

INDUSTRY WATCH

Soap 'n' AI

Robert Dale

Language Technology Group

Email: rdale@language-technology.com

Abstract

It's less than a year since OpenAI's board voted to fire Sam Altman as CEO, in a palace coup that lasted just a weekend before Altman was reinstated. That weekend and subsequent events in OpenAI's storyline provide all the ingredients for a soap opera. So, just in case Netflix is interested, here's a stab at a synopsis of what might be just the first of many seasons of 'The Generative AI Wars'.

Introduction

As I write this in early October 2024, the last couple of months have delivered [signs of rough times ahead for OpenAI](#). There's a sense that whatever technical lead OpenAI once held, others—principally Google and Anthropic—are [catching up](#), leaving OpenAI with no moat; there have been successive reports of [executives leaving the company](#); and Microsoft, the company's principal investor, has indicated that [OpenAI is a competitor as well as a partner](#). All this has been in the context of a more general concern that [generative AI won't be able to deliver on overhyped promises](#), and investors [won't be able to recoup their investments](#). And yet, the company [has just raised US\\$6.6B](#) in funding at a valuation of US\$157B.

It's fascinating to watch. And with only a wee bit of whipping up, there's surely great material here for a docudrama.

If you're developing a TV or movie franchise, it's important to choose its historical entry point well, so as to allow maximum scope for both prequels and sequels. So what better place to start than OpenAI's November 2023 dramatic boardroom coup, which Sam Altman himself [likened to a soap opera](#)? With that as our entry point, here's how the first season of 'The Generative AI Wars' might play out.

The Pilot Episode: One long weekend in November

It's Thursday evening, 16 November 2023. Silence. The downtown San Francisco air is still and just a bit chilly. A drone camera swoops us through the city's concrete canyons, crossing into the Mission District and pulling up at the historic Pioneer Building on 18th Street, home to OpenAI. We focus in on a window of the building, through which we see Ilya Sutskever, OpenAI's chief scientist and co-founder, pull out his cell phone. We watch as he texts Sam Altman, the company's CEO and fellow board member, to schedule a video call the next day.

Fast forward to Friday noon. Sam connects to the scheduled call. Amongst the board members present: Ilya, tech entrepreneur Tasha McCauley, and Helen Toner, the director of strategy at Georgetown University's Center for Security and Emerging Technology. We don't know what else was on the agenda, if anything, but there's one stunning item that is actioned in the course of the meeting: [the board fires Sam](#). CTO Mira Murati is [installed as interim CEO](#) in his place.



Greg Brockman is told that he will no longer be chairman of the board. The company publishes a [blog post](#) that announces the changes, stating that Sam was dismissed because ‘he wasn’t consistently candid in his communications with the board’, and that the board ‘no longer has confidence in his ability to continue leading OpenAI’, but providing no detail beyond that.

And so begins a weekend of chaotic activity. On Friday afternoon, the company holds an all-hands meeting: [Ilya defends Sam’s firing](#), saying it was necessary to protect OpenAI’s mission of making AI beneficial to humanity. Microsoft CEO Satya Nadella posts a statement emphasizing that, as far as Microsoft is concerned, [it’s business as usual](#), and the company’s long-term agreement with OpenAI remains in place—in other words, don’t frighten the horses. But Satya is [said to be furious](#) about being blind-sided by the board’s decision.

On Saturday morning, Greg Brockman tweets that [he is quitting](#). Three other senior researchers [also quit](#). By Saturday evening, investors in OpenAI’s for-profit entity have already discussed working with Microsoft and senior employees at the company [to bring back Sam](#). Sam is reported to be ambivalent about coming back and would want to see significant governance changes; and there are reports that he and Greg are planning a new venture.

By Sunday morning, [the board is in discussions for Sam to return as CEO](#), with Mira and Brad Lightcap, OpenAI’s COO, pushing for his reinstatement. By Sunday evening, Ilya states that Sam will not return; but reports surface that Sam and Greg are open to returning to OpenAI if the remaining board members who fired them step down.

By Monday, over 650 of the company’s 770 employees have [signed a letter saying they may quit](#) if Sam isn’t reappointed. Sam and Greg say they’ll join Microsoft [to lead a new AI research team](#). With a nod to the signatories list, Satya indicates a willingness to employ other OpenAI staff, and Microsoft begins [readying laptops, office space and onboarding procedures](#) at its LinkedIn offices in San Francisco.

On Tuesday November 20, a sudden turnaround: Ilya publishes a post on X saying that [he regrets his participation](#) in the board’s actions, and that he’ll do everything he can to reunite the company. By Tuesday evening, it’s reported that [Sam will return as CEO](#) and several members of the board will depart. OpenAI staff throw an informal party for Sam, but [Ilya is noticeably absent](#).

Phew. We’re back where we started, except for the board membership, of course. The palace coup lasted no more than five days, and SF’s evening air is once more still. For now.

Episode 2: A smell of Musk in the air

It’s March 2024. November’s boardroom shenanigans seem an age away, and a sense of normality has returned to the 18th Street offices. In the intervening period, Andrej Karpathy, another founding member of OpenAI, [announced his departure from OpenAI](#) to work on personal projects, but there are no hints that anything else lies behind his decision. OpenAI, along with dozens of other companies, has [signed a pledge](#) to build AI for the good of humanity, although some see the commitment as [vague and meaningless](#). But nothing much else in OpenAI’s world is drawing attention. Perhaps a special guest appearance is needed to make sure we keep up the audience numbers?

Enter Elon Musk.

Just as nature abhors a vacuum, Elon abhors a lack of controversy. So, Elon, an early investor in OpenAI, spices things up by [suing the company for breach of contract](#): he’s not happy with the company’s shift to a for-profit model, and he accuses it of abandoning the start-up’s original mission to develop AI for the benefit of humanity.

OpenAI responds to Musk’s lawsuit by revealing early emails where it turns out that [Musk acknowledged the need for profit to fund AI resources](#).

And OpenAI accuses Elon Musk of suing the company for his own commercial interests, claiming that his allegations are unfounded and [based on claimed agreements that don’t exist](#).

Elon opens another front, claiming that [OpenAI has been aggressively recruiting Tesla engineers](#) with massive compensation offers, leading to Tesla having to increase compensation for its AI engineering team.

And Musk escalates his fight with OpenAI by [demanding documents from ex-board member Helen Toner](#) about her departure and the firing of Sam as CEO.

Woo, it's getting exciting again.

Episode 3: Safety first?

Ilya Sutskever has been noticeably invisible since his flip-flop around Sam's ousting and return the previous November. But after six months of radio silence, in May 2024 he pops his head above the parapet and [announces that he's leaving OpenAI](#). Also leaving is Jan Leike, the co-lead of the company's superalignment team, who heads off to join Anthropic, a company which has always proclaimed more loudly its concerns about AI safety. There, Leike will [lead a new superalignment team](#) focusing on AI safety and security.

The departure of both Ilya and Jan from OpenAI is said to be due to disagreements over priorities and allocation of resources. OpenAI had promised that [20% of its computing power would be allocated to AI safety](#), but that promise wasn't fulfilled. We later learn that [another OpenAI employee quit over safety concerns](#) hours before the two execs resigned. The departures leave OpenAI's superalignment team 'dead in the water'. *Wired* confirms that the entire team focused on long-term risk [has either resigned or been absorbed into other research groups](#).

Not surprisingly, all this raises concerns about the company's true priorities around safety. Greg Brockman and Sam Altman [respond on X](#) to the concerns raised, and a subsequent company blog post [elaborates its position](#) via a 'safety update'. But without enforceable regulations, [these efforts remain largely symbolic](#).

Episode 4: Don't mess with Black Widow

It's Monday 13 May 2024. During [a livestream demonstration](#), OpenAI shows off its 'Sky' voice assistant, built on its new GPT-4o LLM: we now have a chatbot you can converse with just like a human. Following the demo, Sam Altman issues a one-word tweet: 'her'. *The Verge* is in no doubt as to what he was referring to, headlining [its story covering the demo](#) 'ChatGPT will be able to talk to you like Scarlett Johansson in Her'.

The word *like* is carrying quite a bit of weight there. You could read it as referring to a similarity in capabilities. Or you could see it as referring to the tonal qualities of the voice. Whatever was intended, enough people take the second reading to provoke heated debate. Scarlett Johansson herself [accuses OpenAI of using her voice likeness](#) for its voice assistant without her permission, despite her turning down a previous offer from the company.

OpenAI denies any intentional imitation, but the company [could still face legal repercussions](#).

OpenAI decides to [pull the voice](#).

Episode 5: NDA

And that's not the company's only mid-May misstep. It's revealed that OpenAI's strict non-disclosure agreement prevents past employees from speaking out about their previous employer. And any employee who refuses to sign the NDA on departure is [threatened with revocation of vested equity](#). This information comes to us courtesy of Daniel Kokotajlo, who, it turns out, resigned from his position at OpenAI as a governance researcher in April after [losing confidence that the company would behave responsibly](#) in its attempt to build AGI.

When the NDA's terms come to light, Altman expresses embarrassment and says he wasn't aware of the threats to take away vested equity. But in documents shared with *Vox*, it becomes apparent that [he had signed off on the terms](#).

OpenAI changes its rules, [promising not to take away employees' vested equity](#).

Episode 6: Helen's story

You might remember that Helen Toner and Tasha McCauley were board members when Sam was ousted back in November. After six months of silence, they write [a piece for *The Economist*](#) advocating for government regulation of AI firms, arguing that self-governance fails due to profit-driven incentives.

Helen goes further in [an interview on the TED AI Show podcast](#), fleshing out the earlier statement that Sam was 'not [being] consistently candid' with the board. She [gives three reasons](#) behind the board's decision: Altman's 'failure to tell the board that he owned the OpenAI Startup Fund', providing inaccurate information about the company's safety processes and attempting to remove Toner from the board after she published 'a research paper that angered him'. [That paper](#) was about the potential risks of AI.

In what is presumably a right-to-reply in response to Helen and Tasha's piece, *The Economist* promptly gives space to Bret Taylor and Larry Summers, current OpenAI board members. They [refute the claims by Helen and Tasha](#) about events related to AI regulation and the CEO's replacement.

All of this causes the press to do a bit of digging into Sam's past. *The Washington Post* claims he was [fired from Y Combinator](#) back in 2019, although [this claim is refuted](#) by Paul Graham, Y Combinator's co-founder. Bloomberg reports that Sam had been '[bending the world to his will](#)' long before OpenAI. There's [polarised social media discussion](#) around Sam's alleged lying and manipulation.

Episode 7: Fame, 15 minutes

Back in February 2024, it was revealed that OpenAI was [looking for an 'insider risk investigator'](#) to protect against internal security threats and safeguard the company's assets. Then, in April, the company [dismissed two of its researchers](#), Leopold Aschenbrenner and Pavel Izmailov, for allegedly leaking information. Both were members of—no, surely not—the OpenAI safety team.

Now, in June, Leopold publishes '[Situational Awareness](#)', a provocative manifesto about what's coming down the AI pipe that is styled as a series of essays. Or perhaps it's really an advertisement for his [new investment firm](#). Anyway, he is made suddenly ubiquitous in the media with his claim that '[AGI by 2027 is strikingly plausible](#)'.

At 165 pages, the manifesto looks like the work of a blogger who adheres to the [Zvi Mowshowitz](#) school of TLDR postings, but you don't have to read it for current purposes. If you're keen, you could instead listen to the entire four-hour conversation with Leopold on Dwarkesh Patel's podcast, but the bit that's relevant to our story [starts at 2h30m in](#) on that interview and runs for about 15 minutes.

Long story short: Leopold claims he was fired from OpenAI for raising security concerns and sharing an internal memo with the board, a charge he denies.

And almost simultaneously with Leopold's accusations, a group of current and former employees of OpenAI and Google DeepMind [issue an open letter](#), endorsed by worthies Yoshua Bengio, Geoffrey Hinton, and Stuart Russell amongst others, calling for the 'right to warn' the public about the potential risks of advanced AI technologies.

Meanwhile, Elon returns for a brief cameo appearance, [withdrawing](#) his breach of contract lawsuit against OpenAI a day before a judge was to consider OpenAI's request for dismissal. But

he's not done fighting: in response to Apple's newly announced ChatGPT integration, he [threatens to ban iPhones](#) from Tesla, SpaceX, and xAI, citing security concerns.

Episode 8: Ilya's back!

Out of nowhere, Ilya Sutskever announces in mid-June that [he's starting a new company](#) called Safe Superintelligence, with the aim of developing exactly what it says on the tin.

The company has [a seriously minimalist website](#) consisting of a single page of what looks like hand-written HTML. The page states the company's focus—addressing what it sees as 'the most important technical problem of our time'—and invites potential employees to get in touch.

Just three months later, by the beginning of September Ilya and his co-founders have [raised US\\$1B](#), an amazing feat for a start-up that has no customers and no product.

At a time when smaller players are starting to be consumed in one way or another by the bigger fish—witness Amazon's [acquisition of Adept](#), Microsoft's [effective acquire of Inflection](#), and Google's [hiring of the co-founders of Character.ai](#)—it's hard to see how new entrants can gain any significant ground. Ilya's clearly very investable, but is even US\$1B enough to make a significant dent when Musk's recently announced xAI Memphis training cluster of 100,000 Nvidia H100 GPUs is said to represent a US\$3–4B investment?

Meanwhile, Greg Brockman announces that he's [taking a sabbatical](#) until the end of the year, and Fortune reports that about half of OpenAI's AGI/ASI safety researchers [have recently left the company](#).

Episode 9: Do the math

It's around this time that we start to see increased concern that we might be witnessing a generative AI hype bubble.

The possibility has been visited often in the technical press, but you have to listen when it's the VCs themselves urging caution. In mid-June, prominent venture capital firm Sequoia Capital does the numbers and concludes that the sector's massive hardware spend means the AI industry [must generate US\\$600B annually to break even](#).

A few days later, Jim Covello of Goldman Sachs warns that [a market correction is inevitable](#) as the economic benefits of AI may be overestimated.

And echoing Sequoia, Gartner predicts that significant increases in datacentre and server spending driven by generative AI demand may be [outpacing the software industry's ability to monetise AI advancements](#).

This commentary spurs deeper analysis of OpenAI's financials. OpenAI's US\$3.4B revenue is driven by ChatGPT Plus and other services, but [high operational costs lead to overall financial losses](#) despite optimistic future valuations. The company's training and inference costs [could reach US\\$7B for 2024](#), with the company losing US\$5B this year.

And Elon files [a new federal lawsuit against OpenAI](#), alleging it abandoned its mission for profit. What profit, some might ask?

Season Finale: Money shot

OpenAI is not blind to the aforementioned financial issues. It emerges that the company is in talks [to raise a further US\\$6.5B](#), but it recognises that its unorthodox corporate structure isn't optimally appealing to investors. Sam announces in a company-wide meeting that [OpenAI will shift from a complex non-profit structure to a more traditional for-profit model](#) next year; the company's aim of a US\$150B valuation [depends on restructuring to remove a profit cap](#) for investors.

OpenAI is also rumoured to be [granting Altman a 7% equity stake](#), but Sam [denies this](#).

Amidst the funding discussions, OpenAI's CTO Mira Murati and two other executives, VP of Research Barret Zoph and Chief Research Officer Bob McGrew, [announce their departures](#) from the company. Now only two members of OpenAI's original 11-member founding team remain: Wojciech Zaremba and, of course, CEO Sam Altman. Karen Hao in *The Atlantic* opines that '[Altman's consolidation of power is nearing completion](#)'.

On October 2, OpenAI announces that it has [raised US\\$6.6B in new funding](#) at a US\$157B post-money valuation. Investors were required to [hand over a minimum of US\\$250m](#), and we learn that they were also asked to [refrain from funding five companies](#) OpenAI perceives as close competitors (reportedly Anthropic, Glean, Perplexity, Safe Superintelligence, and xAI).

This is [the largest venture capital raise of all time](#), making the announcement a fitting climax to Season 1. Cue dramatic music and fade.

And that brings us up-to-date. Despite so many high-level crew members jumping ship, Sam and OpenAI have weathered the storm, but not without some tarnishing of reputations. Assuming [legal complications](#) can be overcome, next year will see a restructured company, one probably [not so beholden to the worthy aims](#) of the original structure. What will happen next? There are so many questions:

- Will Sam Altman [become a 10x billionaire](#)?
- Will OpenAI [need another funding round in 2025](#), or might it IPO?
- Will Safe SuperIntelligence be acquired, and by whom?
- Will Greg Brockman [ever return](#) from his sabbatical?
- Will ScarJo play herself in the TV screenplay?
- Will Leopold's prediction of AGI by 2027 hold up?
- Will Elon do another Elon thing?
- Will all this turn out to be [the true prequel](#) to *The Terminator*?

Stay tuned for Season 2, or just watch it unfold in real time.

If you'd like to keep up to date with everything that's happening in the NLP industry, consider subscribing to the comprehensive and free This Week in NLP newsletter at <https://thisweekinnlp.substack.com/>.