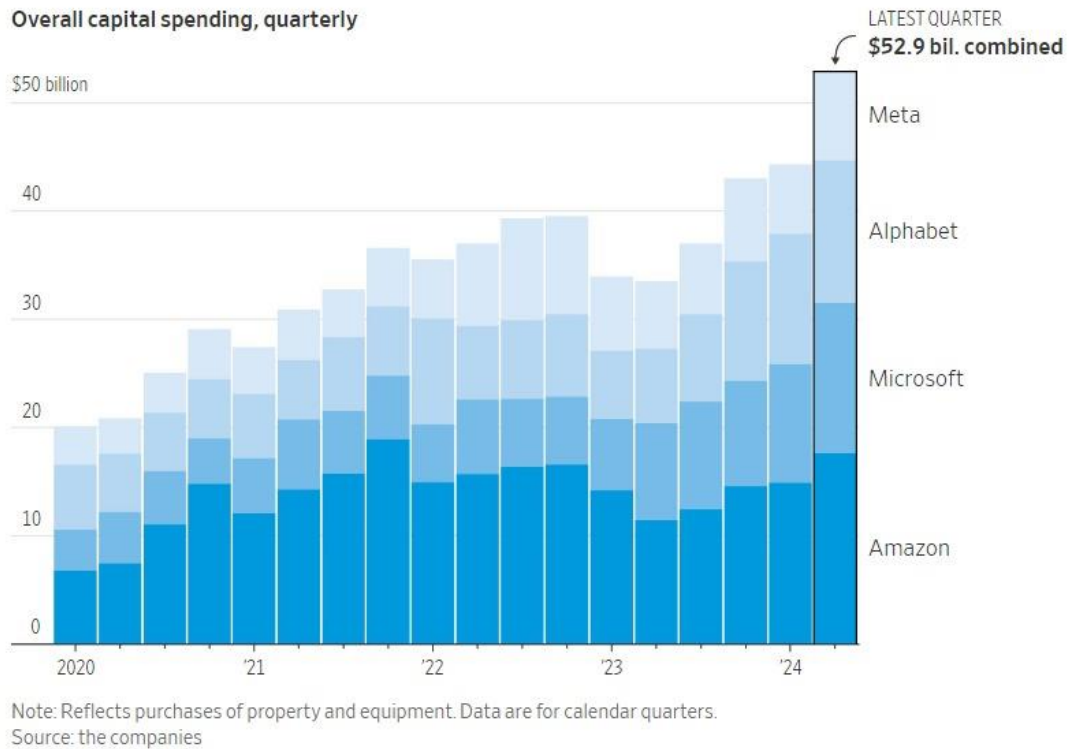


The AI Spending Spree, in Charts

Tech giants and investors are shoveling cash into artificial intelligence amid questions about whether it will pay off



By [Nate Rattner](#) and [Tom Dotan](#)

Generative artificial intelligence has sparked one of the biggest spending booms in modern American history, as companies and investors bet hundreds of billions of dollars that the technology will revolutionize the global economy and one day lead to massive profits.

The question is when, and even whether, all those investments will pay off.

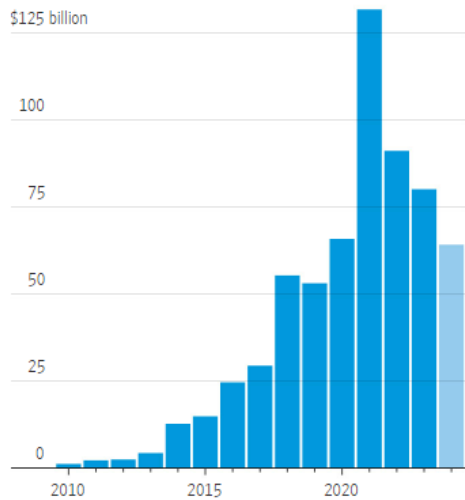
Apps like OpenAI's ChatGPT have attracted hundreds of millions of users, but relatively few people are paying for premium versions and businesses are still experimenting with how [generative AI](#) can increase their productivity. Nonetheless, the biggest tech companies are putting record amounts of money into capital spending, primarily for the hardware needed to develop and run AI models.

"The risk of underinvesting is dramatically greater than the risk of overinvesting," Sundar Pichai, chief executive of Google parent [Alphabet](#), [said on an earnings call](#) in July.

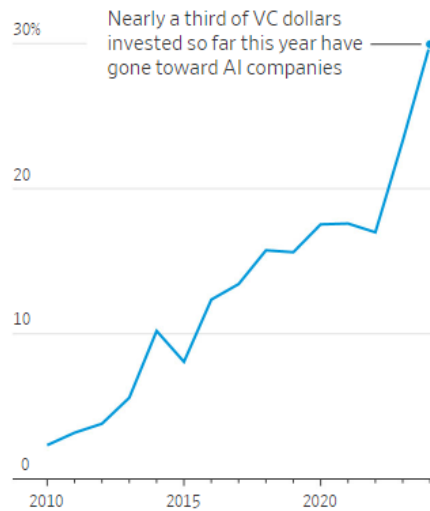
Venture capitalists are similarly betting that at least a few AI startups could one day be worth billions or even trillions, even though most currently aren't profitable. Venture-

capital investments in AI startups are at \$64.1 billion so far this year, putting them on track to approach a peak set during a broader investing upsurge in 2021. And the total share of VC investments going toward AI this year is the highest on record.

Venture-capital investment in artificial intelligence startups...



...as a share of total VC deal value each year

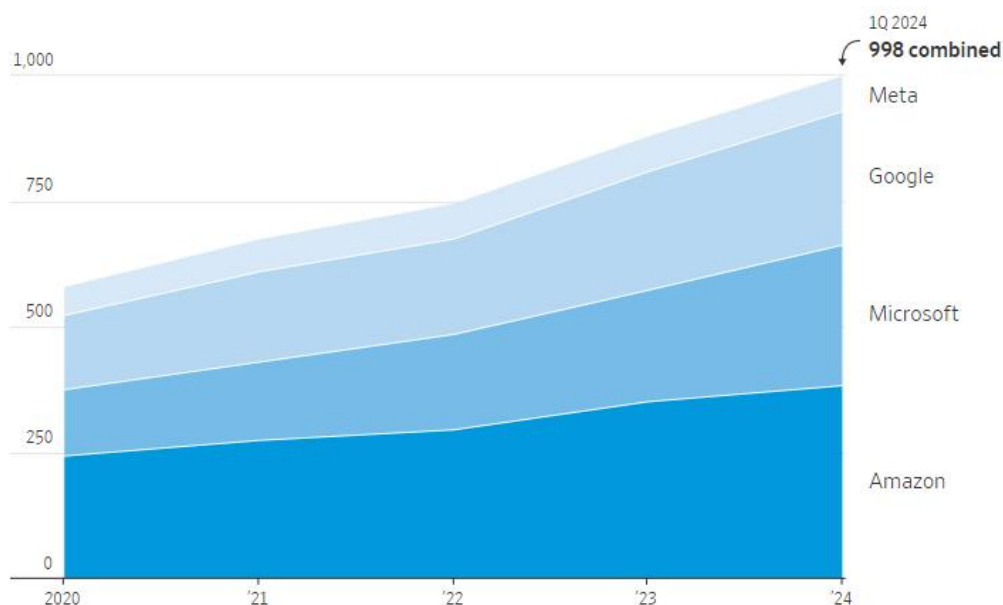


Note: 2024 data as of Aug. 27
Source: PitchBook

The fruits of all that spending can be seen around the U.S., as new data centers are popping up with increased frequency. In the past, data centers were used primarily to remotely store data and run non-AI software. AI-optimized data centers house specialized chips needed to develop and run generative AI applications.

From early 2020 through this year, Microsoft has more than doubled its number of data centers. Google's total is up 80% over the same period. Oracle is heavily focused on the business and plans to build 100 new data centers.

Estimated number of data centers as of the first quarter of each year

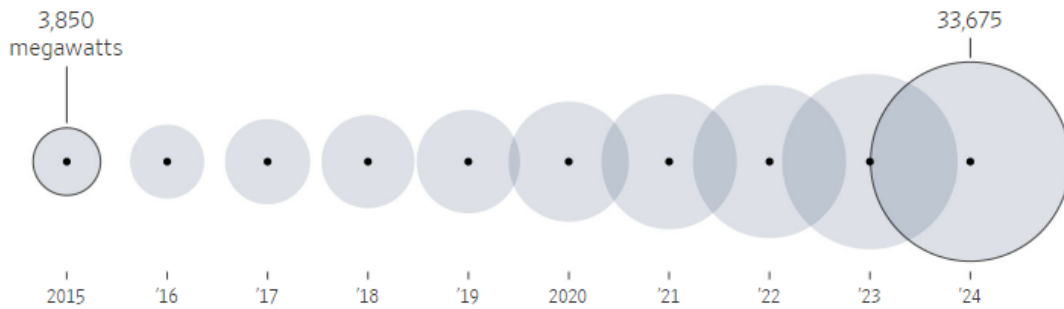


Source: Dell'Oro Group

AI data centers are more power hungry than those built in the past, because AI chips require a constant and reliable source of energy to operate. Even brief dips in power could damage the “training runs” in which AI models improve by analyzing reams of data. For large models, each training run costs tens or hundreds of millions of dollars.

Since 2015, the amount of power that data centers in the U.S. and Canada have ordered from energy companies has increased nearly ninefold.

Commissioned power for U.S. and Canada data centers, yearly

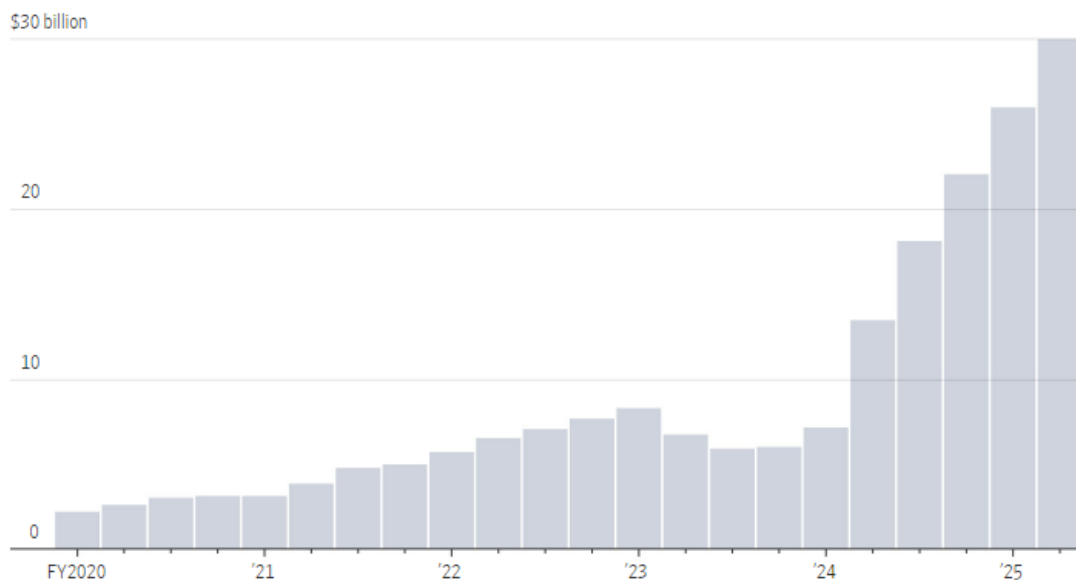


Note: 2024 figure is a full-year projection
Source: datacenterHawk

The chips used to train and operate AI models primarily come from one company: [Nvidia](#). Its graphics processing units, originally designed for videogames, cost tens of thousands of dollars each on the high end. Tech companies building and hosting AI models vie to get the biggest possible allocation from Nvidia.

[Meta Platforms](#) CEO [Mark Zuckerberg](#) has said his company plans to have 600,000 GPUs by the end of 2024. [Tesla](#) chief [Elon Musk](#), who is building his own AI startup xAI, said he hopes to have 300,000 GPUs by next summer.

Nvidia's quarterly revenue



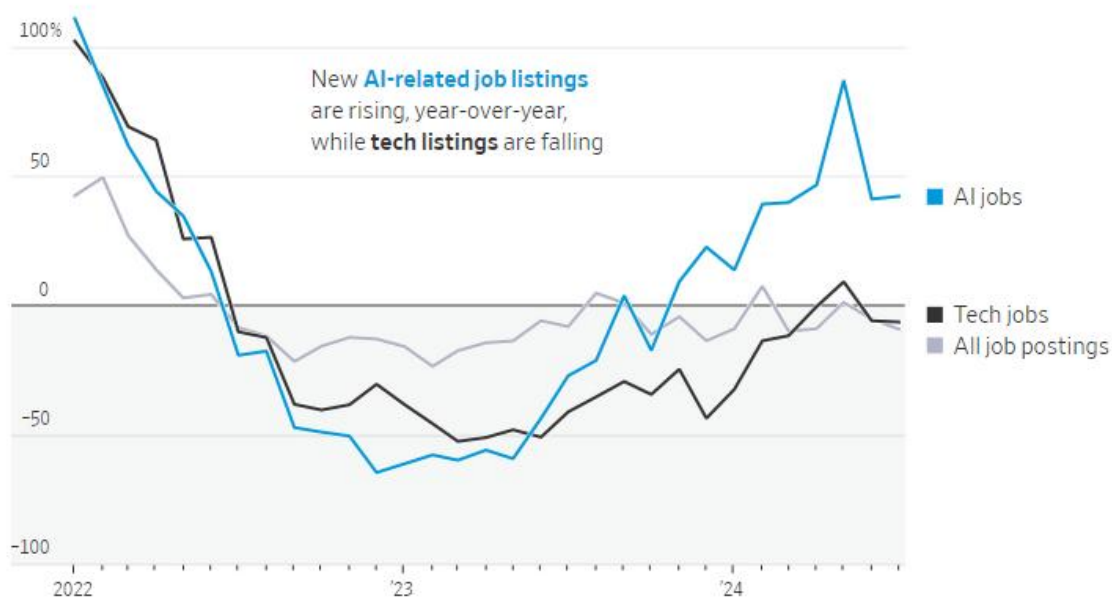
Note: Latest fiscal quarter ended July 28
Source: the company

Highly skilled talent is another scarce commodity. Despite [widespread layoffs](#) that have hit Silicon Valley recently, tech giants are spending many millions of dollars to lock down research scientists who they believe can push AI into new frontiers. Many of these

experts until recently would have worked in academia. Now they are among the most richly compensated technologists in the world.

Even people with a more basic understanding of the machine learning that underpins AI have their pick of six-figure jobs. New listings for AI-related positions in July were up nearly 50% compared with last year, while postings for tech jobs overall were down slightly.

Newly listed U.S. jobs, change from a year earlier



Note: As of July
Source: UMD-LinkUp AI Maps

Investors' patience with Silicon Valley's outside AI spending is likely to wane. They have already penalized the stock of [companies such as Meta](#) and [Microsoft](#) for increases in AI spending without fast-enough revenue growth. A partner at the venture-capital firm Sequoia Capital recently calculated that to justify this year's investment in data centers and chips alone, AI businesses will ultimately need to generate \$600 billion in revenue. Though most companies don't disclose their revenue from AI, analysts have estimated the annual total is at most in the tens of billions.

Worries about whether AI backers have gotten ahead of their skis hark back to the dot-com era a quarter-century ago, when companies poured cash into fiber-optic networks to support bullish estimates of internet use that took [longer to develop](#) than expected.

Executives at the top tech companies are preaching patience. In the recent earnings calls, Zuckerberg said it would be years until AI apps are monetized, and Google's Pichai said "there is a time curve in terms of taking the underlying technology and turning it into meaningful solutions."

Write to Nate Rattner at nate.rattner@wsj.com and Tom Dotan at tom.dotan@wsj.com

Leia o artigo na origem em:

https://www.wsj.com/tech/ai/artificial-intelligence-investing-charts-7b8e1a97?mod=hp_listb_pos1