

United States Senate

WASHINGTON, DC 20510

July 22, 2024

Sam Altman
CEO
OpenAI
3180 18th Street
San Francisco, CA 94110

Dear Mr. Altman,

We write to you regarding recent reports¹ about OpenAI's safety and employment practices. OpenAI has announced a guiding commitment to the safe, secure, and responsible development of artificial intelligence (AI) in the public interest. These reports raise questions about how OpenAI is addressing emerging safety concerns. We seek additional information from OpenAI about the steps that the company is taking to meet its public commitments on safety, how the company is internally evaluating its progress on those commitments, and on the company's identification and mitigation of cybersecurity threats.

Safe and secure AI is widely viewed as vital to the nation's economic competitiveness and geopolitical standing in the twenty-first century. Moreover, OpenAI is now partnering with the U.S. government and national security and defense agencies to develop cybersecurity tools to protect our nation's critical infrastructure. National and economic security are among the most important responsibilities of the United States Government, and unsecure or otherwise vulnerable AI systems are not acceptable.

Given OpenAI's position as a leading AI company, it is important that the public can trust in the safety and security of its systems. This includes the integrity of the company's governance structure and safety testing, its employment practices, its fidelity to its public promises and mission, and its cybersecurity policies. The voluntary commitments that you and other leading AI companies made with the White House last year were an important step towards building this trust.

We therefore request the following information by August 13, 2024:

1. Does OpenAI plan to honor its previous public commitment to dedicate 20 percent² of its computing resources to research on AI safety?

¹ <https://www.nytimes.com/2024/06/04/technology/openai-culture-whistleblowers.html>, <https://www.vox.com/future-perfect/2024/5/17/24158403/openai-resignations-ai-safety-ilya-sutskever-jan-leike-artificial-intelligence>, <https://www.washingtonpost.com/technology/2024/07/13/openai-safety-risks-whistleblower-sec/>, <https://www.nytimes.com/2024/07/04/technology/openai-hack.html>, <https://www.washingtonpost.com/technology/2024/07/12/openai-ai-safety-regulation-gpt4/>.

² <https://openai.com/index/introducing-superalignment/>, <https://fortune.com/2024/05/21/openai-superalignment-20-compute-commitment-never-fulfilled-sutskever-leike-altman-brockman-murati/>.

- a. If so, describe the steps that OpenAI has, is, or will take to dedicate 20 percent of its computing resources to research on AI safety.
 - b. If not, what is the percentage of computing resources that OpenAI is dedicating to AI safety research?
2. Can you confirm that your company will not enforce permanent non-disparagement³ agreements for current and former employees?
3. Can you further commit to removing any other provisions from employment agreements that could be used to penalize employees who publicly raise concerns about company practices, such as the ability to prevent employees from selling their equity in private “tender offer” events?
 - a. If not, please explain why, and any internal protections in place to ensure that these provisions are not used to financially disincentivize whistleblowers.
4. Does OpenAI have procedures in place for employees to raise concerns about cybersecurity and safety? How are those concerns addressed once they are raised?
 - a. Have OpenAI employees raised concerns about the company’s cybersecurity practices?
5. What security and cybersecurity protocols does OpenAI have in place, or plan to put in place, to prevent malicious actors or foreign adversaries from stealing an AI model, research, or intellectual property from OpenAI?⁴
6. The OpenAI Supplier Code of Conduct⁵ requires your suppliers to implement strict non-retaliation policies and provide whistleblowers channels for reporting concerns without fear of reprisal. Does OpenAI itself follow these practices?
 - a. If yes, describe OpenAI’s non-retaliation policies and whistleblower reporting channels, and to whom those channels report.
7. Does OpenAI allow independent experts to test and assess the safety and security of OpenAI’s systems pre-release?⁶
8. Does the company currently plan to involve independent experts on safe and responsible AI development in its safety and security testing and evaluation processes, procedures, and techniques, and in its governance structure, such as in its safety and security committee?
9. Will OpenAI commit to making its next foundation model available to U.S. Government agencies⁷ for pre-deployment testing, review, analysis, and assessment?

³ <https://www.vox.com/future-perfect/351132/openai-vested-equity-nda-sam-altman-documents-employees>.

⁴ <https://www.nytimes.com/2024/07/04/technology/openai-hack.html>.

⁵ <https://openai.com/policies/supplier-code/>.

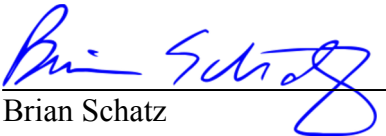
⁶ <https://www.washingtonpost.com/technology/2024/07/12/openai-ai-safety-regulation-gpt4/>.

⁷ <https://www.politico.eu/article/rishi-sunak-ai-testing-tech-ai-safety-institute/>.

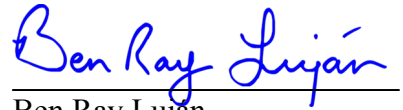
10. What are OpenAI's post-release monitoring practices? What patterns of misuse and safety risks have your teams observed after the deployment of your most recently released large language models? What scale must such risks reach for your monitoring practices to be highly likely to catch them? Please share your learnings from post-deployment measurements⁸ and the steps taken to incorporate them into improving your policies, systems, and model updates.
11. Do you plan to make retrospective impact assessments of your already-deployed models available to the public?
12. Please provide documentation on how OpenAI plans to meet its voluntary safety and security commitments to the Biden-Harris administration.⁹

Thank you very much for your attention to these matters.


Sincerely,



Brian Schatz
United States Senator



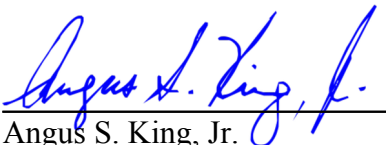
Ben Ray Lujan
United States Senator



Peter Welch
United States Senator



Mark R. Warner
United States Senator



Angus S. King, Jr.
United States Senator

⁸ <https://openai.com/global-affairs/our-approach-to-frontier-risk/>.

⁹ <https://www.whitehouse.gov/briefing-room/statements-releases/2023/07/21/fact-sheet-biden-harris-administration-secures-voluntary-commitments-from-leading-artificial-intelligence-companies-to-manage-the-risks-posed-by-ai/>.