

Congress of the United States

Washington, DC 20515

September 13, 2023

Mr. Satya Nadella
Chief Executive Officer
Microsoft
One Microsoft Way
Redmond, WA 98052

Mr. Andy Jassy
President and Chief Executive Officer
Amazon
410 Terry Ave. North
Seattle, WA 98109

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

Mr. Mustafa Suleyman
Chief Executive Officer
Inflection AI
650 Page Mill Rd.
Palo Alto, CA 94304

Dr. Dario Amodei
Chief Executive Officer
Anthropic
548 Market Street
San Francisco, CA 94104

Mr. Alexandr Wang
Chief Executive Officer
Scale AI
155 5th St Fl 6
San Francisco, CA 94103

Mr. Mark Zuckerberg
Chairman and Chief Executive Officer
Meta
1 Hacker Way
Menlo Park, CA 94025

Mr. Arvind Krishna
Chairman and Chief Executive Officer
IBM
1 Orchard Rd
Armonk, NY 10504

Mr. Sundar Pichai
Chief Executive Officer
Alphabet Inc. and Google LLC
1600 Amphitheatre Parkway
Mountain View, CA 94043

Dear Mr. Nadella, Mr. Altman, Dr. Amodei, Mr. Zuckerberg, Mr. Pichai, Mr. Jassy, Mr. Suleyman, Mr. Wang, and Mr. Krishna,

We write with deep concerns and questions about the working conditions of those who perform your companies' "ghost work" — unseen but critical tasks such as data labeling, without which there would be no artificial intelligence (AI).¹ Despite the essential nature of this work, millions of data workers around the world perform these stressful tasks under constant surveillance, with

¹ Mary L. Gray & Siddarth Suri, *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*, HarperCollins (2019).

low wages and no benefits.² These conditions not only harm the workers, they also risk the quality of the AI systems — potentially undermining accuracy,³ introducing bias,⁴ and jeopardizing data protection.⁵ Tech companies have a responsibility to ensure safe and healthy working conditions, fairly compensated work, and protection from unjust disciplinary proceedings.⁶ Tech companies also must be more transparent about the role data workers play in their AI, so that consumers can make informed choices about the products they use. Unfortunately, many companies have sidestepped these duties, and that must change.⁷

An unseen force behind the meteoric rise of AI products such as ChatGPT, Bard, and Bing is data workers. Contrary to the popular notion that AI is entirely machine-based and autonomous, AI systems in fact depend heavily on human labor — for tasks ranging from labeling training data to rating chatbot responses for accuracy and safety.⁸ Today, millions of workers perform these tasks, paid hourly or for piecework on digital labor platforms like Amazon’s Mechanical Turk (MTurk), and their numbers are growing.⁹

Despite the essential nature of this work for AI, the working conditions are grueling. Wages are low: the median wage of workers on MTurk has been estimated at just \$1.77 per hour.¹⁰ Most workers receive no health insurance or benefits.¹¹ As much as a third of workers’ time is spent on uncompensated work such as searching for tasks, undergoing training, and communicating with requesters,¹² and many workers are periodically unpaid without explanation.¹³ Workers are often under constant surveillance, with keystroke logs, computer screenshots, and even webcam photos taken by the digital labor platform. Supervised by algorithms, some data workers face lockouts

² Adrienne Williams et al., *The Exploited Labor Behind Artificial Intelligence*, NOEMA (Oct. 13, 2022), <https://www.noemamag.com/the-exploited-labor-behind-artificial-intelligence/>.

³ @AlphabetWorkers, Twitter (May 16, 2023 11:06 AM), <https://twitter.com/AlphabetWorkers/status/1658489222695915522>.

⁴ Nata Barbosa & Monchu Chen, *Rehumanizing Crowdsourcing: A Label Framework Addressing Bias and Ethics in Machine Learning*, Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 1 (2019).

⁵ Danica Sergison, *Privacy Risks for Customers and Workers in the Gig Economy*, PIA Blog (Sept. 2, 2018), <https://www.privateinternetaccess.com/blog/privacy-risks-for-customers-and-workers-in-the-gig-economy/>.

⁶ Julian Friedland et al., *A Ghost Workers’ Bill of Rights: How to Establish a Fair and Safe Gig Work Platform*, Cal. Rev. Mgmt. (Jan. 7 2020), <https://cmr.berkeley.edu/2020/01/ghost-workers/>; *Fairwork AI Principles* 23.07, Fairwork (July 2023), <https://fair.work/en/fw/principles/ai-principles/>.

⁷ Fairwork, *Fairwork Cloudwork Ratings 2023: Work in the Planetary Labour Market* (July 2023), <https://fair.work/wp-content/uploads/sites/17/2023/07/Fairwork-Cloudwork-Ratings-2023-Red.pdf>.

⁸ *Responsible Sourcing of Data Enrichment Services*, P’ship on AI (June 17, 2021), <http://partnershiponai.org/wp-content/uploads/2021/08/PAI-Responsible-Sourcing-of-Data-Enrichment-Services.pdf>.

⁹ Gray, *supra* note 1; Milagros Miceli et al., *Studying Up Machine Learning Data: Why Talk About Bias When We Mean Power?*, ArXiv (Sept. 16, 2021), <https://arxiv.org/pdf/2109.08131.pdf>.

¹⁰ Kotaro Hara et al., *Worker Demographics and Earnings on Amazon Mechanical Turk: An Exploratory Analysis*, Extended Abstracts of the 2019 CHI Conf. on Hum. Factors in Computing Sys. 1 (2019).

¹¹ Gray, *supra* note 1.

¹² Carlos Toxtli et al., *Quantifying the Invisible Labor in Crowd Work*, ArXiv (Oct. 1, 2021), <https://arxiv.org/pdf/2110.00169.pdf>.

¹³ Turkopton, *Managed by Code: Worker Problems on Amazon’s Mechanical Turk Platform* (June 27, 2023), https://drive.google.com/file/d/1Krjs0obqHo5KgEG2ySLMKepiBy-s_NtP/view; Rebecca Tan & Regine Cabato, *Behind the AI Boom, an Army of Overseas Workers in ‘Digital Sweatshops’*, Wash. Post (Aug. 28, 2023), <https://www.washingtonpost.com/world/2023/08/28/scale-ai-remotasks-philippines-artificial-intelligence/>.

from their accounts or withheld payments, with no ability to appeal to a human.¹⁴ With little psychological or other support, data workers are exposed to traumatizing content such as torture, rape, and murder.¹⁵ Worker organizing is often penalized.¹⁶ It is not surprising that this work takes a huge toll on workers' mental health, causing anxiety, depression, and PTSD.¹⁷

These precarious labor conditions not only harm workers, but they can also undermine the quality of the AI systems. Workers are expected to screen out dangerous chatbot answers, but they may have little time to assess an answer's safety.¹⁸ Data workers are often given scant training or supervision, which can result in the introduction of bias.¹⁹ One study in the content moderation context found that YouTube banned LGBTQ content not because of an algorithmic error, but because the data workers who were rating content were located in a country where LGBTQ content was censored.²⁰ Finally, there may be few or no data protection processes in place, jeopardizing consumer privacy.²¹

Tech companies must not build AI on the backs of exploited workers. To help us better understand data workers' working conditions, we request written responses to the following questions by October 11, 2023.

For the purpose of the questions, the following definitions apply:

- The term “data work” includes the preparation, cleaning, and labeling of data for AI training, and review of AI outputs, including in real time.
 - The term “data workers” includes company employees, contractors, subcontractors, independent contractors, pieceworkers, and anyone else who your company compensates directly or indirectly for performing data work on any internal platform, or externally.
 - The term “your company” includes its contractors, subcontractors, and the platforms with which it works.
1. Does your company currently make publicly available information about the role data workers play in developing its AI technology? If so, please provide that information. If

¹⁴ *Id.*; Caroline O'Donovan, *This “Creepy” Time-Tracking Software Is Like Having Your Boss Watch You Every Second*, BuzzFeed News (Aug. 7, 2018), <https://www.buzzfeednews.com/article/carolineodonovan/upwork-freelancers-work-diary-keystrokes-screenshot>.

¹⁵ See Billy Perrigo, *OpenAI Used Kenyan Workers on Less Than \$2 per Hour to Make ChatGPT Less Toxic*, TIME (Jan. 18, 2023), <https://time.com/6247678/openai-chatgpt-kenya-workers/>; Karen Hao & Deepa Seetharaman, *Cleaning Up ChatGPT Takes Heavy Toll on Human Workers*, Wall St. J. (July 25, 2023), <https://www.wsj.com/articles/chatgpt-openai-content-abusive-sexually-explicit-harassment-kenya-workers-on-human-workers-cf191483>.

¹⁶ See Issie Lapowsky, *They Helped Score a Raise for Google AI ‘Raters.’ Then They Were Terminated*, Fast Co. (June 9, 2023), <https://www.fastcompany.com/90906492/google-bard-ai-raters-fired>.

¹⁷ See Hao, *supra* note 15.

¹⁸ @AlphabetWorkers, *supra* note 3.

¹⁹ Barbosa and Chen, *supra* note 4.

²⁰ Edd Gent, *The ‘Ghost Work’ Powering Tech Magic*, BBC (Aug. 29, 2019), <https://www.bbc.com/worklife/article/20190829-the-ghost-work-powering-tech-magic>.

²¹ Sergison, *supra* note 5.

not, why not? Please share any plans your company has to be more transparent about the role its data workers play and their working conditions.

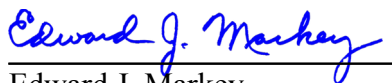
2. Please describe your company's use of data workers in the production of its AI products since 2020:
 - a. How many data workers performed work for your company each year, broken down by type of engagement (e.g., direct employee, contractor, subcontractor, independent contractor, or pieceworker)? If data workers are employed through contractors or subcontractors, please identify these contractors or subcontractors.
 - b. How do you identify data workers that are appropriate for each task?
 - c. What are the main tasks that data workers performed for your company, and which products are data workers used to create?
 - d. Does your company have an internal company platform for data work?
 - e. Does your company use external platforms for data work? If so, please identify them.
 - f. Do data workers perform work for your company both in the United States and abroad? If so, please identify the countries where data workers perform their work.
 - g. Does your company have written policies or requirements for data work on internal or external platforms? If so, please provide copies.
 - h. How does your company provide training and opportunities for professional development for its data workers? Please provide copies of training documents provided to data workers as part of their work.
3. What steps does your company take to ensure that data workers are paid a living wage?
 - a. Does your company have company-wide policies, or separate policies for internal or external platforms, addressing data worker compensation? If so, please provide copies. If not, why not?
 - b. Does your company set an hourly wage for data workers? If so, what hourly wage rates has your company required since 2020?
 - c. What steps does your company take to verify platforms' compliance with any company policies as well as applicable labor law?
 - d. For piecework, how does your company set a fair pay per task? For example, does your company run a task pilot to estimate task length, track time that data workers spend on each task, or use a living-wage calculator? If not, why not?
 - e. If your company estimates the time it will take to complete a task, does it include allowing for training time, messaging with requesters, and breaks? If not, why not?
4. What steps does your company take to ensure that all decisions about data workers, particularly those regarding suspension, withholding of pay, or other disciplinary action, are made with an opportunity to appeal to a human?

- a. Does your company's process to determine worker suspension or other disciplinary action require a human to review the decision? If so, please describe these processes. If not, why not?
 - b. When an algorithm flags a worker account for suspension, is there an opportunity to appeal to a human in all cases? If not, why not?
 - c. Are all stored wages paid out before a worker account is suspended? If not, please identify the annual amount of stored value forfeited by data workers whose accounts have been suspended since 2020.
 - d. Do data workers have the ability to reach a human with questions about their tasks, and receive timely, clear responses?
5. What does your company do to help data workers stay healthy and deal with the physical and psychological costs of their work?
 - a. Does your company make available mental health resources or mental health benefits to data workers dealing with traumatizing content? If so, please describe the resources or benefits. If not, why not?
 - b. Does your company guarantee data workers access to benefits or compensation for injuries sustained doing platform work, such as repetitive stress injury, or back pain? If not, why not?
 - c. Do your data workers have access to employer provided healthcare, sick days, and paid parental leave?
6. How can data workers' working conditions jeopardize model accuracy and introduce bias?
 - a. What steps does your company take to mitigate those risks?
 - b. Does your company have a process in place to ensure time allocated per task is sufficient from a product accuracy standpoint? If so, please describe the process. If not, why not?
 - c. Does your company have a process for ensuring that data workers are adequately trained and instructed on how to complete a task without introducing bias? If so, please describe the process. If not, why not?
 - d. Does your company have a process to detect possible biases that may be emerging from data work? If so, please describe the process. If not, why not?
7. How does your company maintain privacy protections?
 - a. What data does your company collect on your data workers? Does your company conduct surveillance of your data workers during their work, for example by logging keystrokes or taking pictures or video?
 - b. Does your company have a privacy policy for the data it collects from your data workers? If so, please provide a copy. If not, why not?
 - c. Does your company take steps to ensure that consumer privacy is maintained when data is transferred to data workers? If so, please describe those steps. If not, why not?

8. What standards, policies, or procedures does your company have to ensure data workers have freedom of association and collective bargaining rights at work?

For the sake of both the data workers and the AI products you are releasing, we urge you to make the safe working conditions and fair compensation of data workers a top priority. Thank you in advance for your attention to this important matter.

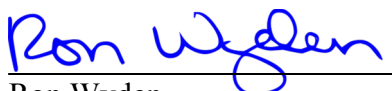
Sincerely,



Edward J. Markey
United States Senator



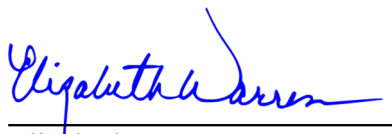
Pramila Jayapal
Member of Congress



Ron Wyden
United States Senator



Jamaal Bowman, Ed.D.
Member of Congress



Elizabeth Warren
United States Senator



Katie Porter
Member of Congress



Bernard Sanders
United States Senator



Mark Pocan
Member of Congress