# OpenAI's Sam Altman is becoming one of the most powerful people on Earth. We should be very afraid

## Sam Altman's ChatGPT promises to transform the global economy.

11 min. read  ·        View original

---

On 16 May 2023, Sam Altman, OpenAI's charming, softly spoken, eternally optimistic billionaire CEO, and I stood in front of the US Senate judiciary subcommittee meeting on AI oversight. We were in Washington DC, and it was at the height of AI mania. Altman, then 38, was the poster boy for it all.

Raised in St Louis, Missouri, Altman was the Stanford dropout who had become the president of the massively successful Y Combinator startup incubator before he was 30. A few months before the

hearing, his company's product [ChatGPT] had taken the world by storm. All through the summer of 2023, Altman was treated like a Beatle, stopping by DC as part of a world tour, meeting prime ministers and presidents around the globe. US Senator Kyrsten Sinema gushed: "I've never met anyone as smart as Sam… He's an introvert and shy and humble… But… very good at forming relationships with people on the Hill and… can help folks in government understand AI." Glowing portraits at the time painted the youthful Altman as sincere, talented, rich and interested in nothing more than fostering humanity. His frequent suggestions that AI could transform the global economy had world leaders salivating.

Senator Richard Blumenthal had called the two of us (and IBM's Christina Montgomery) to Washington to discuss what should be done about AI, a "dual-use" technology that held tremendous promise, but also had the potential to cause tremendous harm – from tsunamis of misinformation to enabling the proliferation of new bioweapons. The agenda was AI policy and regulation. We

swore to tell the whole truth, and nothing but the truth.

Altman was representing one of the leading AI companies; I was there as a scientist and author, well known for my scepticism about many things AI-related. I found Altman surprisingly engaging. There were moments when he ducked questions (most notably Blumenthal's "What are you most worried about?", which I pushed Altman to answer with more candour), but on the whole he seemed genuine, and I recall saying as much to the senators at the time. We both came out strongly for AI regulation. Little by little, though, I realised that I, the Senate, and ultimately the American people, had probably been played.

In truth, I had always had some misgivings about OpenAI. The company's press campaigns, for example, were often over the top and even misleading, such as their fancy demo of a robot "solving" a Rubik's Cube that turned out to have special sensors inside. It received tons of press, but it ultimately went nowhere.

For years, the name OpenAI – which implied a kind of openness about the science behind what the company was doing – had felt like a lie, since in reality it has become less and less transparent over time. The company's frequent hints that AGI (artificial general intelligence, AI that can at least match the cognitive abilities of any human) was just around the corner always felt to me like unwarranted hype. But in person, Altman dazzled; I wondered whether I had been too hard on him previously. In hindsight, I had been too soft.

From left: Christina Montgomery, chief privacy and trust officer at IBM, Gary Marcus and Sam Altman appearing before the Senate judiciary subcommittee meeting on AI oversight. Photograph: Bloomberg/Getty Images

I started to reconsider after someone sent me a tip, about something small but telling. At the Senate, Altman painted himself as far more altruistic than he really was. Senator John Kennedy had asked: "OK. You make a lot of money. Do you?" Altman responded: "I make no... I get paid enough for health insurance. I

have no equity in OpenAI," elaborating that: "I'm doing this 'cause I love it." The senators ate it up.

Altman wasn't telling the *full* truth. He didn't own any stock in OpenAI, but he did own stock in Y Combinator, and Y Combinator owned stock in OpenAI. Which meant that Sam had an *indirect* stake in OpenAI, a [fact acknowledged on OpenAI's website](#). If that indirect stake were worth just 0.1% of the company's value, which seems plausible, it would be worth nearly $100m.

▶ Timeline

### From Loopt to OpenAI: Sam Altman's career in brief

Show

That omission was a warning sign. And when the topic returned, he could have corrected it. But he didn't. People loved his selfless myth. (He doubled down, in a piece with *Fortune*, claiming that he didn't need equity with OpenAI because he had ["enough money"](#).) Not long after that, I discovered OpenAI had made a [deal with a chip company that Altman](#)

owned a piece of. The selfless bit started to ring hollow.

The discussion about money wasn't, in hindsight, the only thing from our time in the Senate that didn't feel entirely candid. Far more important was OpenAI's stance on regulation around AI. Publicly, Altman told the Senate he supported it. The reality is far more complicated.

On the one hand, maybe a tiny part of Altman genuinely does want AI regulation. He is fond of paraphrasing Oppenheimer (and is well aware that he shares a birthday with the leader of the Manhattan Project), and recognises that, like nuclear weaponry, AI poses serious risks to humanity. In his own words, spoken at the Senate (albeit after a bit of prompting from me): "Look, we have tried to be very clear about the magnitude of the risks here… My worst fears are that we cause significant – we, the field, the technology, the industry – cause significant harm to the world."

Presumably Altman doesn't want to live in regret and infamy. But behind closed doors, his lobbyists keep pushing for

weaker regulation, or none at all. A month after the Senate hearing, it came out that OpenAI was working [to water down the EU's AI act](). By the time he was fired by OpenAI in November 2023 for being "[not consistently candid]()" with its board, I wasn't all that surprised.

At the time, few people supported the board's decision to fire Altman. A huge number of supporters came to his aid; many treated him like a saint. The well-known journalist Kara Swisher (known to be quite friendly with Altman) blocked me on Twitter for merely *suggesting* that the board might have a point. Altman played the media well. Five days later he was reinstated, with the help of OpenAI's major investor, Microsoft, and a petition supporting Altman from employees.

But a lot has changed since. In recent months, concerns about Altman's candour have gone from heretical to fashionable. Journalist Edward Zitron [wrote]() that Altman was "a false prophet – a seedy grifter that uses his remarkable ability to impress and manipulate Silicon Valley's elite". Ellen Huet of Bloomberg News, on the podcast *Foundering*,

reached the conclusion that "when [Altman] says something, you cannot be sure that he actually means it". Paris Marx has warned of "Sam Altman's self-serving vision". AI pioneer Geoffrey Hinton recently questioned Altman's motives. I myself wrote an essay called the Sam Altman Playbook, dissecting how he had managed to fool so many people for so long, with a mixture of hype and apparent humility.

Many things have led to this collapse in faith. For some, the trigger moment was Altman's interactions earlier this year with Scarlett Johansson, who explicitly asked him *not* to make a chatbot with her voice. Altman proceeded to use a different voice actor, but one who was obviously similar to her in voice, and tweeted "Her" (a reference to a movie in which Johansson supplied the voice for an AI). Johansson was livid. And the ScarJo fiasco was emblematic of a larger issue: big companies such as OpenAI insist their models won't work unless they are trained on all the world's intellectual property, but the companies have given little or no compensation to many of the artists, writers and others

who have created it. Actor [Justine Bateman described]() it as "the largest theft in the [history of the] United States, period".

Meanwhile, OpenAI has long paid lip service to the value of developing measures for AI safety, but several key safety-related staff recently departed, claiming that promises had not been kept. Former OpenAI safety researcher Jan Leike said the company [prioritised shiny things over safety](), as did another recently departed employee, [William Saunders](). Co-founder Ilya Sutskever departed and called his new venture Safe Superintelligence, while former OpenAI employee Daniel Kokotajlo, too, has warned that promises around safety [were being disregarded](). As bad as social media has been for society, errant AI, which OpenAI could accidentally develop, could (as Altman himself notes) be far worse.

'He dazzled.' Sam Altman at the Station F in Paris on 26 May 2023. Photograph: Joel Saget/AFP/Getty Images

The disregard OpenAI has shown for safety is compounded by the fact that the

company appears to be on a campaign to keep its employees quiet. In May, journalist Kelsey Piper uncovered documents that allowed the company to [claw back vested stock from former employees](#) who would not agree to not speak ill of the company, a practice many industry insiders found shocking. Soon after, many former OpenAI employees subsequently signed a letter at [righttowarn.ai](#) demanding whistleblower protections, and as a result the [company climbed down](#), stating it would not enforce these contracts.

Even the company's board felt misled. In May, former OpenAI board member Helen Toner [told the *Ted AI Show* podcast](#): "For years, Sam made it really difficult for the board... by, you know, withholding information, misrepresenting things that were happening at the company, in some cases outright lying to the board."

By late May, bad press for OpenAI and its CEO had accumulated so steadily that the venture capitalist Matt Turck posted [a cartoon on X](#): "days since last easily avoidable OpenAI controversy: 0."

Yet Altman is still there, and still incredibly powerful. He still runs OpenAI, and to a large extent he is still the public face of AI. He has rebuilt the board of OpenAI largely to his liking. Even as recently as April 2024, homeland security secretary Alejandro Mayorkas travelled to visit Altman, to recruit him for [homeland security's AI safety and security board](#).

A lot is at stake. The way that AI develops now will have lasting consequences. Altman's choices could easily affect all of humanity – not just individual users – in lasting ways. Already, as OpenAI has acknowledged, its tools have been used by [Russia and China for creating disinformation](#), presumably with the intent to influence elections. More advanced forms of AI, if they are developed, could pose even more serious risks. Whatever social media has done, in terms of polarising society and subtly influencing people's beliefs, massive AI companies could make worse.

Furthermore, generative AI, made popular by OpenAI, is having a massive environmental impact, measured in

terms of electricity usage, emissions and water usage. As Bloomberg recently put it: "AI is already wreaking havoc on global power systems." That impact could grow, perhaps considerably, as models themselves get larger (the goal of all the bigger players). To a large extent, governments are going on Altman's say-so that AI will pay off in the end (it certainly has not so far), justifying the environmental costs.

Meanwhile, OpenAI has taken on a leadership position, and Altman is on the homeland security safety board. His advice should be taken with scepticism. Altman was at least briefly trying to attract investors to a $7trn investment in infrastructure around generative AI, which could turn out to be a tremendous waste of resources that could perhaps be better spent elsewhere, if (as I and many others suspect) generative AI is not the correct path to AGI [artificial general intelligence].

Finally, overestimating current AI could lead to war. The US-China "chip war" over export controls, for example – in which the US is limiting the export of

critical GPU chips designed by Nvidia, manufactured in Taiwan – is impacting China's ability to proceed in AI and escalating tensions between the two nations. The battle over chips is largely predicated on the notion that AI will continue to improve exponentially, even though data suggests current approaches may recently have reached a point of diminishing returns.

Altman may well have started out with good intentions. Maybe he really did want to save the world from threats from AI, and guide AI for good. Perhaps greed took over, as it so often does.

Unfortunately, many other AI companies seem to be on the path of hype and corner-cutting that Altman charted. Anthropic – formed from a set of OpenAI refugees who were worried that AI safety wasn't taken seriously enough – seems increasingly to be competing directly with the mothership, with all that entails. The billion-dollar startup Perplexity seems to be another object lesson in greed, training on data it isn't supposed to be using. Microsoft, meanwhile, went from advocating "responsible AI" to

rushing out products with serious problems, pressuring Google to do the same. Money and power are corrupting AI, much as they corrupted social media.

We simply can't trust giant, privately held AI startups to govern themselves in ethical and transparent ways. And if we can't trust them to govern themselves, we certainly shouldn't let them govern the world.

---

I honestly don't think we will get to an AI that we can trust if we stay on the current path. Aside from the corrupting influence of power and money, there is a deep technical issue, too: large language models (the core technique of generative AI) invented by Google and made famous by Altman's company, are unlikely *ever* to be safe. They are recalcitrant, and opaque by nature – so-called "black boxes" that we can never fully rein in. The statistical techniques that drive them can do some amazing things, like speed up computer programming and create plausible-sounding interactive characters in the style of deceased loved ones or historical figures. But such black

boxes have *never* been reliable, and as such they are a poor basis for AI that we could trust with our lives and our infrastructure.

That said, I don't think we should abandon AI. Making better AI – for medicine, and material science, and climate science, and so on – really could transform the world. Generative AI is unlikely to do the trick, but some future, yet-to-be developed form of AI might.

The irony is that the biggest threat to AI today may be the AI companies themselves; their bad behaviour and hyped promises are turning a lot of people off. Many are ready for government to take a stronger hand. According to a June poll by Artificial Intelligence Policy Institute, 80% of American voters prefer "regulation of AI that mandates safety measures and government oversight of AI labs instead of allowing AI companies to self-regulate".

To get to an AI we can trust, I have long lobbied for a cross-national effort, similar to Cern's high-energy physics consortium. The time for that is now.

Such an effort, focused on AI safety and reliability rather than profit, and on developing a new set of AI techniques that belong to humanity – rather than to just a handful of greedy companies – could be transformative.

More than that, citizens need to speak up, and demand an AI that is good for the many and not just the few. One thing I can guarantee is that we won't get to AI's promised land if we leave everything in the hands of Silicon Valley. Tech bosses have shaded the truth for decades. Why should we expect Sam Altman, last seen cruising around Napa Valley in a [$4m Koenigsegg supercar](#), to be any different?

*[Gary Marcus](#) is a scientist, entrepreneur and bestselling author. He was founder and CEO of machine learning company Geometric Intelligence, which was acquired by Uber, and is the author of six books, including the forthcoming Taming Silicon Valley (MIT Press)*