Too Much AI, Too Soon

UNCOVERALPHA
OCT 09, 2025

I decided to share some thoughts about the current state of the market regarding A

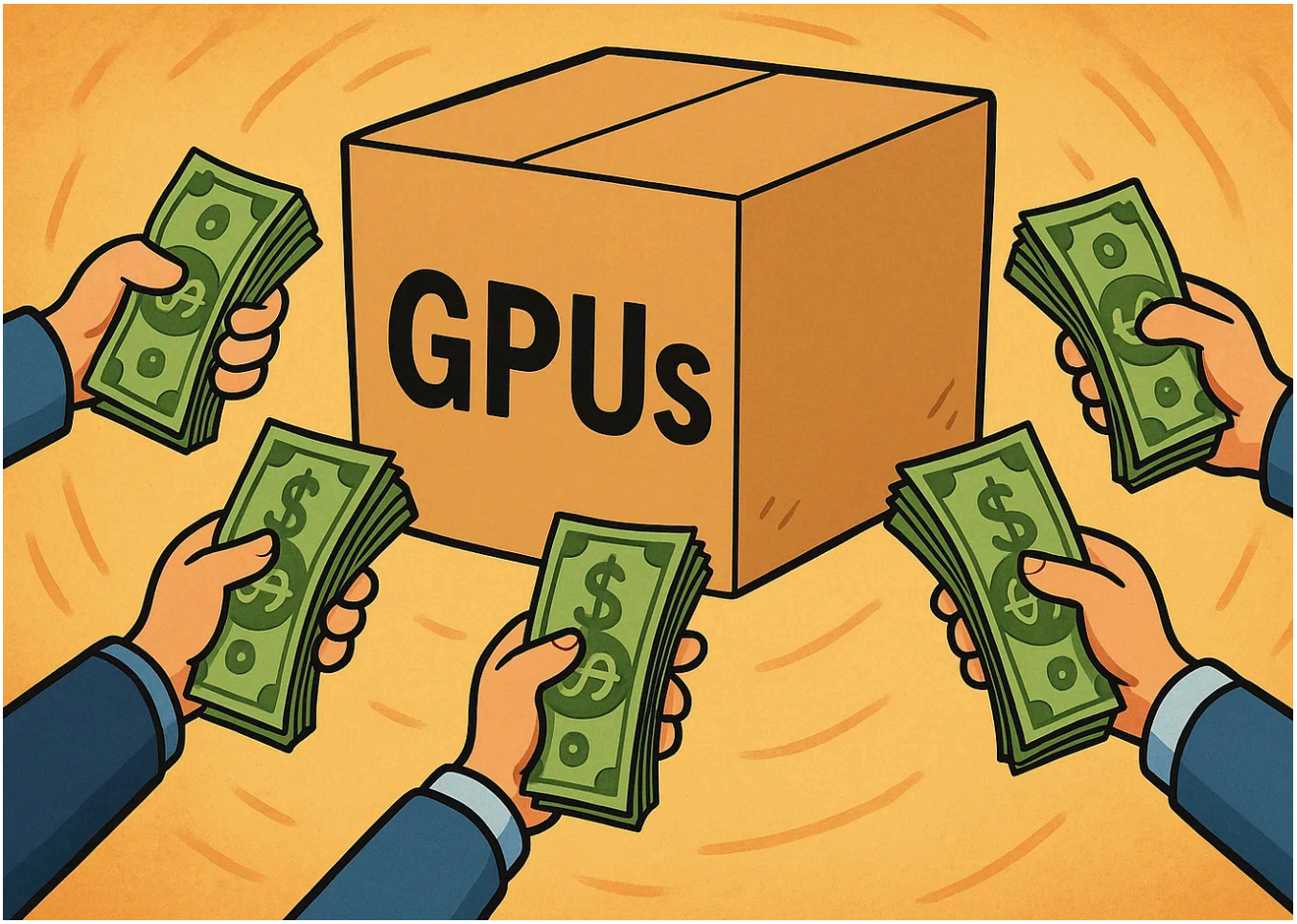
I have become very cautious due to recent financing developments, the projected amount of capital to be raised, and the general valuation levels of many of these companies. As a result, some of you already know, I have trimmed or sold many of tech/AI positions lately.

I want to start off and say that my conviction in AI in the long term has not chang one bit. I continue to think it will be the biggest transformation in history for soci and the economy. That being said, I believe the stock market expectations in the s term have gotten ahead of the reality that we face. The issues I see can be segment into the following categories:

1. We are running out of organic capital and entering the phase of »creative« deals

2. GPUs are a faster depreciating asset than what is thought

3. Valuations are factoring in a very small chance of things slowing down

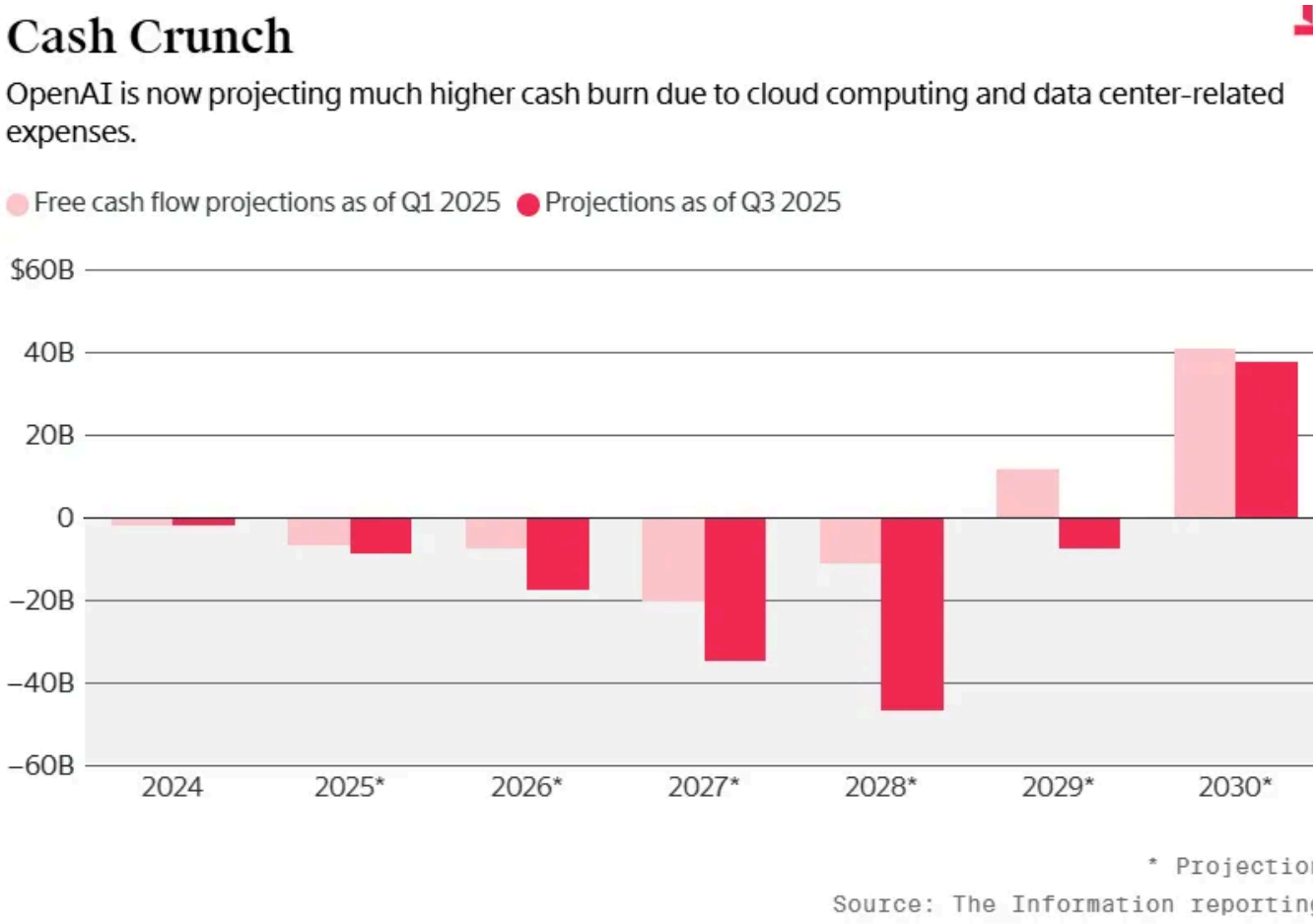


We are running out of organic capital, so »creativity « has taken front stage

When funding unicorn startups, the classical setup used to be: big VC firms at $1B-$10B; at $10-$30B, someone like Softbank, and then you do an IPO. But AI lab like OpenAI and Anthropic don't want the IPO route for now, as going that way means your business model and economics get dissected, and analysts dive deeper what makes sense and what doesn't. Even if they did an IPO, it wouldn't raise near enough capital, as now we are entering a stage where AI labs need over $100 billio new investments on an annual basis. OpenAI, with its deal with Nvidia and AMD top of their Stargate datacenter, plans to build a total of 26GW of data centers in th next few years. And this is just a current number that we know so far. One GW of AI data center costs around $60B, so we are talking about raising more than $ 1.5 trillion. To put that number in perspective, the most profitable business models fr the big tech companies, Amazon, Google, Meta, Microsoft, and Apple, in the last 5 years produced a total of $1.4T in Free Cash Flow. And this was in a pandemic environment where usage and profits soared from the increased demand. So now v

are talking about a company needing to raise more than the combined 5-year Free Cash Flow of Big Tech.

OpenAI is on track to make around $15-$20B in revenue this year. Even if that nur doubles or triples next year, it is not even remotely enough to justify the investmer size, so OpenAI will, of course, have to continue raising capital and possibly debt. top of that revenue, they are expected to lose around $9B, with losses continuing t rise to $47B in 2028. Bloomberg also reported that xAI, another AI lab, is losing around $1B per month.



Financing trillions of CapEx via their own FCF will be very hard to do, so it's clear OAI will need to raise capital and debt, but who is big enough to put down over $ billion?

Nvidia decided to potentially invest $100B in OAI, structured as $10B for each GW power that OAI brings. To me, this deal is concerning, especially as we are now entering a phase where Nvidia is the only possible financier for these types of deal

the FCF of everyone else is already depleted from investing in heavy CapEx to bui datacenters and buy Nvidia chips. In my view, the main reason Nvidia did this dea that, as we will discuss later in this article, OpenAI is key to the entire AI sector ri now. Dylan Patel from Semianalysis recently said in a podcast that OpenAI and Anthropic are the end buyers of 1/3 of all Nvidia GPUs right now.

The problem is not just the circular type of deal itself; the problem is that, at a size over $100B, the only possible investor is a company like Nvidia or perhaps Apple. the hyperscalers have all of their cash already committed to their own CapEx, whi hitting the $70-$100B annual range. Even the hyperscalers are on their limits wher comes to spend as the CapEx is rising much faster than the revenues and FCFs. O of it, even with Nvidia's $100B OAI, it still needs $1.4T? Who will finance that?

Another concerning sign for me is that debt financing has started to roll in in thes deals. We just got information on $20B financing for xAI, where the $20B is provic by an SPV. Of that $20B, $12.5B is debt and $7.5B is capital, with Nvidia contribut $2B of the $7.5B. xAI will then rent those chips from the SPV for 5 years, where th GPUs act as collateral. Meta has also raised $29B for a data center recently, with $ of that $29B being debt, and that data center is expected to be the collateral. Oracl has also completed a $38B debt raise. Hyperscalers like Microsoft and others are g to the neoclouds. Nebius has signed a $17.4B (potentially expandable to $19.4B) de with Microsoft. Looking at the deal, debt is in play again:

»Nebius expects to finance the capital expenditure associated with the contract through a combination of cash flow coming from the deal and the issuance of debt secured against the contract in the near term, at terms enhanced by the credit qua of the counterparty. «

So we have entered a market stage where debt and a company like Nvidia, with its motivations, act as the lender of last resort.

The creative deals where chips are collateral are also a big problem, as I will expla further here. I expect to see more of these GPU collateral deals until the market fi figures out the problems with those.

GPUs are a faster depreciating asset than what is thought

The life of the current generation of GPUs is shorter than most think, and what m companies are projecting in their amortization plans. We are entering the inferen phase of the AI cycle, where we are running out of data centers and energy. The m important metric has become tokens per watt. Nvidia has also moved to a 1-year upgrade cycle, which means that each year you will get a much more capable and energy-efficient accelerator than the previous generation. And this is not at the sc of anything we had in history with Moore's law and chips. Jensen said it himself: between Hopper and Blackwell, they are driving the cost of tokens down 10 to 20x Moore's law would have achieved that by just 20%, so this is much faster, and the amortization of these GPUs should be much, much faster than what the neoclouds hyperscalers are modeling. On a recent podcast, Jonathan Ross, the CEO of Groq a one of the founders of Google TPUs, said that at Groq, they are using a 1-year cycl terms of amortization, as the people who are using 3-5 year amortization cycles ar wrong. With chips, you don't just have an upfront investment in CapEx; you also i the OpEx of running that chip, along with the electricity and water costs that com with it. Not to forget, electricity costs are going up because these AI factories requ a lot of electricity. Looking at the statements and financials of Neoclouds and hyperscalers, you can see that their numbers differ. The hyperscalers follow a 3-4 y amortization cycle for GPUs, whereas Corewave and some neoclouds follow a 6-ye depreciation of Nvidia GPUs, as stated by their leadership. The losses on these neoclouds would have been much, much bigger if the amortization cycle were 1-2 years instead of 6, which is another concerning pressure point in the whole ecosys

But some might say, well, you still see people renting Nvidia H100, which are chip that Nvidia started selling 3 years ago. Yes, but there are two factors to that. The f one is that you have two clients pushing demands sky high, as they are subsidizing end users, as the computing to do the services that they offer is much more expens than the price that they are charging the end users. This works out only to the poi

where investors are willing to give you the money to continue doing that. And the second, even more important point is that the H100 is still useful despite being 3 y old, because NVDA switched to a 1-year product cycle between H100 and Blackwe so this is in late 2024. Before that, the cycle was 18-24 months. So, in terms of cycl times, the chip isn't that old from a generation perspective compared to looking a in years. However, with Nvidia now on a one-year product cycle, this change affec things significantly. In my view, the real amortization of these chips should be in 1 years.

For the sake of math, let's take Coreweave's amortization of 6 years. This means th when Nvidia Vera Rubin comes out in late 2026, people will still want to rent Amp A100, which started shipping in late 2020. That is crazy and not going to happen. I the hyperscalers ammorization rates of 3-4 years are a stretch in my view, especial we go to a world where we don't have any »free« AI data centers waiting around, w have to build new ones and get new power to them, which takes time, so for all the comapanies that will want to scale they will have to switch up their old GPUs at th data centers they already have running for new GPUs to get more tokens per watt their watt usage is limited.

The problem with extending your amortization cycle is that it shows higher profit today than they really are. So here is another concern of mine, as the profits of all hyperscalers in the cloud space are going to come under pressure as the true amortization rate shows up in the coming years. It becomes a broader industry problem when investors start to focus more on this and see that the neoclouds are losing even more money than they state. Again, for the AI circle to continue, inves need to pour money into these neoclouds as well…

Nvidia is well aware of this problem, so this comment from an NVDA employee doesn't come as a surprise:

»…taking out the old ones and put in the new ones, and those old ones we'll actually buy b If a customer has A100s and they want to go to H100s, we'll buy back the servers and the and then resell them overseas.

Source: AlphaSense

My speculation is that overseas means China, but now that they can't sell to the Chinese market, the question is, who will buy these old chips? At the end, someon has to take those useless chips on their books. And Nvidia is already committed tc taking on some of these potential problems if they arise. As recently reported in a CoreWeave deal, Nvidia is obligated to pay the company up to $6.3 billion through 2032 if the cloud provider has unsold capacity. The agreement was actually signed 2023, but was only publicly revealed in an SEC filing this month. So Nvidia is alrea acting as a backstop to some extent, although Coreweave's debt by itself is much l than the $6.3B.

Why do you think Microsoft is doing deals with Neoclouds? Because they are seei surge in demand for compute from their clients. Microsoft wants to maintain the client relationship and keep the client happy, but they are not confident enough ir CapEx growing even further, so they would rather offload some of the risk to some else. The client doesn't know or care that Microsoft doesn't own the physical infrastructure, and when the hype fades, Microsoft doesn't have to write those old chips off as a loss, since the neoclouds have taken over that risk. It's a win for Microsoft as they keep the client, and if the demand turns out to be durable in the term, they have more than enough time to build out their own data center and swi back to their own infrastructure. In the meantime, in the frenzy cycle we are in rig now, they can offload the risk of chips becoming obsolete faster than expected. Or the main reasons Microsoft wants to work with neoclouds is that they are uncerta about CapEx and prefer to take OpEx.

On top of everything already stated about these creative deals, we are now even dc deals where GPUs are in SPVs that serve as collateral. As already stated before, if real GPU depreciation rate is 1-2 years, which I believe is correct, then the collate on many of these deals will be a problem.

Valuations are factoring in a very small chance of things slowing down

Current valuations of many of the technology companies are factoring in very littl risk. First is the customer concentration risk. Groq CEO said that 35-36 companie currently responsible for 99% of token spending in AI right now. And even among those 35 companies, 2 are by far the most significant spenders: OpenAI and Anthr We already mentioned the stat from Patel that 1/3 of Nvidia GPUs end up going to OpenAI or Anthropic. The demand from these two companies is reflected not onl Nvidia but also throughout the semiconductor chain and in the revenues of hyperscalers and neoclouds. This means that a big chunk of the market is depende on the success and progress of these two companies.

Both OAI and Anthropic need to continue growing at a very high clip, in terms of users, user engagement, and model performance. In addition, both of these compa (OAI & Anthropic) have to continue to raise enormous amounts of new capital at +$500B valuations, which we already talked about, and it is going to be very challenging, to say the least. We haven't even mentioned the rate of progress of th models. I am not an AI tech skeptic, but I believe that, as with anything, there is a of things not working out. Right now, the market is pricing in a perfect execution future roadmaps. It is also telling that Microsoft, which had complete access to O (even their IP), chose not to fulfill OAI's future compute demand needs at the rate wanted. Keep in mind, Microsoft has rights of first refusal, meaning they could ha the Oracle cloud purchase order if they wanted it, but they didn't. One has to at le think about why that is. Microsoft's Satya has, over the years, proven to us that he one of the best CEOs out there.

Also, I don't see a future where 5 companies are spending on $100B training runs the next frontier AI. I believe that it will become much more narrow in the future 3 or even fewer players forming the market, which means that a lot of the current compute spend for training is being wasted as they create similar functioning mod and in terms of the model performance layer, the moat doesn't seem to be sticky o long-lasting.

The market is also discounting the risk of disruption for many public technology companies, in my view. When it comes to disruption, everyone thinks only of Goo Search, but this potential disruption has now expanded further. The business mod

of companies like OAI, Anthropic, and xAI are expanding into areas such as socia media, e-commerce distribution, productivity tools, and even cloud infrastructure Information retrieval (the Google Search alternative) is only the first step.

If we just look at the cloud market, most of us, including myself, thought that we would have an oligopoly of Amazon, Microsoft, and Google just a few years ago, a was unimaginable to expect that anyone would raise enough funds to invest +$100 build out an AI cloud infrastructure. Well, today, if companies like OAI actually achieve at least half of what they have in plan, they will have the same, if not even more, capacity than some of those hyperscalers. The direct deals they are doing wi Nvidia, AMD, and SK Hynix also mean they are skipping cloud providers. A curre employee at Nvidia even said that xAI's goal is to actually become a compute prov

---

NVDA employee: *»xAi Elon's company. They're building up a tremendous salesforce. Th probably called me like 10 times in the last 6 months, and they're building out there. They to make a massive disruption...*

Analyst: *They want to become Oracle?*

NVDA employee: *Bingo.«*

source: AlphaSense

We also have companies like Oracle, which are willing to take big risks, with debt OAI orders to build out capacity. We have the neoclouds. So, for the three hyperscalers, if the market doesn't soon cool down in terms of funding for these neoclouds and AI labs, they could face serious competitors down the line.

The flip side is that when the market cools down, hyperscalers with positive FCF have opportunities to buy some of these competitors today, as they might become distressed assets. Nonetheless, the disruption risk with this technological shift is

significant, affecting the entire technology industry, and the valuations currently c not reflect that, in my view.

What also doesn't get enough attention is that much of the spending by current te leaders is not tied to new revenue streams but actually to defend the moats and business models they already have. They are in a race that has gotten out of hand, as Meta's Zuckerberg has recently stated, the risk of overspending a few hundred billion on infrastructure is smaller than the risk of being left out. I agree with Mar this point and understand why all of these companies have to be in this race. How the capital market's job would be to properly reflect that risk in valuation multiple and right now, they are not.

To be clear, I am not calling for a 2000s-like bubble drawdown of +50%, but I do believe that we are reaching financial limits that will cause the market to reevalua some of the multiples it has given to companies today, and that we are about to en consolidation phase. In this phase, it will also become much clearer who has a sustainable moat and what the new business opportunities are.

For AI to reach its economic potential, we need better and more efficient hardwar and more efficient software for inference and training these models. I believe we v get that, but right now the market is in a race with itself, and short-term expectati have gone far too high, especially as we consider that most tech companies are goi to go through heavy CapEx cycles, and profits and FCFs will shrink. On top of tha you have moats being shaken all across the industry, and many will even question only the moats but the capex-light business models, as everyone needs AI infrastructure. The trigger point for stopping this AI race is in the hands of the ca markets. Once they decide we will no longer fund this at this pace, it will signal to both private and public companies that the normalization phase has begun, and I believe we are very close to that point.

As always, I hope you found this article valuable. I would appreciate it if you could share it with people you know who might find it interesting. I also invite you to become a paid subscriber, as paid subscribers get additional articles covering both

tech companies in more detail, as well as mid-cap and small-cap companies that I interesting.

Thank you!

Disclaimer:

Nothing contained in this website and newsletter should be understood as investn or financial advice. All investment strategies and investments involve the risk of lo Past performance does not guarantee future results. Everything written and expre in this newsletter is only the writer's opinion and should not be considered investı advice. Before investing in anything, know your risk profile and if needed, consult professional. Nothing on this site should ever be considered advice, research, or aı invitation to buy or sell any securities.

112 Likes · 15 Restacks

Discussion about this post

Comments Restacks

Write a comment...

Alexander Stahel 7d *Edited*

Liked by UncoverAlpha

"OpenAI, with its deal with Nvidia and AMD on top of their Stargate datacenter, plans to build a 26GW of data centers in the next few years."

26 GW is about the installed capacity of Switzerland, one of the most electrified countries in the if not the most. Took us 100 years.

Hard to guess the replacement cost in today's CHF. An easier data point: UK's Hinkley Point C v get to £50bn for 3.2GW. The construction time will be 13-15 years by EDF, ex permits et al.

If they go for combined cycle gas turbines, they must come from Siemens Energy, Mitsubishi H GE Vernova. All are basically booked out and will pass on huge cost inflation plus a natural mor premium - such turbines are, perhaps, the most difficult tech kid humans have ever created.

Hope: GE wants to double capacity I read. We will see.

Coal plants would likely be the path of least cost resistance, except that there is nobody left to them except the Chinese or perhaps a Russian firm. No plant was build in the US since 2012. Ev longer for Europe which is phasing it out by law.

They are all dreaming with their timetables. They also drive cost inflation with it. Their own vani their biggest enemy.

PS: ordinary people will resist data centers sooner rather than later if their electricity bills goes u etc

LIKE (5) REPLY

1 reply

Matthieu Vandenbussche 6d

Liked by UncoverAlpha

Great analysis Tx!

LIKE (2) REPLY

1 reply by UncoverAlpha

14 more comments...