

Congress of the United States

Washington, DC 20515

August 8, 2024

Mr. Sam Altman
Chief Executive Officer
OpenAI
3180 18th Street, Suite 100
San Francisco, California 94110

Dear Mr. Altman:

We write with concern regarding reports that OpenAI “is putting a priority on profits and growth” while “prevent[ing] workers from voicing their concerns about [its] technology.”¹ You have stated that addressing the dangers posed by artificial intelligence (AI) “should be a global priority.”² Yet OpenAI’s board members and employees have repeatedly warned that OpenAI has sacrificed safety in the pursuit of profit, including by suppressing employees’ ability to notify regulators and the public about emerging threats.³ Given the discrepancy between your public comments and reports of OpenAI’s actions, we request information about OpenAI’s whistleblower and conflict of interest protections in order to understand whether federal intervention may be necessary.

Recent advances in AI have raised significant safety concerns. You yourself have acknowledged that AI models have the potential to be “significantly destabilizing for public safety and national security” and that you “can’t anticipate every . . . potential abuse, or failure of the technology.”⁴ You participated in Leader Schumer’s AI forums and highlighted safety concerns in your statements at a Senate Judiciary Committee hearing, describing AI regulation as essential and emphasizing that it is “vital that AI companies . . . adhere to an appropriate set of safety requirements, including internal and external testing prior to release and publication of evaluation results.”⁵

¹ New York Times, “OpenAI Insiders Warn of a ‘Reckless’ Race for Dominance,” Kevin Roose, June 4, 2024, <https://www.nytimes.com/2024/06/04/technology/openai-culture-whistleblowers.html>.

² Center for AI Safety, “Statement on AI Risk,” open letter, <https://www.safe.ai/work/statement-on-ai-risk>.

³ New York Times, “OpenAI Insiders Warn of a ‘Reckless’ Race for Dominance,” Kevin Roose, June 4, 2024, <https://www.nytimes.com/2024/06/04/technology/openai-culture-whistleblowers.html>.

⁴ U.S. Senate Committee on the Judiciary, “Written Testimony of Sam Altman,” May 16, 2023, <https://www.judiciary.senate.gov/imo/media/doc/2023-05-16%20-%20Bio%20&%20Testimony%20-%20Altman.pdf>.

⁵ *Id.*; Wired, “Inside the Senate’s Private AI Meeting with Tech’s Billionaire Elites,” Matt Laslo and Khari Johnson, September 14, 2023, <https://www.wired.com/story/senate-ai-forum-musk-zuckerberg/>.

Yet multiple reports indicate that OpenAI’s leadership has prioritized profits over safety.⁶ In 2022, the then-unreleased AI tool GPT-4 was tested as part of a new version of Microsoft’s Bing search engine in India, even though the tool had not yet been approved by OpenAI’s safety board.⁷ Microsoft initially denied that this secret testing occurred, before reversing course following a former OpenAI’s employee’s whistleblowing report and press scrutiny.⁸ A year later, OpenAI’s board voted to oust you as CEO, effective immediately, based in part on “concerns over commercializing advances before understanding the consequences.”⁹ Board members said you had not been “consistently candid in [your] communications,”¹⁰ adding that they “just couldn’t believe things that [you were] telling [them]” in large part due to a history of “withholding information [and] misrepresenting things that were happening at the company.”¹¹ You later rejoined the company alongside a dramatically different board; three of the four directors who had voted to remove you were replaced.¹²

Even after this turmoil, OpenAI appears to be prioritizing commercial success over transparency and safety. Last month, a group of current and former OpenAI employees published an open letter announcing they had “lost confidence that OpenAI will behave responsibly” in the race to replicate human intelligence.¹³ The open letter came after revelations that OpenAI routinely included non-disparagement clauses in employees’ contracts to restrict them, “for the rest of their lives, from criticizing their former employer.”¹⁴ As a result of the clauses, departing employees would lose all vested equity earned during their time at OpenAI if they criticized the company,

⁶ New York Times, “OpenAI Insiders Warn of a ‘Reckless’ Race for Dominance,” Kevin Roose, June 4, 2024, <https://www.nytimes.com/2024/06/04/technology/openai-culture-whistleblowers.html>; Reuters, “OpenAI researchers warned board of AI breakthrough ahead of CEO ouster, sources say,” Anna Tong, Jeffrey Dastin and Krystal Hu, November 23, 2023, <https://www.reuters.com/technology/sam-altmans-ouster-openai-was-precipitated-by-letter-board-about-ai-breakthrough-2023-11-22/>.

⁷ New York Times, “OpenAI Insiders Warn of a ‘Reckless’ Race for Dominance,” Kevin Roose, June 4, 2024, <https://www.nytimes.com/2024/06/04/technology/openai-culture-whistleblowers.html>; *see also* OpenAI, “Lessons learned on language model safety and misuse,” March 3, 2022, <https://openai.com/index/language-model-safety-and-misuse/#risk-impact-measurement> (explaining OpenAI’s 2022 “model deployment approach” including the standard practice of “[p]re-deployment risk analysis” as part of a commitment “to investing in safety and policy research even when they trade off with commercial utility”); OpenAI, “OpenAI safety update,” May 21, 2024, <https://openai.com/index/openai-safety-update/> (affirming OpenAI’s existing practice, as of May 2024, of the Safety Advisory Group “review[ing] model capability reports and mak[ing] recommendations ahead of deployment”).

⁸ New York Times, “OpenAI Insiders Warn of a ‘Reckless’ Race for Dominance,” Kevin Roose, June 4, 2024, <https://www.nytimes.com/2024/06/04/technology/openai-culture-whistleblowers.html>.

⁹ Reuters, “OpenAI researchers warned board of AI breakthrough ahead of CEO ouster, sources say,” Anna Tong, Jeffrey Dastin, and Krystal Hu, November 23, 2023, <https://www.reuters.com/technology/sam-altmans-ouster-openai-was-precipitated-by-letter-board-about-ai-breakthrough-2023-11-22/>.

¹⁰ Ars Technica, “OpenAI fires CEO Sam Altman, citing less than ‘candid’ communications,” Kyle Orland and Benj Edwards, November 17, 2023, <https://arstechnica.com/ai/2023/11/openai-fires-ceo-sam-altman-citing-less-than-candid-communications/>.

¹¹ CNBC, “Former OpenAI board member explains why CEO Sam Altman got fired before he was rehired,” Hayden Field, May 29, 2024, <https://www.cnbc.com/2024/05/29/former-openai-board-member-explains-why-ceo-sam-altman-was-fired.html>.

¹² Wall Street Journal, “Sam Altman Rejoins OpenAI Board Along With Three New Directors,” Deepa Seetharaman, Tom Dotan, March 8, 2024, <https://www.wsj.com/tech/ai/openai-board-brings-sam-altman-back-adds-three-new-members-bbdb4807>; The New York Times, “Sam Altman Asserts Control of OpenAI as He Rejoins Its Board,” Cade Metz, Tripp Mickle, Mike Isaac, March 8, 2024, <https://www.nytimes.com/2024/03/08/technology/sam-altman-to-return-to-openais-board-of-directors.html>.

¹³ New York Times, “OpenAI Insiders Warn of a ‘Reckless’ Race for Dominance,” Kevin Roose, June 4, 2024, <https://www.nytimes.com/2024/06/04/technology/openai-culture-whistleblowers.html>.

effectively forcing employees to choose between their financial future and their ability to speak out on safety concerns.¹⁵ Following these revelations, OpenAI subsequently announced that the company will cease using or enforcing non-disparagement clauses.¹⁶ However, it is unclear whether current or former employees have sufficient protections to act as whistleblowers regarding the potentially unique risks posed by AI tools.

To date, OpenAI employees appear to have publicly aired their safety concerns only after a product has been released. For example, an OpenAI executive resigned in the wake of the release of GPT-4o in May 2024,¹⁷ claiming “safety culture and processes have taken a backseat to shiny products.”¹⁸ Another former employee described “rushed and not very solid” safety work “in service of meeting [product] shipping date[s].”¹⁹ These concerns culminated in OpenAI whistleblowers filing a complaint with the Securities and Exchange Commission (SEC) in July 2024 alleging that OpenAI’s use of non-disclosure and related agreements reflected “systemic violations” of federal law.²⁰ The employees noted that “even if OpenAI is making reforms in light of public disclosures of their illegal contracts . . . enforcement action is critical” to ensure that AI companies do not continue to “send[] direct or indirect messages to employees that they must honor a ‘code of silence’ that has resulted in so many disasters in the past.”²¹

Robust safeguards for whistleblowers at OpenAI are all the more important given your numerous reported conflicts of interest. You reportedly operate a “sprawling investment empire that is becoming a direct beneficiary of OpenAI’s success.”²² This empire is worth at least \$2.8 billion and includes a growing number of startups that do business with OpenAI.²³ For example, you control 7.6% of Reddit as the third-largest outside shareholder.²⁴ When OpenAI announced a partnership with Reddit to use its content to train its AI models, Reddit’s stock rose by 10.5%, which boosted your personal stake by \$69 million.²⁵ You also personally invested over one

¹⁴ Vox, “ChatGPT can talk, but OpenAI employees sure can’t,” Kelsey Piper, May 18, 2024, <https://www.vox.com/future-perfect/2024/5/17/24158478/openai-departures-sam-altman-employees-chatgpt-release>.

¹⁵ *Id.*

¹⁶ CNBC, “OpenAI sends internal memo releasing former employees from controversial exit agreements,” Hayden Field, May 23, 2024, <https://www.cnbc.com/2024/05/24/openai-sends-internal-memo-releasing-former-employees-from-non-disparagement-agreements-sam-altman.html>; Tweet by Sam Altman, August 1, 2024, <https://x.com/sama/status/1818867964369928387>.

¹⁷ Open AI, “Hello GPT-4o,” May 13, 2024, <https://openai.com/index/hello-gpt-4o/>.

¹⁸ Tweet by Jan Leike, May 17, 2024, <https://x.com/janleike/status/1791498184671605209>.

¹⁹ Washington Post, “OpenAI promised to make its AI safe. Employees say it ‘failed’ its first test,” Pranshu Verma, Nitasha Tiku and Cat Zakrzewski, July 12, 2024, <https://www.washingtonpost.com/technology/2024/07/12/openai-ai-safety-regulation-gpt4/>.

²⁰ Letter from OpenAI Whistleblowers to SEC Chair Gensler Re: OpenAI Violations of Rule 21F-17(a) and Implementation of E.O. 14110, July 1, 2024, p.1, https://www.washingtonpost.com/documents/83df0e55-546c-498a-9efc-06fac591904e.pdf?itid=lk_inline_manual_4.

²¹ *Id.*, p.5.

²² Wall Street Journal, “The Opaque Investment Empire Making OpenAI’s Sam Altman Rich,” Berber Jin, Tom Dotan and Keach Hagey, June 3, 2024, <https://www.wsj.com/tech/ai/openai-sam-altman-investments-004fc785>.

²³ *Id.*

²⁴ *Id.*

²⁵ Reuters, “OpenAI Strikes Deal to Bring Reddit Content to ChatGPT,” May 16, 2024, <https://www.reuters.com/markets/deals/openai-strikes-deal-bring-reddit-content-chatgpt-2024-05-16/>; Wall Street Journal, “The Opaque Investment Empire Making OpenAI’s Sam Altman Rich,” Berber Jin, Tom Dotan and Keach Hagey, June 3, 2024, <https://www.wsj.com/tech/ai/openai-sam-altman-investments-004fc785>.

million dollars in RainAI, a developer of AI chips.²⁶ Later, OpenAI signed a letter of intent to buy \$51 million of chips from this startup.²⁷ Your position on both sides of these deals may strengthen your incentive to release products as quickly as possible, regardless of safety concerns, and undermine your responsibility as CEO to meet OpenAI’s mission of “ensur[ing] that artificial general intelligence benefits all of humanity.”²⁸

Whistleblower protections are vital tools in combatting corporate misconduct. This is why we have previously introduced legislation such as the *COVID-19 Whistleblower Protection Act*, the *FTC Whistleblower Act*, and the *SEC Whistleblower Reform Act* to protect those who shine a light on corporate misconduct.²⁹ The revelations from your former employees³⁰ make clear that robust whistleblower protections are especially necessary in the AI sector. A growing, bipartisan group in Congress have echoed these concerns.³¹

To better understand whether Congressional intervention is necessary to protect whistleblowers warning of AI-related harms and to protect against misaligned incentives created by your financial conflicts of interest, we ask that you provide answers to the following questions no later than August 22, 2024:

1. In a letter addressed to a group of U.S. Senators dated July 31, 2024, OpenAI indicated that in March 2024 it introduced an “Integrity Line” to allow OpenAI personnel to anonymously “raise concerns, including about AI safety and/or violations of applicable law and policy.”³²
 - a. What actions did OpenAI take to inform employees of the introduction of the hotline?
 - b. How many reports has the hotline received since its introduction?
 - c. Of the total number of reports, how many were escalated to OpenAI’s Board of Directors?
 - d. What was the ultimate outcome of these incidents?
 - e. Were any of these employees – either formally or informally – penalized or disciplined?
2. The same letter indicated that OpenAI’s Employee Handbook now “details various means by which employees can raise issues,” “expressly prohibits harassment and

²⁶ CCN, “Altman’s Conflict of Interest? OpenAI Buys Chip from a Startup Its CEO Invested In,” Giuseppe Ciccomascolo, December 6, 2023, <https://www.ccn.com/analysis/openai-altman-rain-chip/>.

²⁷ *Id.*

²⁸ Open AI, “About,” <https://openai.com/about/>.

²⁹ COVID-19 Whistleblower Protection Act, S.268, <https://www.congress.gov/bill/117th-congress/senate-bill/268>; FTC Whistleblower Act of 2021, H.R.6093, <https://www.congress.gov/bill/117th-congress/house-bill/6093>; SEC Whistleblower Reform Act of 2023, S.811, <https://www.congress.gov/bill/118th-congress/senate-bill/811>.

³⁰ New York Times, “OpenAI Insiders Warn of a ‘Reckless’ Race for Dominance,” Kevin Roose, June 4, 2024, <https://www.nytimes.com/2024/06/04/technology/openai-culture-whistleblowers.html>.

³¹ Letter from Senators Schatz, Luján, Welch, Warner, and King to Sam Altman, July 22, 2024, https://www.schatz.senate.gov/imo/media/doc/letter_to_openai.pdf; Letter from Senator Grassley to Sam Altman, August 1, 2024, https://www.grassley.senate.gov/imo/media/doc/grassley_to_openai_-_ndas.pdf.

³² Letter from Jason Kwon, Chief Strategy Officer at Open AI to Senators Schatz, King, Luján, Warner, and Welch, July 31, 2024, p.6, https://www.washingtonpost.com/documents/cd2edf3d-c892-4293-b2d7-41c882d2e1fb.pdf?itid=lk_inline_manual_23.

retaliation against employees who decide to do so,” and “makes clear that employees may make reports or disclosures to any national, federal, state, or local government agency.”³³

- a. Please provide copies of the current Employee Handbook, as well as any previous versions of the Employee Handbook used from January 2022 to the present.
 - b. A letter accompanying the SEC complaint filed by OpenAI whistleblowers allege that OpenAI’s employment agreements violated federal law due to “[n]on-disparagement clauses that failed to exempt disclosures of securities violations to the SEC,” “[r]equiring prior consent from the company to disclose confidential information to federal authorities,” “[c]onfidentiality requirements with respect to agreements, that themselves contain securities violations,” and [r]equiring employees to waive compensation that was intended by Congress to incentivize reporting and provide financial relief to whistleblowers.”³⁴
 - i. In the case of a conflict between the Employee Handbook and employment agreements (including severance, non-disparagement, and non-disclosure documents), which document governs?
 - c. What further steps will OpenAI take to ensure that the ‘code of silence’ referenced in the SEC whistleblowers’ letter³⁵ does not implicitly or explicitly discourage potential whistleblowers from raising complaints in the future?
3. In the letter addressed to a group of U.S. Senators dated July 31, 2024, OpenAI stated that the company voided non-disparagement provisions for all current and former employees and informed employees of the changes in May 2024 and also sent a follow-up communication and updated the relevant documents in July 2024.³⁶
- a. Please provide copies of the May 2024 communication and the July 2024 communication.
 - b. Please provide copies of the updated documents issued in July 2024.
 - c. Has OpenAI undertaken any additional steps to inform former and current employees of these changes?
4. Has OpenAI changed its process for ensuring that new products undergo a pre-deployment safety review, following Microsoft’s admission that an early version of GPT-4 was used in India prior to safety approval and an OpenAI employee’s warnings that GPT-4o was rushed through a compressed safety timeline?³⁷
- a. Please provide a detailed accounting of all instances where OpenAI’s products previously bypassed safety protocols.
 - b. Please outline what circumstances, if any, would permit a product to bypass safety review.

³³ *Id.*

³⁴ Letter from OpenAI whistleblowers to SEC Chair Gensler Re: OpenAI Violations of Rule 21F-17(a) and Implementation of E.O. 14110, July 1, 2024, p. 4, https://www.washingtonpost.com/documents/83df0e55-546c-498a-9efc-06fac591904e.pdf?itid=lk_inline_manual_4.

³⁵ *Id.*, p. 5.

³⁶ Letter from Jason Kwon, Chief Strategy Officer at Open AI to Senators Schatz, King, Luján, Warner, and Welch, July 31, 2024, pp.5-6, https://www.washingtonpost.com/documents/cd2edf3d-c892-4293-b2d7-41c882d2e1fb.pdf?itid=lk_inline_manual_23.

³⁷ New York Times, “OpenAI Insiders Warn of a ‘Reckless’ Race for Dominance,” Kevin Roose, June 4, 2024, <https://www.nytimes.com/2024/06/04/technology/openai-culture-whistleblowers.html>; Washington Post, “OpenAI promised to make its AI safe. Employees say it ‘failed’ its first test,” Pranshu Verma, Nitasha Tiku and Cat Zakrzewski, July 12, 2024, <https://www.washingtonpost.com/technology/2024/07/12/openai-ai-safety-regulation-gpt4/>.

- c. Please provide a flowchart detailing the pre-deployment safety review process and highlighting key decision-makers and explaining in what circumstances each decision-makers' have veto power over any product releases, and when these determinations can be overruled.
 - d. OpenAI has indicated that a "Deployment Safety Board" will "approve decisions to deploy models above a certain capability threshold."³⁸ What is that threshold? What is the decision process for this Board?
 - e. OpenAI has stated that it is "moving towards an agreement" with the U.S. AI Safety Institute in which it will "provide early access to [OpenAI's] next model."³⁹ Will the agreement provide the U.S. AI Institute the authority to delay or halt the release of a model?
5. After your return to OpenAI, the company reportedly established an audit committee and implemented a new conflicts policy.⁴⁰
- a. Please provide a copy of OpenAI's new conflicts policy.
 - b. Please provide a list of any investigations conducted by the audit committee into your conflicts of interest, including a description of each investigation and the resolution of the investigation.
 - c. What specific protections are in place to protect OpenAI from your financial conflicts of interest?
 - i. Has OpenAI's audit committee or another entity required you to divest from any of your outside holdings or recuse yourself from any decisions at OpenAI because of conflicts of interest?

Sincerely,



Elizabeth Warren
United States Senator



Lori Trahan
Member of Congress

³⁸ Letter from Jason Kwon, Chief Strategy Officer at Open AI to Senators Schatz, King, Luján, Warner, and Welch, July 31, 2024, p.1, https://www.washingtonpost.com/documents/cd2edf3d-c892-4293-b2d7-41c882d2e1fb.pdf?itid=lk_inline_manual_23.

³⁹ *Id.*, p.2.

⁴⁰ Wall Street Journal, "The Opaque Investment Empire Making OpenAI's Sam Altman Rich," Berber Jin, Tom Dotan and Keach Hagey, June 3, 2024, <https://www.wsj.com/tech/ai/openai-sam-altman-investments-004fc785>.