James Yang

DATA 512A

10/20/22

Crowdwork Ethnography

## Step 1: A Record of Prior Knowledge of MTurk

To my understanding, Amazon Mechanical Turk is a platform where people can do tasks and get rewarded for the completion of these tasks. From my perspective, these completions are a way for the company to collect data on humans through medial tasks that simulate different cognitive thinking. If I were to host a huge catalog of tasks with varying weighted rewards, my purpose behind this tool would be to collect trafficable data through categorical outputs. According to a lecture from Dr. McDonald, he states that Amazon Turk is a crowdsourcing platform where crowdsourcing is a service of obtaining information from many people via the internet.

Something that crowdsourcing platforms typically boast about is the ease of the tasks and the flexibility that they provide. They most likely won’t require the worker to go into a physical office, so a lot of their tasks could be completed from home. There is most likely a lack of a deadline as well, meaning the projects could be started and finished at the convenience of the worker.

## Step 2: Preview or Complete as least 10 Mechanical Turk Hits

The title of the tasks such as “Research Rewards” and “Adam Pragmatics” don’t necessarily reveal what the tasks truly are. At first glance, I would personally not be interested in these hits as a potential real Turker because they aren’t exactly stimulating activities. From my understanding, a lot of the tasks within the preview are simply answering questions and figuring out simple layout puzzles. More expensive hits such as “VacationRentalAPI” require a qualification for the task and then a short description of the requirements. To simplify the format, I have introduced a list of the hits I have analyzed.

### Tweet Classification / Vacation Rental API

The HIT is asking the user to perform a classification label for sets of tweets given to them. For Vacation Rental API, it is essentially looking at the API text and identifying what they would look at first. In order to complete the HIT successfully, we need to have a fluent understanding of English and contextualization when analyzing a tweet that contains an ad, sarcasm, etc. Based on the time allotted of 60 minutes and the money reward of $0.01 and 60 minutes for $7.50 respectively, I can safely say that this task is not worth an MTurk’s time. Based on the preview, the HIT represents almost a perfect example of what I thought MTurk would be. The reason is because it is simply having the user classify labels for training datasets with very little cost. The instructions are clear and I would be able to complete this HIT correctly and quickly based on the instructions because it is a fairly medial task that any fluent English speaker could complete. Based on the preview, this is not work that I would be interested in doing because it is not stimulating. I would not be willing to work on any amount hits like this one because it is not worth developing my career behind. I think that this hit is reasonable for some, but not personally for me. There was nothing behind it that was uncomfortable or unpleasant. The purpose of the HIT is clear to me, I understand that this is simply a way of generating more labelled data for models to be used behind NLP.

### Danna / CopyText Inc

Danna is a hit that requires you to identify objects brought up on the screen through a series of multiple choice options. CopyText Inc is similar in interpreting the text from a receipt and want the worker to copy the text. In order to complete this, the worker just must have the cognitive ability to identify objects seen on the screen. If they can identify a water bottle or a scooter, then they are able to complete the task. Based on the preview and time allotted, the reward of $0.06 for 30 minutes and 60 minutes for $0.05 seems fair enough for the amount of work that it requires. Based on the preview, the hit represents a task that would be on mechanical turk because it is an identification problem for models to label training sets of data. The instructions are clearly provided and I think I would be able to quickly finish this hit because it requires very basic cognitive ability. Based on the preview, I would not be interested in doing work on Amazon mechanical turk because it is simply not something I would want to invest my time on. I would not be willing to work on 10 hits of this or even 100 or 1000 because it simply would not be worth my time. There is nothing about this hit that seems unreasonable or inappropriate to me, and neither is there anything that I find unpleasant or uncomfortable. The purpose of the hit is clear to me and I understand how this could generate good information for computer vision training sets.

### SFR / IIS Academia

SPR is a hit that is asking you to perform identification of image naming. For IIS Academia it is asking you to show agreement between a pair of sentences. In order to this, you need to have a fundamental understanding of general naming convention for photos found on a computer as well as the identifiable situation of sentences. Based on the preview and the time allotted, $0.10 for 30 minutes and 120 min for $0.27 seems somewhat better in comparison to the other turk tasks. The task is what I thought would be on mechanical turk because it is an identification process for NLP that allows for training data to be labelled for output of picture values and what is the most common value assigned to each image. The instructions are clear and I would be able to complete this hit correctly and quickly because I would be able to identify the image and label it as such. Based on the preview, I would still not be interested in doing Amazon turk because it is not something I want to invest my time in. I would not be willing to work on 10 hits or 100 or even 1000 because it is not something I would want to invest my time in. There is nothing that is unreasonable or inappropriate to me and it doesn’t seem unpleasant or uncomfortable. The purpose of hit is clear to me, it is used to collect data for labels for both nlp and computer vision models.

### cZqjwwRWXGZVZDzfgbyw / Adam Pragmatics

This work is asking to listen to audio clips and make judgements based on what the article is saying. They are asking for how the clips sound with regards to the biological sex and what kind of accent is found and how natural the voice is. In the Adam Pragmatics, the task is to chat with computer assistants and interpret what they are saying to get helpful leads. In order to do this hit, we will need to have a sense of deciphering in the universal accents and understand how words are being enunciated. For the request of 120 min on Adam Pragmatics, $0.20 is simply not enough. For the audio listening, based on the preview, the hit represents a task that I thought would be on mechanical turk because it is showcasing an example of collecting labeled training data for audio recognition models. The instructions provided are a bit convoluted within the descriptions based on what the overarching response is. I think that this work may be a bit difficult to finish because its sometimes difficult understanding audio recordings and deciphering what they are saying. Based on the preview, this work is not something I would be interested in doing mechanical turk because $0.15 for 20 minutes of time is simply not enough money for me to invest on the internet. I would not be willing to do this 10 or 100 or 1000 times because it’s not something that I believe would be beneficial in my career development. There may be something that is unreasonable or inappropriate depending on the language used within the auto clips. The purpose of the hit is clear to me in what data they are trying to collect.

### BBL /MLDataLabeler

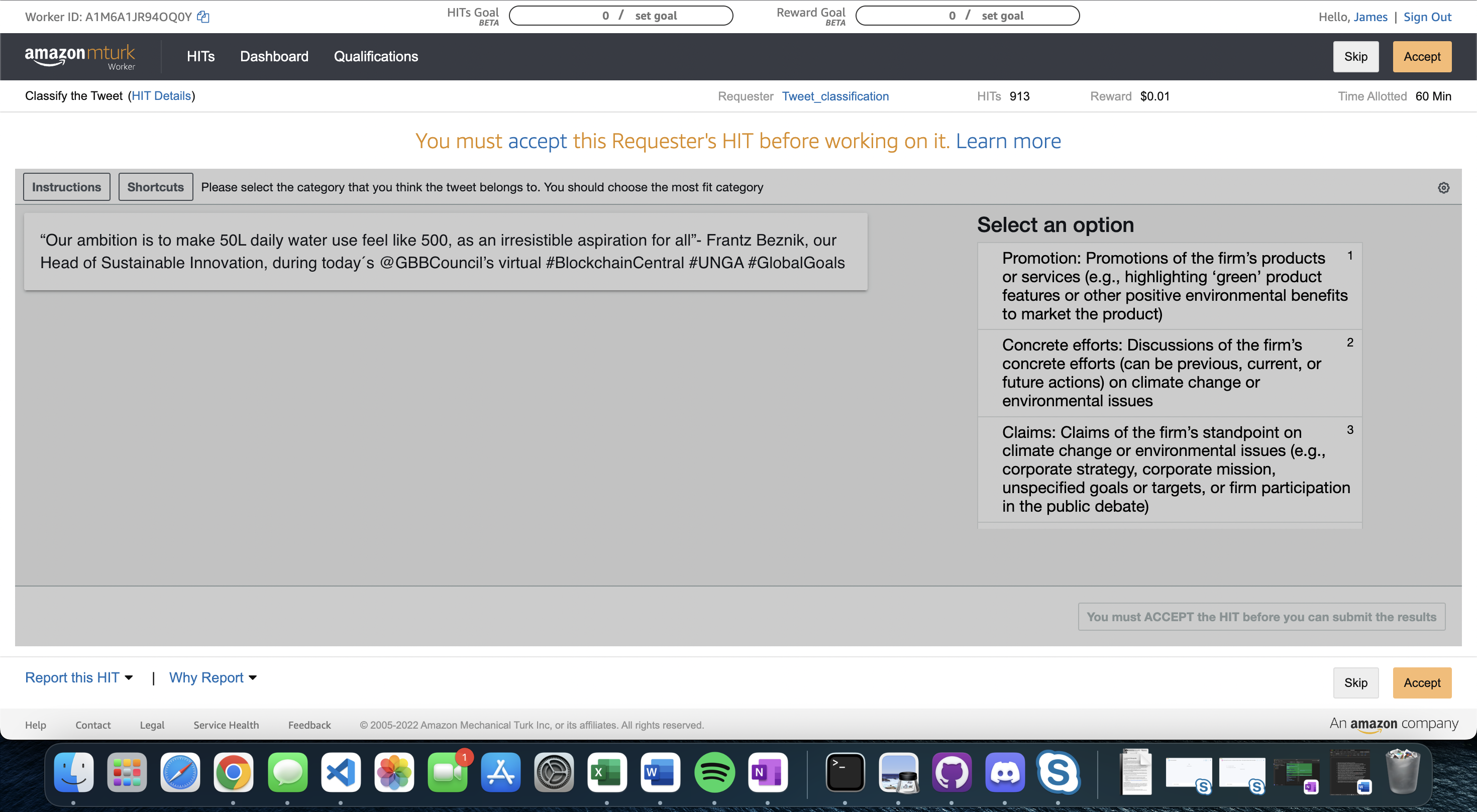
This hit is asking to analyze two videos and give opinions on what was seen. In order to do this hit successfully, we need to interpret what was on the screen. In the MLDataLabeler, it gives a video of a man and noise and asks the user to draw a box around the man. Based on the preview and the time allotted, $0.53 for 30 minutes seems more fair than the previous turk hits however it isn’t what I believe to be the best output for the time invested. In the MLDataLabeler, it is asking for $0.03 for every 5 minutes which translates to about $0.15 for the same 30 minutes from the other task. The instructions provided on the hit are complete and clear. I believe that I could complete the hit because it is very simplistic and open ended based on user response. There is no formulaic response. Based on the preview, I still would not be interested in doing mechanical turk because it is not something I would want to invest my time in. Watching videos that I do not inherently have interest in isn’t something that sounds enjoyable with my time. I would not be willing to work on 10 hits like this one or even 100 or 1000 because it does not appeal to my time. Nothing in this hit seems unreasonable or inappropriate to me, and nothing seems unpleasant or uncomfortable. The purpose of this hit isn’t quite clear to me because interpretation of videos can be so broad that it could lead to uncorrelated responses. This could lead to significant and unsignificant results.

### Well Labs

This hit is asking to label categories, item names, and quantity amounts based on the image given in the prompt. In order to complete this, the worker must be able to successfully interpret an image and answer the general questions found on the label of the item. Based on the preview, $0.08 for 20 minutes of time isn’t the best use of time. This task represents what I thought mechanical turk would be because it is a way of putting a label on training sets for labels and computer vision. The instructions provided are clear because it is a simple input to output result that allows for easy enabling of opinion. I personally would not be willing to be work on 10 hits like this one or 100 or 1000 because it is a bit too tedious for my liking and I don’t think I would be developing my career very well by spending time on things such as this. There isn’t anything that seems unreasonable or inappropriate about this hit and it doesn’t make me unpleasant or uncomfortable. The purpose of this hit is clear to me. I understand that this is used to develop labelled data.

### Step 2 Reflection

When reflecting upon the hit previews, I learned that mechanical Turk workers complete tasks which require minimal amounts of stimulating work. A lot of these tasks can be completed with little focus, and many times have no regard for whether an input is correct or not.



**Figure 1.1: Depicts an example of the twitter classification Hit found on Amazon Turk.**

The twitter classification Hit for example requires a select of choices that will give you $0.01 after answering a few questions in an allotted time of 60 minutes. Workers would have to answer 100 times to get a dollar which means that they would need to answer 1500 times an hour to barely get a minimum wage rate. From the perspective of a worker, they are not doing this for the sake of payment but rather boredom and random amounts of enjoyment. It is a crowdsourcing platform that requires little amounts of work to receive little amounts of payment; when a person dedicates enough time and accrues enough reward, they can sometimes buy a small purchase to feel a sense of accomplishment. From the perspective of a requester, this is one of the best platforms to pull trained label data from because it is coming directly from an anonymous unbiased human. For NLP, computer vision, mathematical scaling, and many other types of machine learning models, this is a great source for requestors. The cost of putting one of these hits will most likely cost very little and the return that they need to give to the mechanical Turk workers is also very little. Something that I have also learned from previewing these hits is that I am at a crossroad ethically with the amount of reward that they are receiving. If a worker spent 30 minutes a day on these hits and received $0.05 each day, they would accumulate $1.50 after a month and 15 hours of investment in time. That 15 hours could most likely have been put into a minimum wage job that would have given them $225 instead, or a coursera class to benefit their career.

## Step 3: Read the Documentation of how to be a good requester

When reading the documentation, I noticed a reoccurring theme of making the worker feel comfortable about the task that they are completing.