

OpenAI drama – Our read, and implications.

Research Analysts

Pierre Ferragu

+1 646 681 4616

pierre@newstreetresearch.com

Rolf Bulk

+65 6016 5450

rolf@newstreetresearch.com

Antoine Chkaiban

+44 20 7375 9126

antoine@newstreetresearch.com

**Benjamin Harwood**

+1 929 379 4139

ben@newstreetresearch.com

Derek Mogull

+1 646 681 4638

derek@newstreetresearch.com

Piyush Kalyani

+44 20 7375 9145

piyush@newstreetresearch.com

Tobias Bähr

+44 20 7375 9111

tobias@newstreetresearch.com

We summarize in this short note our understanding of what happened at OpenAI since Friday and give a framework to understand implications for Microsoft and beyond.

The facts

When founded in 2015, OpenAI was meant, according to its charter, to “avoid enabling uses of AI or AGI that harm humanity or unduly concentrate power”¹. It was designed as a non-profit organization, controlled by a board, under the best (but naive) intentions of developing AI carefully, away from the pressure of profit-seeking villains. The main underlying ideology was that people developing AI at the bleeding edge had to be very careful about alignment (making sure AI has good “intentions”), before releasing it for public use.

In 2019, under the pragmatic leadership of Sam Altman, and realizing that developing AI requires colossal financial means, mostly to be spent on Nvidia GPUs, OpenAI went, in broad daylight, into the most absurd direction of keeping its non-profit status but welcoming investors in a for-profit subsidiary... fully controlled by OpenAI and its board. In 2019, Ilya Sutskever, the chief scientific officer of OpenAI and the leader of Friday’s coup, said in a documentary in 2019: “*It is pretty likely that the entire surface of the earth will be covered with solar panels and datacenters*”. You need money to do that, and non-profit is not the best way to get money.

In the last 8 months, 3 members of the board of directors of OpenAI (out of 9) resigned, for various reasons, which, in our view, all boiled down to the same underlying root issue: The non-profit setup doesn’t make sense here. Reid Hoffman went to found his own AI startup, Shiron Zilis probably wondered what she had to do with the whole thing if Elon Musk was not involved anymore, and Will Hurd went back to real politics.

In recent months, tensions, which had built up for years, between “pragmatics” who wanted to build commercial momentum and scale over OpenAI technological achievements, and “prudents”, who didn’t want commercial consideration to interfere with their fear of misalignment, escalated, to the point that out of the remaining 6 members of the board, on Friday, 4 ousted the other 2, Sam Altman and Greg Brockman, who were working hard at raising capital at an \$80-90bn valuation, around three times the one of the last round².

On Saturday, 3 senior OpenAI researchers resigned, and the press came out with reports that the OpenAI board was engaging with Sam Altman to reinstate him.

As we write these lines, the story is still unfolding, to our greatest delight! For what happens in the next few hours and implications, see our analysis on the next page.

¹ <https://openai.com/charter>

² [WSJ](#) reported details of the latest discussions in September.

Our analysis

Open AI, as it existed on Thursday, is dead.

We are surprised to be alone in noting the preposterous contradiction underpinning OpenAI: four individuals with absolute unconstrained power over an organization owning the most advanced AI technology and defending in its charter that it is meant to avoid enabling AI or AGI that unduly concentrates power.

This model, now that its implications are clear for all to see, is untenable. Our expectation is that either the quartet will be forced to resign, or OpenAI staff will run away, and the quartet will be left by themselves.

What caused the coup?

Speculations on that front have been wild in the last 36 hours, from: “they have nailed AGI and disagree over what to do with it”, down to “it is an immature board, prone to factionalism”.

Our view, supported by an extensive review of what people close to the situation have said in the last few weeks, is that tensions grew over the pace at which Sam was pushing the deployment of more advanced products. It frustrated those keen to preserve resources for research, especially alignment research, as OpenAI became severely constrained in terms of compute capacity, after the product launched at the recent DevDay event. These tensions likely resulted in frustrating communications and Sam making some operational decisions without keeping the board fully aware. The coup, and the sibylline associated blog post, about Sam not being “consistently candid in his communications with the board, hindering its ability to exercise its responsibilities” resulted from this situation.

We would note a few important views underpinning this view.

First, the above is very consistent with what we know of human nature. The quartet was left with unreasonable power, and very limited visibility and glamor, compared to Sam. Only very few individuals can resist that type of dissonance. It had to happen, in the lexical field of Star Wars, we call that falling to the dark side of the force.

Second, human nature, again, tends to imagine the wildest reality, when reality is not visible. (This is the root cause of all conspiracy theories.) It is tempting to imagine OpenAI on the brink of an inflection point or other singularity, in terms of product development, but

it is, based on what we know from publicly available information, objectively unlikely.

Microsoft is at risk.

The OpenAI deal was transformative for Microsoft and propelled the company to the forefront of the generative AI revolution. We find the OpenAI-Microsoft relationship very powerful: OpenAI has the most advanced generative AI technology, and Microsoft has the most interesting product portfolio, when it comes to tangible, near-term monetization of generative AI. A winning partnership, no doubt.

The only issue is the non-profit status of OpenAI and the risk it represents to Microsoft. As the board of OpenAI controls the non-profit and the for-profit entities, it controls collaboration with Microsoft, and therefore the future of Microsoft.

Microsoft has to act now to ensure this structural risk is removed. This reality must be already fully understood by management, who, we imagine, are actively engaged in current negotiations. Investors owning Microsoft should forget about the delusion of what the value of a stake in the for-profit OpenAI could be, and carefully monitor that Microsoft eliminates its dependency on the non-profit board. Now is the time to address this issue!

What does this tell us about the future of AI?

The ambition to develop AI as a non-profit in order to control its inherent risks is naive. It can only grow in the mind of people who have not witnessed the reality of non-profit organizations, which are, from a governance standpoint, the most recklessly “political”, in the most negative sense of the term: a fight for the maximization of individual power.

If AI technology is kept inside a non-profit controlled by a board of four, it is bad. For instance, alignment matters would be controlled only by these four people. If AGI is achieved within that institutional structure, it can only end in a totalitarian way, as these four people would decide what is good and what is bad. AGI alignment, if not public and transparent, is totalitarian.

Our view is that events of the last few days will make these realities clearer to everybody and hopefully lead to a healthy competitive and transparent market, in which pressure to open-source will increase, and regulation efforts fructify over time – We all know

regulation never comes ahead of the curve, and we have to live with that other anthropological reality.

Our conviction is that this development is a good thing, on two counts.

First, it is the best way to make AI safe. At the end of the day, AI is a tool, or, when it comes to detrimental use, a weapon. Anyone understands a weapon is most dangerous if only one side of a conflict has it. If it is widely available, it deters everybody from using it. It ought to be the same for AI, oughtn't it?

Second, it is the best way to make AI successful. As pressure mounts to make AI competitive (non-monopolistic) and open-source, it appears obvious that it will accelerate its developments. We cannot find a single technology which, being developed both in an open, transparent, competitive environment, and in a closed, non-profit environment, did better in the latter. The invisible hand is still at work, even more so for AI.

Covering the earth with datacenters and solar panels.

Last but not least, and what matters the most to our coverage universe: developments of the last few days are excellent for datacenter spending.

Between 1960 and 1973, the US spent the equivalent of ~\$20bn per year in today's value of the currency³ on the Apollo program. This was the cost to conquer the moon. Microsoft alone will likely spend over \$50bn in capex per year in the foreseeable future to conquer generative AI. We are witnessing the largest-scale investment humanity has ever made in a new technological frontier.

As technology business leaders digest the events of the last few days, they will understand the importance of avoiding any excessive dependency in the world of AI, and this will inevitably lead to a multiplication of investments in infrastructure.

Ilya, the one willing to slow the adoption of AI, might be looking for a new job on Monday morning, and in addition, there is one thing he nevertheless said in 2019: *"It is pretty likely that the entire surface of the earth will be covered with solar panels and datacenters."*

The root causes of the dispute and the coup are also telling. In the last few weeks, OpenAI, after announcing a few new products at their DevDay event, ran out of compute capacity. We have all seen orders at Nvidia and Broadcom for AI chips going through the roof, but we are still dealing with a situation where compute is vastly running behind.

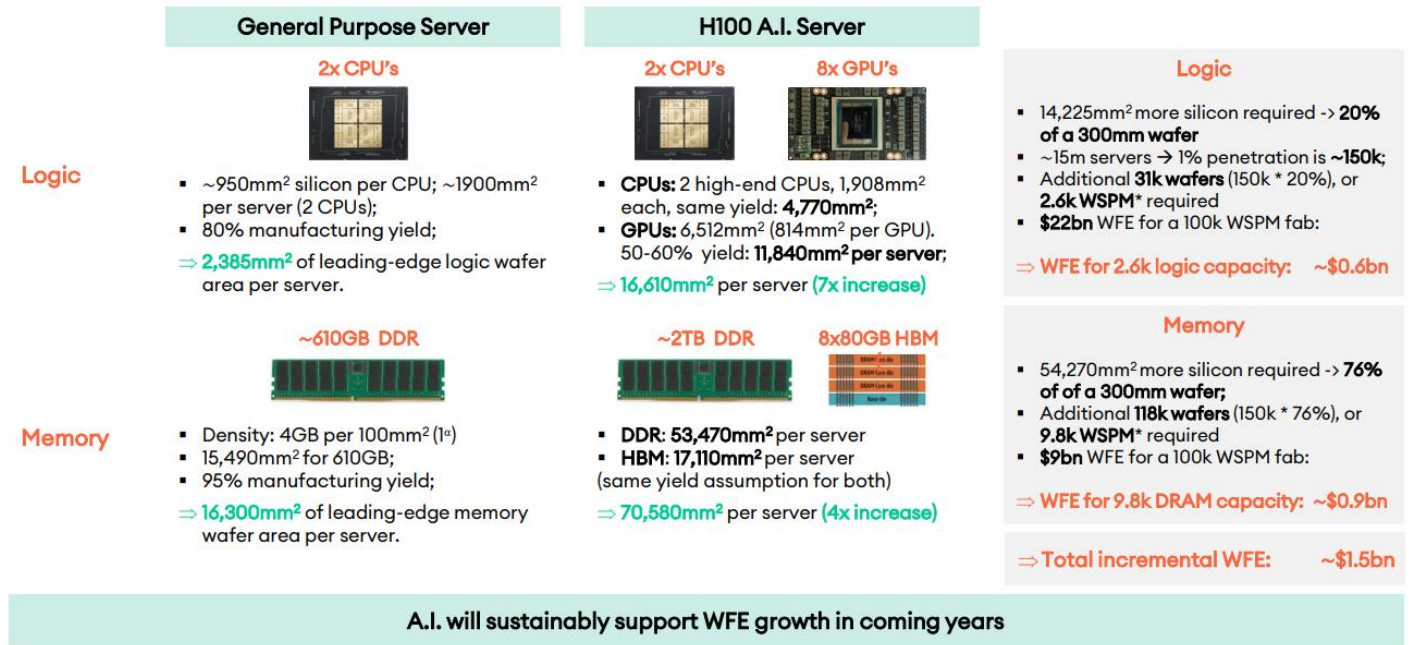
In our coverage universe, most reliable and sustainable long-term beneficiaries are key franchises and near-monopolistic suppliers of the datacenter infrastructure value chain: **Semicap** manufacturers (see [exhibit 1](#) for an analysis on how much a 1% increase in A.I. server penetration adds to WFE spending), **TSMC**, **Nvidia**, **Arista**, **Broadcom**, **Intel**, **AMD** (see our work on [AI clusters deployment](#) for more details), and **Micron** are the most obvious names, but we would highlight also those exposed to power semis: in our coverage, **Infineon** and **Wolfspeed**.

Outside of our coverage, most notable names we have considered in our research are other memory manufacturers **Hynix** and **Samsung**, **Marvell** as an alternative to Broadcom for custom silicon and networking silicon, **Accton** for networking, **Coherent** for optical components.

³ The United States spent \$25.8 billion on Project Apollo between 1960 and 1973, or approximately \$257 billion when adjusted for inflation to 2020 dollars.

Ex 1 – A 1% increase in AI server penetration results in \$1.5bn incremental WFE spending

A.I. servers require 5x more leading-edge wafer area (7x more Logic, 4x memory)



A.I. will sustainably support WFE growth in coming years

Source: Intel, Nvidia, SK Hynix, Micron and NSR estimates and analysis.

* KWSPM = Thousand Wafer Starts Per Month

Relevant Research

- Microsoft unveiled two in-house chips. Our take. – (16 November 2023) - [Link](#)
- OpenAI DevDay: implications for tech infrastructure – (7 November 2023) - [Link](#)
- The latest in advanced packaging - take-aways, replay & preferred slides – (17 October 2023) - [Link](#)
- Who can afford that many AI clusters? - Conference Call Slides. – (2 October 2023) - [Link](#)
- Generative AI: Updated thoughts on the back of our Big Idea Conference. – (25 September 2023) - [Link](#)
- Nvidia: which clients drove most of the surge in datacenter revenues? – (6 September 2023) - [Link](#)
- Nvidia, Arista, Broadcom: How much more room for A.I. revisions? – (15 August 2023) - [Link](#)
- How much does a 1% increase in A.I. server penetration add to WFE spending? – (3 August 2023) - [Link](#)
- Semicon West 2023: Our takeaways – (18 July 2023) - [Link](#)
- What does TSMC doubling CoWoS capacity tells us about Nvidia and AI chip growth? – (14 June 2023) - [Link](#)
- Meta's MTIA? Best illustration that succeeding with an in-house chip is challenging... – (23 May 2022) - [Link](#)
- Cluster economics of Nvidia GPU vs. Google TPU: How different? Implications for market structure? – (22 May 2023) - [Link](#)
- Power semiconductors: Our updated thoughts following the PCIM conference – (16 May 2023) - [Link](#)
- AMD & Microsoft designing a custom A.I. chip? Our quick take. – (5 May 2023) - [Link](#)
- Amazon shareholder letter: Our read on the momentum of in-house accelerators – (17 April 2023) - [Link](#)
- Memory share of Wallet in AI: lower than in traditional servers – (14 April 2023) - [Link](#)
- Is Memory a good way to play the ramp of A.I. clusters and Hyperscale Capex? – (6 April 2023) - [Link](#)
- What drives Hyperscale capex going forward? A deep dive in the economics of Cloud – (27 March 2023) - [Link](#)
- Intel DCAI webinar: our 3 takeaways. – (30 March 2023) - [Link](#)
- Nvidia GTC 2023: a few very interesting things we learnt... – (22 March 2023) - [Link](#)
- ChatGPT early deployments this year: potentially very material for Nvidia – (7 February 2023) - [Link](#)
- ChatGPT - Implications of large language models for tech infrastructure (NVDA, TSMC, ...). – (24 January 2023) - [Link](#)
- Tesla A.I. Day: 7 take-aways you won't read elsewhere. – (3 October 2022) - [Link](#)
- Semiconductor Big Idea Conference 2022: Our 9 key take-aways & slides. – (23 September 2022) - [Link](#)
- Sapphire Rapids, Genoa, ARM: Updated thoughts on server CPU roadmaps – (17 August 2022) - [Link](#)
- Dojo, TPU, Cerebras, IPU, GPU et al. Future of Exascale computing architectures – (25 March 2022) - [Link](#)
- Giant AI chips & computers, quantum... First Semiconductor Big Ideas Conference. - (13 September 2021) - [Link](#)
- Packaging innovation (III) – Networking: the 5-year journey towards the single-chip switch - (24 June 2020) - [Link](#)
- Packaging innovation (II) – Can chiplet architecture challenge the CPU Status Quo? - (28 January 2020) - [Link](#)
- Packaging innovation (I) – the upcoming datacenter GPU war - (16 January 2020) - [Link](#)
- Chip Architecture and manufacturing: Why does Intel have 97% share in Server? - (24 September 2019) - [Link](#)
- Technology Infrastructure Coverage initiation - (03 May 2018) - [Link](#)

Disclosures

12 month historical recommendation changes are available on request

This report was produced by New Street Research LLP, 100 Bishopsgate, 18th Floor, London EC2M 1GT Tel: +44 20 7375 9111
Regulatory Disclosures: This research is directed only at persons classified as Professional Clients under the rules of the Financial Conduct Authority ('FCA'), and must not be re-distributed to Retail Clients as defined in the rules of the FCA.

This research is for our clients only. It is based on current public information which we consider reliable, but we do not represent that it is accurate or complete, and it should not be relied on as such. We seek to update our research as appropriate, but various regulations may prevent us from doing so. Most of our reports are published at irregular intervals as appropriate in the analyst's judgment.

This research is not an offer to sell or the solicitation of an offer to buy any security in any jurisdiction where such an offer or solicitation would be illegal. It does not constitute a personal recommendation or take into account the particular investment objectives, financial situations, or needs of individual clients.

All our research reports are disseminated and available to all clients simultaneously through electronic publication to our website.

© Copyright 2023 New Street Research LLP

No part of this material may be copied, photocopied or duplicated in any form by any means or redistributed without the prior written consent of New Street Research LLP.

New Street Research LLC is neither a registered investment advisor nor a broker/dealer. Subscribers and/or readers are advised that the information contained in this report is not to be construed or relied upon as investment, tax planning, accounting and/or legal advice, nor is it to be construed in any way as a recommendation to buy or sell any security or any other form of investment. All opinions, analyses and information contained herein is based upon sources believed to be reliable and is written in good faith, but no representation or warranty of any kind, express or implied, is made herein concerning any investment, tax, accounting and/or legal matter or the accuracy, completeness, correctness, timeliness and/or appropriateness of any of the information contained herein. Subscribers and/or readers are further advised that the Company does not necessarily update the information and/or opinions set forth in this and/or any subsequent version of this report. Readers are urged to consult with their own independent professional advisors with respect to any matter herein. All information contained herein and/or this website should be independently verified.

All research is issued under the regulatory oversight of New Street Research LLP.

Sales:

Charlie Gaynor

charlie@newstreetresearch.com

Zach Monsma

zach.monsma@newstreetresearch.com

Ethan Lacy

ethan.lacy@newstreetresearch.com

Mark Franks

mark@newstreetresearch.com

Michael Chambers

michael@newstreetresearch.com

Daniel Gilroy

daniel.gilroy@newstreetresearch.com

Dimitri Livchits

dimitri@newstreetresearch.com

Heather Broffman

heather@newstreetresearch.com