

# October 2024 Investment Update

Rockets, robots, algos and chips



Michael Frazis  
October 18, 2024



Dear investors and well wishers,

The fund contracted 2% in September, which took us back to +79% for the calendar year-to-date.

I'm hosting a webinar next Monday at 11am Sydney time, [register for access here](#) and please send through any questions.

## Rockets

Elon Musk, that political Rorschach test, has once again outdone himself by catching a falling rocket.

This saves all the weight associated with landing gear, and conveniently returns the booster to where it needs to be for its next mission.



Last week Elon also announced a Tesla taxi, a Tesla bus, and humanoid robots.



Last week Elon also announced a Tesla taxi, a Tesla bus, and humanoid robots.



Remote controlled robots serving drinks

These robots are going to get very good, very fast.

Our world is built for humans, and a robot that does human tasks well should in theory be able to do manual jobs harder, faster, stronger, and longer than any of us. This may only be a few years away.

It raises interesting economic questions.

If you have robots being paid and spending money, does that count towards GDP?



Can growth then uncouple from (soon-to-be-shrinking) population?

And a question for the ages is how one person can lead so many companies executing across so many domains, literally including rocket science and brain surgery.

*"Why don't we try to use the tower to catch it?" he [ELON] asked. He was referring to the tower that holds the rocket on the launchpad. Musk had already come up with the idea of using that tower to stack the rocket; it had a set of arms that could pick up the first-stage booster, place it on the launch mount, then pick up the second-stage spacecraft, and place it atop the booster. Now he was suggesting that these arms could also be used to catch the booster when it returned to Earth.*

*It was a wild idea, and there was a lot of consternation in the room. "If the booster comes back down to the tower and crashes into it, you can't launch the next rocket for a long time," Bill Riley says. "But we agreed to study different ways to do it."*

*A few weeks later, just after Christmas 2020, the team gathered to brainstorm. Most engineers argued against trying to use the tower to catch the booster. The stacking arms were already dangerously complex. After more than an hour of argument, a consensus was forming to stick with the old idea of putting landing legs on the booster. But Stephen Harlow, the vehicle engineering director, kept arguing for the more audacious approach. "We have this tower, so why not try to use it?"*

*After another hour of debate, Musk stepped in. "Harlow, you're on board with this plan," he said. "So why don't you be in charge of it?"*

Human nature

## Markets

The last three months have been choppy. Risk management bailed us out of a number of positions, transferring what would have been substantial losses onto someone else's balance sheet.

Managers have long-gripped about the difficulty of trading in a market dominated by algos. And that was pre-ChatGPT. The situation today is significantly more severe.

Active fund managers used to underperform the market by 1-2% on average (to a first approximation they *are* the market, minus costs and fees).

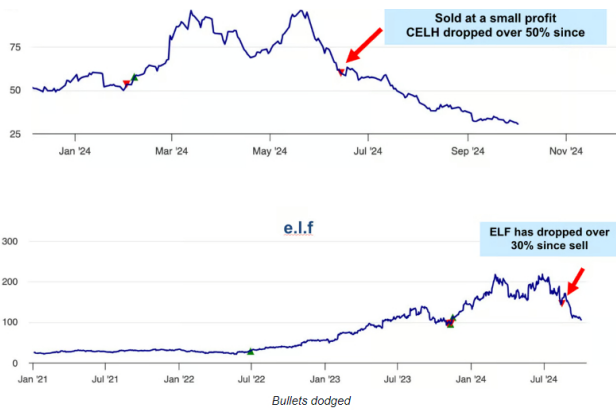
But in recent years, firms like Citadel, which handles roughly a quarter of US equity volume, and quant shops like Jane Street, which Hoover up the smartest Olympiad kids, have stripped additional fortunes out of the active management profit pool, like sharks feeding off an ever-shrinking pod of floundering, discretionary whales.

If you buy or sell a US equity, there's a decent chance a Ken Griffin algo is on the other side.

This is the future, for better or worse.

From our side it's been humbling to see how much these tools have improved our own decision-making, managing our largest wins, and protecting those gains when markets turned over the last few months.

Now, these same models are giving buy signals across the board, both in semiconductors, and for the first time in quite a while, software.



We hired [Saurabh Bhatia](#) as founding CTO in our venture studio to help build out our software capabilities. Saurabh has over 20 years experience at some of Australia's leading startups, most recently at the development agency Paloma.

He has worked on Aussie household names like Afterpay, flatmates.com.au and Airtasker. And along with the venture studio work, is helping build out our internal tooling, some of which we will make available online.

Now that Alphasense, a slow, poorly designed platform, funded by 'growth' funds, has purchased and digested my two favourite platforms, Sentio and Tegus, we will have to rely on our own tools more and more.

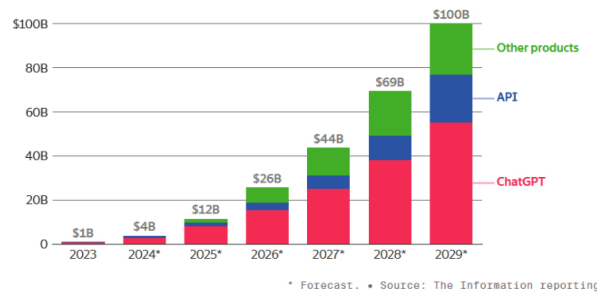
And if you want to build an MVP, or your own internal tooling, reach out and I'm sure [Ace Venture](#) can make it happen.

### Open AI

OpenAI closed a round valuing the firm at US\$157 billion.

### OpenAI's Bold Revenue Forecasts

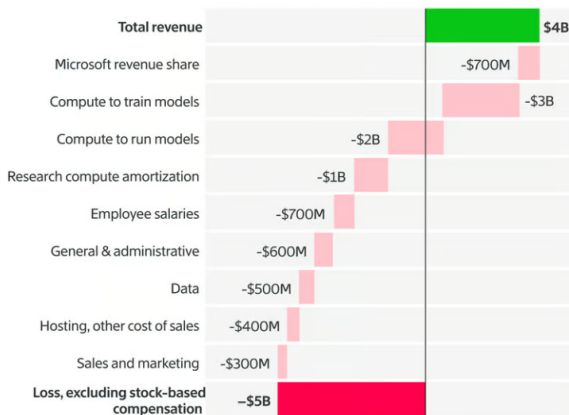
The company projects annual revenues to increase 100-fold by 2029, driven mostly by ChatGPT, and later, other products like video generation, search and other software.



There was interesting detail around their cost structure:

### OpenAI's Projected 2024 Costs

The company expects to lose \$5 billion this year excluding stock-based compensation.



This gives roughly \$1.3 billion of gross profits, subtracting Microsoft's share and compute running costs from revenues.

Gross margins are not great - but in the company's favour it almost certainly has serious unexplored pricing power.

ChatGPT offers *significantly* more value than software like Bloomberg or Salesforce, which charge around ~100x more. ChatGPT is only US\$300 per year.

Underpricing a product and operating at a loss doesn't always work (witness the various VC blow ups of recent years) but in this case it makes clear sense, helping make their tools as ubiquitous and indispensable to users as fast as possible.

### Neoclouds

The rise of neoclouds is captivating the investment world. Companies like Coreweave and Lambda Labs have made major purchases of Nvidia's latest GPUs, backed by leasing contracts with the world's largest tech companies.

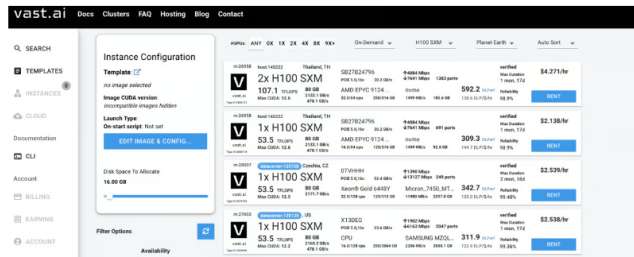
By taking on balance sheet risk when others want to run costs through their expense lines, they've been some of the fastest growing companies in the market. But once contracts roll off, as owners of the equipment, the pricing risk is theirs.

Even Microsoft rents Nvidia GPU capacity from Coreweave, which then uses those long term contracts to buy GPUs from its own shareholder, Nvidia. Similarly, Oracle is leasing GPU capacity off Crusoe Energy to lease to Microsoft and OpenAI.

This circularity has attracted the attention of a certain kind of short-seller - and there's likely truth to the criticism.

There is now a flowering of business models and creative finance. Venture Capital firms are even [setting up their own GPU clusters](#), for exclusive (or at least priority) use by portfolio companies.

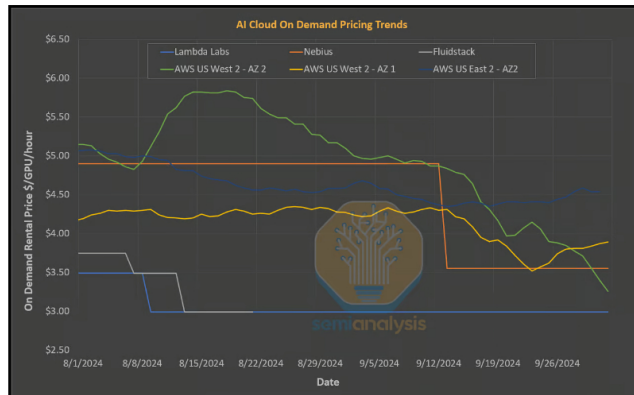
GPU rental costs are tracked online by aggregators like vast.ai, so we can see in realtime how the different clouds are faring:



Leasing rates and secondhand prices were real-time indicators of demand during prior GPU booms and busts, and I'm sure will play the same role this time around.

GPU rental prices are definitely going down, so there is clearly excess capacity being dumped on the market. Payback periods for H100s are now well over a year, so any buyers must take the risk on their future value.

GPUs last 4-5 years, but now that Nvidia has moved to an annual upgrade cycle, practical obsolescence will come faster. This excess capacity will make inference extremely cheap in coming years.



Excess capacity at the cost of speculators is great for the rest of us.

By some counts, inference is already 99% cheaper now than it was a year ago.

Finance is mindblowingly good at getting things built, even as speculators are periodically wiped out.

The market is still nascent and most market participants haven't heard of one another.

the market is still nascent and most market participants haven't heard of one another, meaning the shopping, advertising and discovery/deal flow process involves a lot of Google Ads (many Neoclouds purchase keywords on their competitors' names, wisely avoiding paying up for the surely expensive "H100 rental" keyword!), cold email, word of mouth, browsing gpulist.ai, twitter marketing, friend recommendations, SemiAnalysis recommendations, or even referrals from Nvidia if you are a cloud partner. A typical shopping process could involve buyers connecting with as many as 20 providers (we know over 30 that we shop deals to), ultimately exploring pricing with a subset of this and conducting some trial deployment on shortlisted Neoclouds to ensure proper configuration and reliability.

*You can expect exotic financial innovation and old lessons learned.*

At least one question is now reasonably settle: only the largest companies will train foundational models.

But inference demand is continually being revised upwards. Inference compute on a singly queries could feasibly go up 10, 100, or 1000x from here.

To give one example of how much this might increase, instead of simply choosing the highest probability next token, language models could explore more and more of the possible answer space, like looking ahead a hundred moves or more instead of one.

And the greenfield opportunity in applications is unparalleled, perhaps since Salesforce and that generation of software companies built enormous businesses on databases.

Today, the vast majority of people and businesses don't use LLMs, and we have not even *begun* the era of agents, which will be yet another step change increase in demand for compute.

Along with healthcare, semiconductors remains one of our highest conviction bets.

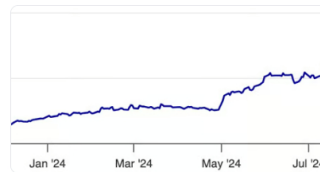
If the season changes, we'll be using tight risk management to protect our portfolio. Best case scenario, we will rip another 4x out of the sector.

Michael

### Keep reading



Clarity Pharmaceuticals with Chairman Dr Alan Taylor



September 2024 Investment Update  
The first rate cuts, a new scaling vector, and healthcare booms



China rally - is this one for real?