'Huge egos are in play': behind the firing and rehiring of OpenAI's Sam Altman

OpenAI's messy firing and rehiring of its powerful chief executive this week shocked the tech world. But the power struggle has implications beyond the company's boardroom, AI experts said. It throws into relief the greenness of the AI industry and the strong desire in Silicon Valley to be first, and raises urgent questions about the safety of the technology. "The AI that we're looking at now is immature. There are no standards, no professional body, no certifications. Everybody figures out how to do it, figures out their own internal norms," said Rayid Ghani, a professor of machine learning and public policy at Carnegie Mellon University. "The AI that gets built relies on a handful of people who built it, and the impact of these handfuls of people is disproportionate." The tussle between Sam Altman and OpenAI's board of directors began on Friday with the unexpected announcement that the board had ousted Altman as CEO for being "not consistently candid in his communications with the board". The blogpost appeared with little warning, even to OpenAI's minority owner Microsoft, which has invested about $13bn in the startup. The board appointed an interim CEO, Mira Murati, then the chief technology officer of OpenAI, but by Sunday had tapped another, the former Twitch CEO Emmett Shear. Altman returned to the startup's headquarters for negotiations the same day; that evening, Microsoft announced it had hired him to lead a new artificial intelligence unit. On Monday, more than 95% of OpenAI's roughly 750 employees signed an open letter asserting they would quit unless Altman were reinstated; signatories included Murati and the man many believed was the architect of Altman's ouster, OpenAI's co-founder and chief scientist, Ilya Sutskever. By Wednesday, Altman was CEO once again. OpenAI's board had been reconstituted without Altman and the company president, Greg Brockman, who had quit in solidarity but also was rehired, and without two members who had voted to fire them both. In the absence of substantive regulation of the companies making AI, the foibles and idiosyncrasies of its creators take on outsized importance. Asked what OpenAI's saga could mean for any upcoming AI regulation, the United Kingdom's Department for Science, Innovation and Technology (DSIT) said in a statement: "Because this is a commercial decision, it's not something for DSIT to comment on." In the US, the White House also did not provide comment. Senators Richard Blumenthal of Connecticut and Josh Hawley of Missouri, chairs of the US Senate subcommittee that oversaw Altman's testimony earlier this year, did not respond to requests for comment; Blumenthal and Hawley have proposed a bipartisan AI bill "to establish guardrails". In a more mature sector, regulations would insulate consumers and consumer-facing products from the fights among the people at the top, Ghani said. The individual makers of AI would not be so consequential, and their spats would affect the public less. "It's too risky to rely on one person to be the spokesperson for AI, especially if that person is responsible for building. It shouldn't be self-

regulated. When has that ever worked? We don't have self-regulation in anything that is important, why would we do it here?" he asked. The political battle over AI The struggle at OpenAI also highlighted a lack of transparency into decision-making at the company. The development of cutting-edge AI rests in the hands of a small, secretive cadre that operates behind closed doors. "We have no idea how a staff change at OpenAI would change the nature of ChatGPT or Dall-E," said Ghani. At the moment, there's no public body running tests of programs like ChatGPT, and companies aren't transparent about updates. Compare that to an iPhone or Android's software updates, which list the changes and fixes coming to the software of the device you hold in your hand. "Right now, we don't have a public way of doing quality control. Each organization will do that for their own use cases," he said. "But we need a way to continuously run tests on things like ChatGPT and monitor the results so as to profile the results for people and make it lower risk. If we had such a tool, the company would be less critical. Our only hope is that the people building it know what they're doing." Paul Barrett, the deputy director of the center for business and human rights at New York University's business school, agreed, calling for regulation that would require AI makers to demonstrate the safety and efficacy of their products the way pharmaceutical companies do. "The fight for control of OpenAI provides a valuable reminder of the volatility within this relatively immature branch of the digital industry and the danger that crucial decisions about how to safeguard artificial intelligence systems may be influenced by corporate power struggles. Huge amounts of money – and huge egos – are in play. Judgments about when unpredictable AI systems are safe to be released to the public should not be governed by these factors," he said. Acceleration v deceleration The split between Altman and the board at least partly seemed to fall along ideological lines, with Altman and Brockman in a camp known as "accelerationists" – people who believe AI should be deployed as quickly as possible – and "decelerationists" – people who believe it should be developed more slowly and with stronger guardrails. With Altman's return, the former group takes the spoils. "The people who seem to have won out in this case are the accelerationists," said Sarah Kreps, a Cornell professor of government and the director of the Tech Policy Institute in the university's school of public policy. Kreps said we may see a reborn OpenAI that fully subscribes to the Meta chief executive Mark Zuckerberg's "move fast and break things" mantra. Employees voted with their feet in the debate between moving more quickly or more carefully, she noted. "What we'll see is full steam ahead on AI research going forward. Then the question becomes, is it going to be totally unsafe, or will it have trials and errors? OpenAI may follow the Facebook model of moving quickly and realizing that the product is not always compatible with societal good," she said. What's accelerating the AI arms race among OpenAI, Google, Microsoft and other tech giants, Kreps said, is vast amounts of capital and the burning desire to be first. If one company doesn't make a certain discovery, another will – and fast. That leads to less caution. "The former leadership of OpenAI has said all the right things about being cognizant of the risk, but as more money has poured into AI, the more incentive there is to move quickly and be less mindful of those risks," she said. Full speed ahead The Silicon Valley wrestling match has called into question the future of the prominent startup's business and its flagship product, ChatGPT. Altman had been on a world tour as an emissary for AI in the preceding weeks. He had spoken to Joe Biden, China's Xi Jinping and other diplomats just days before at the Apec conference in San Francisco. Two weeks before, he had debuted the capability for developers to build their own versions of ChatGPT at a splashy demo day that featured the Microsoft chief executive, Satya Nadella, who has formed a strong partnership with Altman and cast his company's lot with the younger man. How could nations, strategic partners and customers trust OpenAI, though, if its own rulers would throw it into such disarray? "The fact that the board hasn't given a clear statement on its reasons for firing Altman, even to the CEO that the board itself hired, looks very bad," said Derek Leben, a professor of ethics at Carnegie Mellon's business school. Altman, Leben said, came out the winner in the public relations war, the protagonist in the story. Kreps agreed. In favor of decelerationists, Leben said, is that this saga proved they are serious about their concerns, if ham-fisted in their expression. AI skeptics have criticized Altman and others for prophesying doom-by-AI in the future, arguing such concerns overlook real harms AI does in the present day and only aggrandize AI's makers. "The fact that people are willing to burn down the company suggests that they're not just using safety as a smokescreen for an ulterior motive. They're being sincere when they say they're willing to shut down the company to prevent bad outcomes," he said. One thing OpenAI's succession war will not do is slow down the development of AI, the experts agreed. "I'm less concerned for the safety of AI. I think AI will be safe, for better or for worse. I'm worried for the people who have to use it," said Ghani. Johana Bhuiyan contributed reporting