



Managed by code: Worker problems on Amazon's Mechanical Turk platform

TURKOPTICON

Organizing mutual aid, resources, and advocacy to make Amazon Mechanical Turk work a good job while also improving conditions for all workers

By Turkopticon

June 27, 2023

This report, including research and analysis, has been compiled by Amazon Mechanical Turk workers and academic collaborators. For questions and follow up on these issues, please contact Lead Organizer Krystal Kauffman at policy@turkopticon.net.

Amazon Mechanical Turk Workers lack accountability and redress when they are managed at scale, and by code

The AI industry is worth trillions of dollars, producing “scientific innovation” and “artificial intelligence” through the work of hundreds of thousands of data workers. Amazon operates a platform, Mechanical Turk (AMT), with workers from around the world, but the majority are in the United States. Researchers estimate that as late as 2019, there were at least 250,000 AMT workers worldwide.¹ Like many gig work platforms, Amazon Mechanical Turk asserts that workers are independent contractors. As a result, they don’t enjoy a minimum wage or occupational health and safety protections. The International Labor Organization found in 2016 that the majority of Amazon Mechanical Turk workers had a college degree or higher. Overall, 37% reported that it is their primary income.²

When Amazon launched the service in 2006, Jeff Bezos explained that engineers could use workers just as they might write a piece of computer code.³ Engineers can create jobs, set prices, and incorporate work output straight into their code from Application Programmer Interfaces (APIs). Today, the platform forms a key piece of Amazon’s machine-learning web services, such as the SageMaker platform. Workers experience a lack of accountability and redress in practice from those who hire them on AMT, or Amazon itself. Both Amazon and some researchers who rely on our work on AMT collaborate with and/or receive funds from the US federal government through the National Science Foundation and other research and technology development agencies.

The following stories exemplify problems workers face when managed by code on AMT. Many of these problems are not so much consequences of the kinds of statistical artificial intelligence we hear so much about, but from complex layers of more familiar kinds of computer code that affect what work workers can access, whether or how they get paid, or whether they can retain their livelihoods through account access.

Turkopticon is not a union, but a worker-led group that advocates, organizes mutual aid, and creates resources to make AMT a good job while improving conditions for all workers.

¹ Robinson, Jonathan, Cheskie Rosenzweig, Aaron J. Moss, and Leib Litman. “Tapped out or Barely Tapped? Recommendations for How to Harness the Vast and Largely Unused Potential of the Mechanical Turk Participant Pool.” PLOS ONE 14, no. 12 (December 16, 2019): e0226394. <https://doi.org/10.1371/journal.pone.0226394>.

² Berg, J. (2015). Income security in the on-demand economy: Findings and policy lessons from a survey of crowdworkers. *Comp. Lab. L. & Pol’y J.*, 37, 543.

³ Irani, L. C., & Silberman, M. S. (2013, April). Turkopticon: Interrupting worker invisibility in amazon mechanical turk. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 611-620).

When requesters “mass reject” work and workers

AMT workers describe “mass rejections” as a devastating but all too common event when a requester – the employer seeking data work – rejects every HIT (Human Intelligence Task) a worker has submitted, or a large number of them. Though Amazon permits requesters to do this, assuming that “rejected” and unpaid for work is bad work, there is no process on the platform to verify this. It is highly unlikely that all workers submitting work have submitted bad work; thus, mass rejections are usually understood to be bad faith acts by the requester. Rejections negatively impact the workers' approval rating. This, in turn, limits the amount and quality of work available to the worker. Rejections stay on a worker's record indefinitely, even when the rejection is unwarranted. To rebuild their approval rating, workers often have to take the least desirable, most poorly paying tasks until their rating improves and gives them access to better work.

“As a Turker, I experienced a mass rejection when I returned to MTurk after several years. I had an excellent rating that was quickly tanked after I completed many HITS for a requester who rejected all of them. I have had to work very slowly to rebuild my rating over the past six months. After completing thousands of HITS, my rating has not even made it back up to 99%. I haven't had a single HIT rejected since experiencing the mass rejection. I try to be a careful and reliable worker, and at this rate, it may take me years, if ever, to recoup my rating. This limits what HITS I can qualify for and greatly reduces my options for work.”

“I'm new to MTurk (just started yesterday), and I first did 6 HITS by [requester], which were rejected. I did 2 more before I googled where the rejections were...all from [the requester]. I'm taking this as a learning lesson and have blocked [the requester] but now I need to get 7k or 8k approved HITS to bring it to 99%” (modified to anonymize)

“Seungho Kim, a requester, has been doing mass rejections for a while. I'm not certain why. I didn't have a single one of mine approved out of I believe it was 57 HITS. From reading other reviews, it's an ongoing issue...If I had those 57 HITS approved, I would not have had to work so hard just to get my approval rating up. I am not back to a 99% rating yet. I have 5k [tasks] until I get back to 99%. This mass rejection took me down to 96%...I reached out to this requester numerous times to hear nothing back.

If you see one of their HITS, it's pretty easy, but they leave some questionable grey spots, I guess. I reached out to them and asked for clarification on what exactly I did wrong to understand what they wanted. (i.e., if I could see it was a tiger, but the image was blurry, is it okay even though it's blurry and I can make it out?) I also asked if they would consider overturning my rejections. It has been, I believe, a month and a half since my rejections. I did not contact Amazon about this; honestly, I'm not even sure I know where to find their contact information for situations like this.”



In some countries, Amazon pays in company scrip

Amazon does not offer bank deposits to all workers, some are only able to redeem Amazon gift cards for their work – the same work that others get actual income from. While it appears to be based on workers' country of residence, it is not always clear to workers why they are not able to be paid in actual funds.

“I started working on MTurk between 2015 and 2016, and since then, I have been waiting to be allowed to withdraw my earnings to my US bank account. To be clear, I'm not a US resident. I can only use my MTurk earnings on Amazon.com, and I'm trapped in an unending profit cycle for Amazon: first, being a Turker, and second, being an Amazon.com customer (not by choice). Besides, although I have an excellent approval rating (99.9% with +210K approved HITs) and follow all the rules and Terms of Services, I'm still afraid that my account could be unfairly suspended (like it has happened to other Turkers) and that I might lose all of my accumulated earnings throughout all of these years. To use my funds in real life out of the Amazon ecosystem, I'm constantly searching for members of my family and close friends who want to purchase products from the Amazon.com website so that I can buy these for them in exchange for their money. It's unfair and unethical”



Outdated data can mean workers' peril

Demographics are an important part of the Mturk platform. Amazon collects worker characteristics such as age, gender, race, and education (optionally) that open access to, or restrict, access to certain work. Certain personal information about workers, such as gender and location, cannot be updated in Amazon's system. Workers understand this as a measure to prevent "scamming," or changing characteristics to access work (e.g. surveys) meant for certain demographics. However, this prevents workers from legitimately updating their characteristics such as when they affirm their gender, or when they move homes. As a result, workers who transition cannot qualify for work based on their actual, current status. For example, a worker whose address changes sometimes cannot qualify for work based on their actual, current location.

When the platform detects a difference between recorded data and detected data, such as an IP address location and recorded address, it can also flag the worker's account as suspicious. In some cases, this can result in worker account suspensions, blocking workers' access to their accounts, their work, and their accrued earnings.

"...the fact that Mechanical Turk can't/won't update locations (which are used as a qual for many studies) seems to exemplify how little effort they put into the platform. There are numerous people who have mentioned this so I am not surprised that 16 months after I moved from VA to CA, mTurk still lists me in VA, despite my updating my account immediately (& ordering things to be delivered here)"

“Masters”: an “black boxed” algorithmic promotion that benefits the few


AMT has a system called “work qualifications” to allow requesters to control what kinds of workers have access to their work. Qualifications can include whether the worker is an iPhone or Android user, whether the worker has passed a particular skill test, or whether the worker has been named a “Master” worker by the Amazon Mechanical Turk platform.

Amazon’s “Masters” qualification has been a particular frustration for workers on the platform. Amazon assigns the qualification to workers without explaining what assessments of behaviors resulted in the assignment. Requesters are able to easily create jobs that only go to Masters workers on the Amazon Mechanical Turk website. Amazon explains the Masters qualifications to requesters as follows: “Mechanical Turk has built technology which analyzes Worker performance, identifies high performing Workers, and monitors their performance over time. Workers who have demonstrated excellence across a wide range of tasks are awarded the Masters Qualification. Masters must continue to pass our statistical monitoring to retain the Mechanical Turk Masters Qualification.”

Most workers never get the Masters qualification, even when they have performed tens of thousands of tasks at extremely high rates of approval. Workers are also unsure if Amazon is even qualifying workers as Masters anymore, despite offering the qualification filter to requesters on the platform as a way to ensure high-quality workers. Since 2019, Turkopticon has not heard self-reports of new workers attaining this status on the platform. Amazon has never explained who gets the qualification or why they choose not to give it to others.

Many requesters believe that Masters automatically means higher quality data, which isn't necessarily the case. It only restricts workers from accessing their work dramatically. We argue that there are better ways for requesters to get quality data without excluding so many workers. Better methods that are more transparent to workers, such as total HITS completed by the worker, higher approval ratings, or “closed quals” – skill-specific tests requesters can have workers pass before hiring them on.

“it is understandable that the master’s qualification should be earned, but there is no way to know how, and unfortunately because it is offered without any real explanation of what it is, good workers who would otherwise be qualified to complete tasks do not get the opportunity to complete them”



Worker A: *In fairness, the fact that Masters hasn't been granted in years should make researchers think twice about using it as a required qualification.*

Worker B: *I don't think Amazon should offer it to researchers!*

Worker C: *I think it's more the labeling. They think they're getting something premium like when products put labels on their packaging that say "New and Improved" when the only thing new is the label.*

Worker D: *...Since Amazon charges them to use quals, I doubt seriously [that] they're being told that it's not as useful a qual [qualification]... Researchers (especially new ones, but seasoned ones too I would imagine) see "Masters" and likely think "oh, great -- better quality responses ... lemme use that real quick". :confused:*

...

Worker Z: *a lot of times they will end up removing the qual and reposting without it, so that goes to show it really does not help*

Worker E: *For them? No. For Amazon, yes, because Amazon gets a cut for the masters queue. [The worker is referencing the fact that Amazon takes a higher fee from requesters when it connects requesters to Masters workers.]*

Account suspensions: fired by algorithm

Amazon develops algorithmic supervision of workers by using feedback from job posters who have incentives and unchecked to punish workers for unsatisfactory work. Workers flagged by algorithms find themselves locked out of their accounts with no explanation.

What makes a worker suspicious to algorithmic management used by Amazon or requesters? First, requesters can withhold pay for work they do not like or understand and this, in turn, lowers workers' approval rating on the platform. Amazon's patents also suggest that the platform also rates workers by subjecting them to hidden tests and comparing their answers to other workers. Employers often fail to train workers to produce the kinds of results they want, and offer no pathways for workers to repair mistakes or misunderstandings. Workers lack formal processes for contesting these evaluations and associated wage theft. Amazon then uses these evaluations – job poster pay withholding and hidden tests – to develop machine learning models that guess whether a worker is a “bad actor.”⁴

One worker Turkopticon helped had lost her account and access to her earnings, stored on the platform, for two weeks over what Amazon later admitted it was a mistake. The platform's algorithm flagged her account because it detected two people – her and her son, it turned out – logged in and working on the same wifi network. It may also be the case that the algorithm had “learned” that low quality workers shared wifi and flagged the account. On her suspension, the worker appealed to Amazon but got no response until Turkopticon, a Turk worker organization, approached Amazon collectively. Only then did Amazon re-investigate and admit that it made a mistake. An analysis of Amazon patents reveals that Amazon uses a surveillance of a wide range of behaviors to decide whether to trust the worker logged in behind the screen – a worker they will almost never meet – or to suspend them.

⁴ Irani, Lilly. “When Managers Rely on Algorithms of Suspicion: Fraud Logics and Their Fallouts.” Centre for International Governance Innovation (blog), July 4, 2022. <https://www.cigionline.org/articles/when-managers-rely-on-algorithms-of-suspicion-fraud-logics-and-their-fallouts/>. For fuller explanation of these issues and analysis of Amazon patents on algorithmic control, see Irani, Lilly. “Algorithms of Suspicion: Authentication and Distrust on the Amazon Mechanical Turk Platform.” SSRN Working Paper 4482508. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4482508