgan.[4] The banks merely arranged the deal; the anchor investor was Blackstone, the private equity behemoth with investors such as pension funds, endowments, foundations, sovereign wealth funds, and insurance companies.

That deal was dwarfed by the Meta-backed $30 billion Hyperion neocloud project built on a flood plain. Meta made an in-kind contribution to earn its 20% equity of the project while troubled asset manager Blue Owl Capital took 80%, contributing $7 billion in equity and selling $27 billion of A+ rated bonds with maturities of 24 years, the largest private-credit transaction ever executed. The lead investor, PIMCO, which took $18 billion of the deal, manages money on behalf of central banks, foundations and endowments, pension funds, and insurance companies.[5]

1 Gardizy, Anissa et al. "Nvidia in Talks With OpenAI to Guarantee $250 Billion Financing for Data Center." 26 Jul 2026. *The Wall Street Journal*.

2 CoreWeave, Inc. Form 8-K filed September 15, 2025.

3 Kinder, Tabby. "Wall Street frenzy creates $11bn debt market for AI groups buying Nvidia chips." 4 Nov 2024. *The Financial Times*.

4 Seligson, Paula and Bailey Lipschultz. "CoreWeave Raises $8.5 Billion GPU Loan Backed by Meta Deal." 31 Mar 2026. *Bloomberg*.

5 "Meta + Blue Owl's $27B Bet: Is This the US Blueprint for Financing AI Data Centers?" 27 Oct 2025. Global Data Center Hub.

NOTE: This material is for discussion purposes only. This is not an offer to buy or sell or subscribe or invest in securities. The information contained herein has been prepared for informational purposes using sources considered reliable and accurate, however, it is subject to change and we cannot guarantee the accurateness of the information.

Under the deal, Meta leases and operates the campus but has no exposure to the natural catastrophe risk and can end its lease every four years, which keeps the liability off of its balance sheet (even though if Meta cancels the lease it has to pay the difference between the lease rate of the next tenant and the amortization of the debt).[1]

The giant private equity firms, BlackRock, Blackstone, Apollo, KKR, Brookfield, etc., are all vying for AI infrastructure deals and their associated fees. According to *Bloomberg*: "The amount of debt tied to artificial intelligence has ballooned to $1.2 trillion, or 15% of the investment-grade market."[2] On August 10, the firms listed above announced a syndicate to source $500 billion in additional financing for Nvidia: "The deals will use compute power as collateral for the debt."[3]

The PE firms do not retain exposure to the debt. In fact, according to *Bloomberg*: "The Securities and Exchange Commission has made it easier for data center owners to sell asset-backed securities, potentially opening the door for more debt sales as tech firms scour Wall Street for ways to pay for artificial intelligence. The SEC said a major subset of data-center securitizations don't need to have disclosures and investor protections that similar deals require. That includes risk retention, a requirement that companies issuing asset-backed securities retain some of the debt to better align their interests with investors."[4]

The AI risk sits with PE investors, especially insurance companies, which introduces systemic risk. Insurance companies faced with Bernanke's low interest rates post-2008 had to reach for yield and became natural buyers of high-yield, high-risk private debt. The PE funds that create and issue private debt soon realized that the best way to force more deal flow through the insurance companies was to buy them outright. Whereas PE funds generally have ten year terms, and each PE firm has to go to market each year to raise a new fund as the 10-year vintage fund expires, life insurance company assets represent liabilities decades into the future and are constantly replenished by premiums and thus become a source of permanent capital. The PE firm makes money as the owner of the insurance company, and by charging the insurance company fees for asset management and other related services, and for originating the deals that the insurance company buys; the insurance company bears the economic risk while its upside is capped because all it gets is the debt. The insurance company also becomes a dumping ground for illiquid assets held in the PE firms' other funds, which the PE firms gets to transfer at mark-to-model prices that the PE firm itself determines.

Every state has a Life and Health Insurance Guaranty Association that operates like a quasi-FDIC, guaranteeing policies generally up to $300,000 (a bit higher in some states), competitors must fund the bailout, and then state law generally provides a tax credit to the competitor companies that allows them to recoup the bailout costs over time, placing the residual burden on the public.[5]

This arrangement is worse than the FDIC framework under which banks must contribute *before* they fail and contributions are risk-weighted to punish riskier banks. The state life insurance bailout funds charge the healthier companies *after*

1 Tan, Eli and Maureen Farrell. "How Meta Got Everything It Wanted in a Secret Louisiana Data Center Deal." 27 Jul 2026. *The New York Times* and Wirz, Matt. "Three AI Megadeals Are Breaking New Ground on Wall Street." 11 Nov 2025. *The Wall Street Journal*.

2 Mutua, Caleb. "At $1.2 Trillion, More High-Grade Debt Now Tied to AI Than Banks." 7 Oct 2026. *Bloomberg*.

3 Brush, Silla et al. "Nvidia Taps Wall Street for $500 Billion Funding Commitment ." 10 Aug 2026. *Bloomberg*.

4 Trapanick, Jack and Scott Carpenter. "SEC Exempts Data-Center Bonds From Key Securitization Rules." 10 Aug 2026. *Bloomberg*.

5 Drall, Pranjal and Andrew Granato. "Private Credit's State Backstop: How Private Equity Socializes Risk Through Insurers." 9 Aug 2026. Available at SSRN: 23–24.

NOTE: This material is for discussion purposes only. This is not an offer to buy or sell or subscribe or invest in securities. The information contained herein has been prepared for informational purposes using sources considered reliable and accurate, however, it is subject to change and we cannot guarantee the accurateness of the information.

a collapse in proportion to their volume of policies, which creates an incentive for weaker companies to expand rapidly to raise cash before a collapse and the healthier ones to contract in order to reduce threatened assessments.[1]

Because losses will fall on policy holders who have policies above the $300,000, on competitors, and on the public, the PE owners' incentive is to grab as much exorbitant deal fee as possible before the balance sheets of their insurance company subsidiaries implode. This arrangement should sound familiar. As Mark Carney, governor of the Bank of England, professed about 2008: "when the financial crisis hit, the world's largest banks were shown to be operating in a 'heads-I-win-tails-you-lose' bubble."

From 2009 to 2024, PE ownership of life insurance companies went from near zero to more than $700 billion in assets across 134 insurers, or roughly 7% of the $11 trillion industry.[2] Another $300 billion of life insurer assets are held as portfolio companies within PE funds.[3] McKinsey puts the PE-controlled asset figure at $1.5 trillion as of the end of 2025.[4] According to a recent study, "empirical evidence suggests that PE-owned insurers buy identical assets at higher prices when the seller is an affiliate [PE]." Even though interest rates are now much more attractive than during Bernanke's reign, which should allow a return to more traditional assets, "49.5% of new PE-owned insurers' investment was in privately-placed instruments, compared to just 13.9% for independent insurers."[5]

The most active PE firms have the largest insurance stakes: Apollo has placed $227 billion of its deals into Athene US Life, its captive life insurance company; KKR has placed $163 billion of its deals into its Global Atlantic; Blackstone has put $209 billion into its Fidelity & Guaranty, Everlake, and Resolution Life; Brookfield has placed $90 billion into its American National, and so on.[6]

In another similarity to the housing bubble, insurance companies are sensitive to bond ratings because higher-rated bonds attract lower capital charges, which allows the insurance company to use additional leverage. As a result, the *Financial Times* reports, "some of the world's biggest asset managers, including Blackstone and Apollo, are now among the most frequent users of ratings from firms beyond the big three."[7] One private equity sponsor is currently under investigation for sending $8 million to a ratings firm, Egan-Jones.[8] Two former employees recently sued Egan-Jones alleging that they were pressured to inflate ratings to gain business.

And if the rating agencies don't cooperate, there is another way for the insurers to hide the composition of their assets: reinsurance. Reinsurance is not new—as far back as the Middle Ages insurers of ships could spread their risk by reinsuring portions of large contracts with other insurers, thereby opening a distinction between the sales operation and the financial analysis. The practice spread to life insurance in the early twentieth century, which enabled the issuance of larger contracts for the more affluent.

The essence of reinsurance is spreading the risk, so a reinsurer that is a subsidiary of the insurer is a walking contradiction. Yet Apollo-backed Athene US Life has reinsured substantially all of its $200 billion in annuity liabilities with Athene

1 Drall: 28–30.
2 https://www.federalreserve.gov/publications/2026-may-financial-stability-report-leverage.htm Table 3.1
3 Drall et al: 40.
4 Torbey, Henri et al. "Beyond $1 trillion: The next chapter for insurance and private capital." 15 Apr 2026. *McKinsey & Company.*
5 Ibid: 45, 47.
6 Harris, Lee. "How insurance became the lifeblood of private credit." 24 Feb 2026. *The Financial Times.*
7 Harris, Lee et al. "The new crop of rating agencies behind the private credit boom." 10 Nov 2025. *The Financial Times.*
8 Harris, Lee. "Walter insurers paid millions of dollars to credit rating provider Egan-Jones." 2 Aug 2026. *The Financial Tim*

NOTE: This material is for discussion purposes only. This is not an offer to buy or sell or subscribe or invest in securities. The information contained herein has been prepared for informational purposes using sources considered reliable and accurate, however, it is subject to change and we cannot guarantee the accurateness of the information.

Holding Ltd. Bermuda. According to *Bloomberg*, shifting liabilities to a captive reinsurer creates "advantages that can be myriad and arcane, including more flexible capital requirements, alternative accounting standards and fewer disclosures."[1] According to the BIS: "By 2023, US life insurers had ceded [reinsured] $2.1 trillion in reserves, up from $500 billion in 2017, representing 25% of their total assets. Offshore reinsurers accounted for 40% of these risks, up from 14% in 2017, with some of this activity occurring in jurisdictions with less stringent regulatory frameworks."[2]

The entire private credit universe is $2 trillion, $849 billion of which is now owned by life insurers.[3] So while the market is too opaque to track exactly how much AI debt is inside the life insurance companies, the fact that these entities hold 42% of the entire private credit market and are owned and managed by the PE shops placing the AI debt, it's a good bet that they have more than their fair share. And the balance is not exactly safe either: those conversant in business and financial affairs understand that PE-financed companies temporarily expand margins by cutting costs and quality and outsourcing employment to third worlders, which degrades the customer experience and the company's long-term viability.

Life insurance companies in U.S. currently report $11.0 trillion in assets and $10.6 trillion in liabilities, giving the entire industry only a 4% equity cushion—even a small downturn in the credit markets will push the insurers toward insolvency.[4] Apollo's Athene or KKR's Global Atlantic will not be the first to go. They are too large, too careful. But every free-money machine attracts the less capitalized imitator who pushes the model too aggressively and drags down the senior players—Ralph Cioffi's internal fund at Bear Stearns failed in June of 2007; Lehman would not collapse until fourteen months later. As Senator Root put it in his 1913 speech: "finally someone whose judgment was bad, someone whose capacity for business was small, breaks; and as he falls he hits the next brick in the row, and then another, and then another, and down comes the whole structure. That, sir, is no dream. That is the history of every movement of inflation since the world's business began, and it is the history of many a period in our own country."

The PE firms captured the insurance companies for their permanent capital, except that the capital is not permanent—policy holders may, in fact, demand their policies be cashed out. There are two sources of friction that bank depositors do not face: first, policy holders face surrender penalties of typically 10%, but these go to zero after ten to fifteen years, so the oldest accounts—the ones that have grown to be the largest—have low or no penalties for redemption; second, gains are taxed instead of being passed along to beneficiaries tax-free at the death of the holder. But paying a tax on gains is a lot better than losing everything when the AI debt collapses. And if policy holders start redeeming, the insurer must sell what it can, which will be its liquid holdings first. Those who delay redeeming will face ever greater exposure to the illiquid, mark-to-model private debt. Then, as Bernanke observed after the 2008 panic: "If one bank is having problems, people at the bank next door might begin to worry about problems in their bank. And so, a bank run can lead to widespread bank runs or a banking panic, more broadly." So, too, for life insurance companies. Just as governments had to cancel the right to redeem liquid gold during World War I, they will have to cancel the right to redeem insurance policies or else print the money to make up the losses.

1 Schoenberg, Tom et al. "The Offshoring of America's Retirement Savings." 16 Nov 2025. *Bloomberg*.

2 Aquilina, Matteo et al. "The transformation of the life insurance industry: systemic risks and policy challenges." Oct 2025. BIS Papers No 161: 27.

3 Meisenzahl, Ralf. "Life Insurers' Private Credit Investments and Annuity Market Share Capture." 27 Apr 2026. Federal Reserve Bank of Chicago.

4 https://www.federalreserve.gov/releases/z1/current/html/S1281_s.htm

NOTE: This material is for discussion purposes only. This is not an offer to buy or sell or subscribe or invest in securities. The information contained herein has been prepared for informational purposes using sources considered reliable and accurate, however, it is subject to change and we cannot guarantee the accurateness of the information.

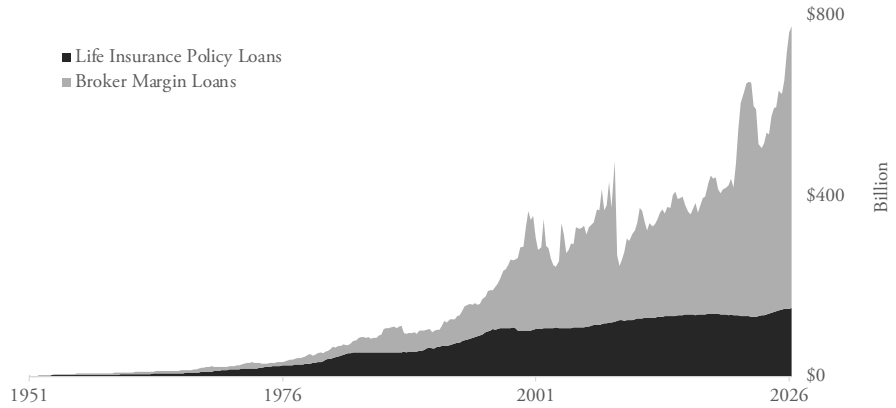
It was, in fact, the threat of policy-holder redemptions that prompted the Fed to nationalize A.I.G. in 2008. As Andrew Ross Sorkin reported at the time: "A.I.G. has more than 81 million life insurance policies with a face value of $1.9 trillion globally. If policyholders lost faith in A.I.G. and rushed to cash in their policies all at once, the entire life insurance industry could falter … there probably would not be enough money to pay out to consumers from what's known as a guarantee fund. Other regulated insurance companies, which have been weakened by credit losses, would be required to pay money into the fund to cover the shortfall, weakening them further and in some cases bankrupting them."[1] And that was when insurance companies held liquid investments.

Warsh Will Submit to the Market and Print

Nvidia's suppliers' $8 trillion in increased market cap since ChatGPT's release in late 2022, Nvidia's own $5 trillion increase in market cap, the hundreds of billions of dollars of neocloud equity valuations, a substantial portion of the $8 trillion increase in the market caps of the hyperscalers since ChatGPT, the $2 trillion in off-balance sheet hyperscaler liabilities, the $2.5 trillion valuation of the AI labs, the value of the start-ups relying on free AI handouts, the $1.2 trillion issued in bonds relying on AI assets as collateral, a trillion dollars of life insurance policy value, the 2.4% of US GDP being directed toward the AI buildout, all depend on products that regularly produce catastrophic errors, on continued meteoric growth in demand even after pricing adjusts radically to reflect costs, on the assumption that the government will erect barriers against Chinese competition. The collapse will threaten the financial system and shake America's social foundations.

The hyperscalers plus Nvidia plus the three U.S.-based Nvidia suppliers comprise 29.7% of the S&P500—eighty million working Americans every two weeks have part of their paychecks shunted into retirement accounts that buy that index. The inflated market caps provide the collateral that backs over $600 billion in margin loans, used either to boost consumer spending or to fund more stock speculation. An additional $152 billion of consumer debt has been borrowed by life insurance policy holders against their policies, that is, debt borrowed against the debt owned by the insurance companies, increasingly comprised of illiquid private credit held offshore.[2]

LEVERAGE ON LEVERAGE



1 Sorkin, Andrew Ross. "The Case for Saving A.I.G., by A.I.G." 2 Mar 2009. *The New York Times*.

2 Board of Governors of the Federal Reserve System (US), Life Insurance Companies; Policy Loans; Asset, Level [BOGZ1FL543069405Q].

NOTE: This material is for discussion purposes only. This is not an offer to buy or sell or subscribe or invest in securities. The information contained herein has been prepared for informational purposes using sources considered reliable and accurate, however, it is subject to change and we cannot guarantee the accurateness of the information.

Warsh declared at the ECB Forum: "the AI boom is showing itself first and very prominently in the United States. I'd certainly rather have this problem, where we have massive capital expenditures instead of thinking years back where we had financial engineering, where we had companies that weren't able to deliver profits, so they would do shareholder buybacks—well, right now they're investing in the future because their expectation is the supply side of the economy will expand, and if it does, that has huge implications for monetary policy." *The members of the Federal Reserve Board will not be free of it. They are human.*

AI is incredible technology, the stuff of science fiction made real. What is less obvious is that at the center of every bubble lies a confusion between utility and value. Dial the calendar back two centuries and edit Warsh's statement by replacing "AI" with "canal" to see the problem. Canals were indeed very useful, but governments continued to back them even after it became obvious that railroads would make them obsolete. Four decades later the railroad bubble saw competing firms build two or three railroads between two cities, just as the microchip manufacturers, the hyperscalers, and the neoclouds are all constructing redundant data centers. When the bubble popped, the tracks remained extremely *useful*—but, having eliminated scarcity, the railroads lost most of their *value*.

The web browser bubble was based on this same confusion: Netscape had revenues of $11 million and was losing money. But the financials did not matter: the market saw the utility of breaking Microsoft's monopoly on computer software and assigned the company a $2.2 billion valuation, which tripled within six months. Lycos was granted a market cap of $300 million based on $1 million in revenue. But the browsers were based on protocols that were in the public domain—there was no bottleneck to capture, no value could be extracted. Even today, Google's Chrome browser is free, offered to entice users to subscribe to Google's other, paid features. If Google began charging for Chrome, users would simply switch to Safari or Firefox or Brave or a dozen others.

AI will follow a similar trajectory. Microchips are commodities, priced on supply and demand. Nvidia may own its market, but overinvestment or innovations that make more efficient use of its microchips could radically reduce demand and, therefore, pricing power—the same way that innovations just between 1996 and 2000 allowed forty times more data to be sent through fiber optic cables, turning scarcity into a glut and crushing telecommunication stocks.[1]

The neoclouds will collapse as microchip overcapacity make debt payments unsustainable. The AI labs will have to reorganize as the reality of Solow's Paradox makes the market understand that AI, while real, will have a much slower effect on productivity than currently anticipated. They will be unable to meet even a fraction of their commitments.

The hyperscalers will not disappear—they control large, real businesses. But their market caps must adjust to remove the AI premiums, much of their cash flow from the past two years will turn out to have been wasted, and they will have to cut dividends and buybacks to pay off their new on and off balance sheet liabilities.

Burns, Volcker, Greenspan, Powell all joined the Fed with new, optimistic theories on how to run the central bank and choose the correct interest rate for the country while controlling inflation. The market swiftly tested each of them with market crashes and bank failures. They folded. They had no choice. The interest rate

1 Couper, Elise and Hejkal, John P. and Wolman, Alexander L. "Boom and Bust in Telecommunications." FRB Richmond Economic Quarterly, Vol. 89, No. 4, Fall 2003: 10.

NOTE: This material is for discussion purposes only. This is not an offer to buy or sell or subscribe or invest in securities. The information contained herein has been prepared for informational purposes using sources considered reliable and accurate, however, it is subject to change and we cannot guarantee the accurateness of the information.

necessary to support malinvestments and fund the government is also the rate that is too low for the rest of the economy, and the government will always choose the policy that supports the banking system over that which maintains the value of the currency.

Warsh is having his honeymoon with smiley speeches about how easy it will be to fix his predecessor's obvious failings. He promises an interest low enough to ensure AI gets funding while high enough to contain inflation, the same way politicians promise to cut taxes and also increase public spending.

Then the stock market will crash, the boomer nest egg will crack, the AI-backed bonds will implode, pension funds will start drowning, along with endowments and foundations and sovereign wealth funds. Insurance companies will face runs and dump assets until the government quasi-nationalizes the industry, freezes policies, blames the free market, and demands the Fed fund the bailout. Warsh will have to engage in extraordinary interventions to keep the bubble alive: "Interest rates should be the dominant means through which we make monetary policy. If we're in a crisis, there could be a different set of rules." There is actually only one rule: print or die.

Republicans will get the blame and the populist impulse that has been fueling Trump will energize the radical socialists that run the Democrat Party. PE titans, who issued but are not exposed to the toxic debt, will buy the surviving assets at pennies on the dollar, and societal wealth will concentrate further. For the rest, gold may not be the only safe haven, but it is certainly the clearest and the simplest and the best.

The information transmitted is intended only for the person or entity to which it is addressed and may contain confidential and/or privileged material. Any review, retransmission, dissemination or other use of, or taking of any action in reliance upon, this information by persons or entities other than the intended recipient is prohibited. If you received this in error, please contact the sender and delete the material from your computer.

The material contained herein is for discussion purposes only and is not an offer to buy or sell securities. It has been prepared using sources considered reliable and accurate, however, it is subject to change and the accurateness of the material cannot be guaranteed.