
Amazon Mechanical Turk Tutorial

Leo Goldman

leo.goldman@columbia.edu

Columbia Experimental Lab For Social Sciences

<https://celss.iserp.columbia.edu/>

Last Update: April 2020

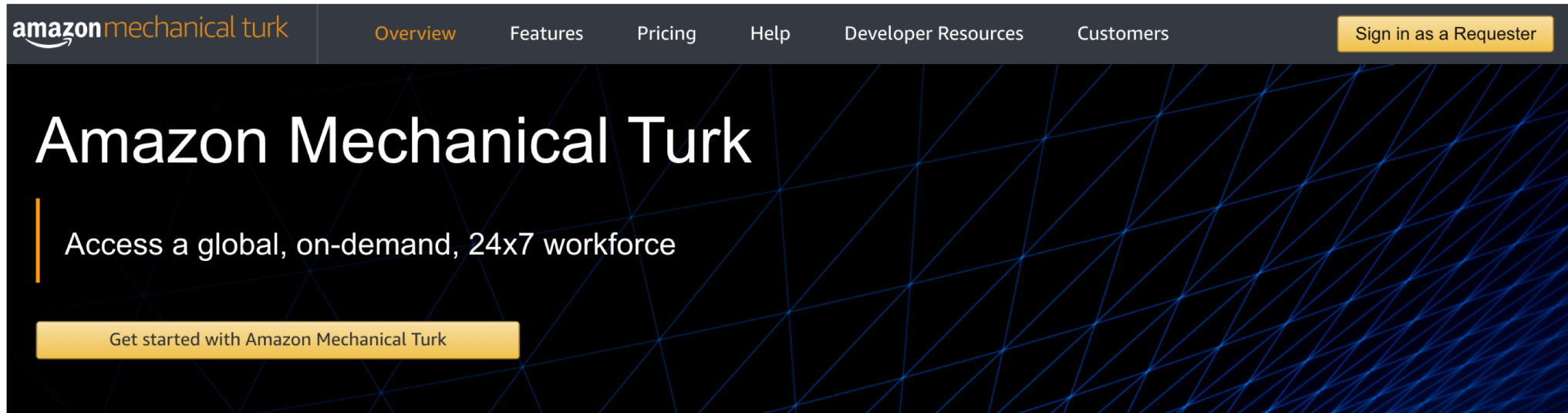
Today's Goals

- Learn about Amazon Mechanical Turk (MTurk)
- Learn how to launch and run your experiment on MTurk
- Keep in mind the necessary tasks post-launch
 - Worker qualifications
 - Bonusing workers
 - https://github.com/leogoldman/Qualtrics_MTurk_Guide

Amazon Mechanical Turk

“Artificial artificial intelligence”

Amazon Mechanical Turk – What is It?

The image shows the top portion of the Amazon Mechanical Turk website. At the top left is the logo 'amazonmechanical turk'. To its right is a navigation menu with links for 'Overview', 'Features', 'Pricing', 'Help', 'Developer Resources', and 'Customers'. On the far right of the navigation bar is a yellow button that says 'Sign in as a Requester'. Below the navigation bar is a large dark blue banner with a white geometric pattern of lines. The banner contains the text 'Amazon Mechanical Turk' in a large white font, followed by the tagline 'Access a global, on-demand, 24x7 workforce' in a smaller white font. At the bottom of the banner is a yellow button that says 'Get started with Amazon Mechanical Turk'.

Amazon Mechanical Turk (MTurk) is a crowdsourcing marketplace that makes it easier for individuals and businesses to outsource their processes and jobs to a distributed workforce who can perform these tasks virtually. This could include anything from conducting simple data validation and research to more subjective tasks like survey participation, content moderation, and more. MTurk enables companies to harness the collective intelligence, skills, and insights from a global workforce to streamline business processes, augment data collection and analysis, and accelerate machine learning development.

While technology continues to improve, there are still many things that human beings can do much more effectively than computers, such as moderating content, performing data deduplication, or research. Traditionally, tasks like this have been accomplished by hiring a large temporary workforce, which is time consuming, expensive and difficult to scale, or have gone undone. Crowdsourcing is a good way to break down a manual, time-consuming project into

Amazon Mechanical Turk – What is It?

- “MTurk”
- Marketplace for work/tasks (more than 100K workers, more than 2K active)
- Outsource computer labor to humans
 - Named after 18th century, “The Turk,” a chess robot that was actually controlled by a chess master
 - Used largely to create labeled data for machine learning
- Tasks range from several seconds (identify photo) to hour-long surveys (maybe yours!)
- Compensation ranges from \$0.01/task to above minimum wage
- Open to almost anyone with an Amazon account
- Not really regulated

<https://www.mturk.com/worker/help>

MTurk – Vocabulary

- Workers and Requesters
- Human Intelligence Tasks, or “HITs”
- Reward (Show-Up/Completion Fee)
 - Baseline amount being offered for completion of a task
- Bonus
 - Additional compensation that can be based on performance
 - Payment can be primarily based on bonus
- Batch
 - Requester creates a batch of HITs

MTurk – Low Wages, High Fees

- Pay can be as low as \$0.01/task
- Amazon takes (high) fees on everything
 - 20% on all rewards and bonuses
 - Additional 20% on batches with more than 10 assignments (i.e. almost all batches)
 - Additional fees for advanced worker restriction
- Research shows median wage ~\$2/hour
 - 4% of workers making more than \$7.25/hour
- Average requester pays \$11/hour but posts far less work

<https://requester.mturk.com/pricing>

[A Data-Driven Analysis of Workers' Earnings on Amazon Mechanical Turk \(2017\)](#)

MTurk – Who Takes My Survey?

- Anyone with an Amazon account is eligible by default
 - In the US, accounts are linked to social security numbers
- May be machines trying to complete your HIT
- Can restrict who takes your survey
 - Only those in US
 - No one who has taken one of my surveys before
 - Only workers with a high HIT Approval Percentage (important)
- Advanced qualifications and restrictions (expensive)
 - MTurk Masters (additional 5%)
 - Restrict by political affiliation (\$0.40/worker)

MTurk – Who Takes My Survey?

- Workers are quick
 - Can get hundreds of responses in a few hours
 - Depends on your pay rate, survey, etc.
- Tradeoff between quality and speed
 - Ideally you want to restrict to workers with >99% HIT approval
 - Noticeable difference between >95% and >99% (>97% seems okay)
- Note that some people may be bad subjects (careless, machines)
 - Incentive good performance (bonuses for getting certain questions right)
 - Add questions to see if people are paying attention

MTurk – Who Takes My Survey?

- Workers qualifications
 - Existing ones include HIT Approval Percentage
- Can create your own
 - Label for previous workers will let you exclude people who have taken your survey before
 - Can create any label to filter towards or against certain workers you've had
- Prev_Worker and Eligible_Worker are very useful qualifications
 - Will show in demo

MTurk – Demographics

- Primarily workers in US and India
- Fairly even gender split
- Tend to be a bit younger than US population
- Fairly even married/single split
- Demographics tracker (below) lets you look at several demographic characteristics, overall and by US/India

<http://www.ipeirotis.com/wp-content/uploads/2017/12/wsdmf074-difallahA.pdf>

<http://demographics.mturk-tracker.com>

MTurk – Worker View

HIT Groups (1-20 of 967)

+ Show Details

- Hide Details

Items Per Page: 20

Requester	Title	HITS	Reward	Created	Actions
+ Shopping Receipts	Extract Data From Shopping Receipt	287	\$0.01	15s ago	Preview Qualify
+ Job Spotter	Verify photos of hiring sign and store name	8	\$0.03	16s ago	Preview Qualify
+ Panel	Identify if two receipts are the same	1	\$0.05	13s ago	Preview Accept & Work
+ James Billings	Market Research Survey	10,116	\$0.01	20s ago	Preview Accept & Work
+ Crowdsurf Support	Review, edit, and score the transcription of up to 35 seconds of media - Earn up t...	2,371	\$0.02	21s ago	Preview Qualify
+ str11223344	Tell us how much this item would cost to replace - General Contents - Batch ID #...	50	\$0.20	12s ago	Preview Qualify
+ Emily Wilson	Continue a conversation with a chatbot	3	\$0.10	38s ago	Preview Accept & Work
+ CopyText Inc.	Type the text from the images, carefully. Productivity and bonuses guaranteed.	21	\$0.01	43s ago	Preview Accept & Work

MTurk – Example HIT

You must [accept](#) this Requester's HIT before working on it. [Learn more](#)

Extract purchased items from a shopping receipt

HIT Reward \$0.03 for all receipt items. Hit Add Row for next Item

Real readable original receipt

Not a receipt or not readable

Rotate Image

Zoom: - 100% +



Enter Retailer Name *

If there is no Retailer Name on the receipt, mark receipt as "Not a receipt or not readable."

Retailer name

If no full address then only supply what is shown on the receipt. Leave fields with no entries empty.

Phone

#####

Street

Street

City

City

State

e.g. Califor

Zipcode

Zipcode

Date (Use 4-digit year)*

mm/dd/yyyy

Time

HH:MMa/p

Please enter entire **Item Description** including UPC/Item number.

Pounds etc. should be added to the **Item Description**, for example a valid Item Description can be "Bananas 2.88lb @ 0.99/lb".

MTurk – Worker View

HIT Groups (1-20 of 904)

+ Show Details

- Hide Details

Items Per Page: 20

Requester	Title	HITs	Reward	Created	Actions	
+ Labor Experiment20	Compensation HIT for A1NM7ZPZ3NH412(~ 1 minutes)	1	\$22.25	28m ago	Preview	🔒 Qualify
+ AB Lab - University of Calgary	Compensation HIT(~ 1 minutes)	1	\$20.00	19d ago	Preview	🔒 Qualify
+ SurveyComet	Make predictions about political events \$16. (~ 120 minutes)	25	\$16.00	24m ago	Preview	🔒 Qualify
+ SurveyComet	Make predictions about political events \$16. (~ 120 minutes)	25	\$16.00	24m ago	Preview	🔒 Qualify
+ SurveyComet	Make predictions about political events \$16. (~ 120 minutes)	25	\$16.00	25m ago	Preview	🔒 Qualify
+ SurveyComet	Compensation HIT for A1EBQ9X6IN5OZC(~ 1 minutes)	1	\$16.00	3d ago	Preview	🔒 Qualify
+ SurveyComet	Make predictions about political events \$16. (~ 120 minutes)	25	\$16.00	10/24/2019	Preview	🔒 Qualify
+ TAP Lab FSU	Understanding people's prior experiences	1	\$15.00	2d ago	Preview	🔒 Qualify
+ TAP Lab FSU	Understanding people's prior experiences	1	\$15.00	3d ago	Preview	🔒 Qualify
+ Peter Gault	Four part English language study - part three(~ 60 minutes)	2	\$15.00	1h ago	Preview	🔒 Qualify
+ UserBob	Upload 25 minute screencast of an Android app	1	\$13.75	8d ago	Preview	🔒 Qualify

MTurk – Creating a Batch

Select a customizable template to start a new project

Survey

- Survey Link
- Survey

Vision

- Image Classification
- Bounding Box
- Semantic Segmentation
- Instance Segmentation
- Polygon
- Keypoint
- Image Contains
- Video Classification
- Moderation of an Image
- Image Tagging
- Image Summarization

Language

- Sentiment Analysis
- Intent Detection
- Collect Utterance
- Emotion Detection
- Semantic Similarity
- Audio Transcription

Survey Link Instructions (Click to collapse)

We are conducting an academic survey about social networks. We need to understand your opinion about social networks. Select the link below to complete the survey. At the end of the survey, you will receive a code to paste into the box below to receive credit for taking our survey.

Make sure to leave this window open as you complete the survey. When you are finished, you will return to this page to paste the code into the box.

Template note for Requesters - To verify that Workers actually complete your survey, require each Worker to enter a **unique** survey completion code to your HIT. Consult with your survey service provider on how to generate this code at the end of your survey.

Survey link:

Provide the survey code here:

You must ACCEPT the HIT before you can submit the results.

MTurk – Creating a Batch

Edit Project

1 Enter Properties

2 Design Layout

3 Preview and Finish

Project Name:

This name is not displayed to Workers.

Describe your survey to Workers

Title

Describe the survey to Workers. Be as specific as possible, e.g. "answer a survey about movies", instead of "short survey", so Workers know what to expect.

Description

Give more detail about this survey. This gives Workers a bit more information before they decide to view your survey.

Keywords

Provide keywords that will help Workers search for your tasks.

MTurk – Creating a Batch

Setting up your survey

Reward per response

This is how much a Worker will be paid for completing your survey. Consider how long it will take a Worker to complete your survey.

Number of respondents

How many unique Workers do you want to complete your survey?

Time allotted per Worker

Maximum time a Worker has to complete the survey. Be generous so that Workers are not rushed.

Survey expires in

Maximum time your survey will be available to Workers on Mechanical Turk.

Auto-approve and pay Workers in

This is the amount of time you have to reject a Worker's assignment after they submit the assignment.

MTurk – Creating a Batch

Worker requirements

Require that Workers be Masters to do your tasks ([Who are Mechanical Turk Masters?](#))

Yes No

Specify any additional qualifications Workers must meet to work on your tasks:

[\(+\)](#) Add another criterion (up to 5)

(Premium Qualifications incur additional fees, see [Pricing Details](#) to learn more)

Project contains adult content ([See details](#))

This project may contain potentially explicit or offensive content, for example, nudity.

Save

Design Layout

MTurk – Creating a Batch Demo

MTurk – Best Practices

- Give a decent wage
 - Some researchers argue that you should pay what you would in a lab
 - I usually recommend to researchers pro-rating \$7.25/hr (federal minimum wage)
- Pay workers in a timely fashion
 - Use a quick auto-approve setting
 - Workers can see how long you take to pay on average

MTurk – Best Practices

- Be responsive
 - Workers can reach out with questions, comments, mistakes
- Be mindful that workers have communities
 - reddit.com/r/mturk
 - See https://github.com/leogoldman/Qualtrics_MTurk_Guide for helpful links

MTurk – Managing Results

- Batch progress/results
- Approve workers
 - "Approving" workers pays out their reward
 - Bonus must be done separately
- Update worker qualifications
 - Keep track of who has taken your experiments and who is eligible
 - Useful if you don't want people to take your survey again, or if you want to allow only certain people to participate in another round.


MTurk – Managing Results


Manage Batches

Click on the name of the batch to see more details

▶ Batches in progress (0)

▼ Batches ready for review (4)

		Review Results	Delete
Created:	June 07, 2019	Assignments Completed:	200 / 200
Time Elapsed:	7 days	Estimated Completion Time:	COMPLETE
Batch Progress:			
	100% submitted	100% published	

		Review Results	Delete
Created:	May 31, 2019	Assignments Completed:	200 / 200
Time Elapsed:	7 days	Estimated Completion Time:	COMPLETE
Batch Progress:			
	100% submitted	100% published	

MTurk – Managing Results

[Manage Batches](#) > **Review Results**

Review Results

Select the check boxes on the left to approve or reject results. You only pay for approved results. To evaluate results offline, select Download CSV.

For additional batch information, [view batch details](#).

Customize View

Filter Results

Upload CSV

Approve All

Download CSV

0 of 200 assignments (FILTER APPLIED: only show assignments that are in 'Submitted' status)

No results match your filter

0 of 200 assignments (FILTER APPLIED: only show assignments that are in 'Submitted' status)

MTurk – Managing Workers

- Might have to update worker qualifications between every batch
 - For example, if you don't want people taking your survey twice
- Paying out bonuses is done individually if done on the website
 - Use MTurk developer to pay out bonuses in bulk
 - https://github.com/leogoldman/Qualtrics_MTurk_Guide

MTurk – Managing Workers Demo

MTurk – Paying Bonuses

- Workers should get a survey code or some kind of ID in from your experiment

Select a customizable template to start a new project

Survey

- Survey Link
- Survey

Vision

- Image Classification
- Bounding Box
- Semantic Segmentation
- Instance Segmentation
- Polygon
- Keypoint
- Image Contains
- Video Classification
- Moderation of an Image
- Image Tagging
- Image Summarization

Language

- Sentiment Analysis
- Intent Detection
- Collect Utterance
- Emotion Detection
- Semantic Similarity
- Audio Transcription

Survey Link Instructions (Click to collapse)

We are conducting an academic survey about social networks. We need to understand your opinion about social networks. Select the link below to complete the survey. At the end of the survey, you will receive a code to paste into the box below to receive credit for taking our survey.

Make sure to leave this window open as you complete the survey. When you are finished, you will return to this page to paste the code into the box.

Template note for Requesters - To verify that Workers actually complete your survey, require each Worker to enter a **unique** survey completion code to your HIT. Consult with your survey service provider on how to generate this code at the end of your survey.

Survey link:

Provide the survey code here:

You must ACCEPT the HIT before you can submit the results.

MTurk – Paying Bonuses

Qualtrics

DN	DO	DP	DQ	DR	DS
Q167	Q60	Random ID	Create New	TreatmentNum	
There is	In general,	Random ID	Create New	TreatmentNum	
:"Import	:"ImportId	:"ImportId":	:"ImportId":	:"ImportId":	:"TreatmentNum"}
		8723325			
		8625861		3	
		9035965			
		1404854			
		4348471			
		4449691			
		5627748			
		5039411		3	
		6192382			
		1048752			
		7354452			

MTurk

AA	AB	AC	AD	AE
Last7DaysAp	Answer.consent	Answer.surveycode	Approve	Reject
100% (1/1)	I agree	6904478		
100% (1/1)	I agree	7854627		
100% (1/1)	I agree	4348471		
100% (1/1)	I agree	5653640		
100% (1/1)	I agree	4449691		
100% (1/1)	I agree	4324067		
100% (1/1)	I agree	4723315		
100% (1/1)	I agree	6563157		
100% (1/1)	I agree	1668177		
100% (1/1)	I agree	5135763		
100% (1/1)	I agree	6658082		

MTurk – Paying Bonuses

- Once you've downloaded both the Qualtrics results and MTurk results, have to link workers to their Random ID
 - Excel VLOOKUP or INDEX/MATCH
 - Stata merge
 - R merge or which()

MTurk – Developer

- Use to avoid paying out bonuses individually
- Requires Amazon Web Services (AWS) account
 - Free, easy to set up
- Can use a variety of languages
 - Command line (Terminal)
 - Python

Simplify using AWS services, including MTurk, in your applications with an API tailored to your programming language or platform. These SDKs are designed to work with all AWS services.

Browser

[Install »](#)
[Documentation »](#)
[Learn more »](#)

Java

[Install »](#)
[Documentation »](#)
[Learn more »](#)

.NET

[Install »](#)
[Documentation »](#)
[Learn more »](#)

Node.js

[Install »](#)
[Documentation »](#)
[Learn more »](#)

PHP

[Install »](#)
[Documentation »](#)
[Learn more »](#)

Python (Boto)

[Install »](#)
[Documentation »](#)
[Learn more »](#)

Ruby

[Install »](#)
[Documentation »](#)
[Learn more »](#)

Go

[Install »](#)
[Documentation »](#)
[Learn more »](#)

C++

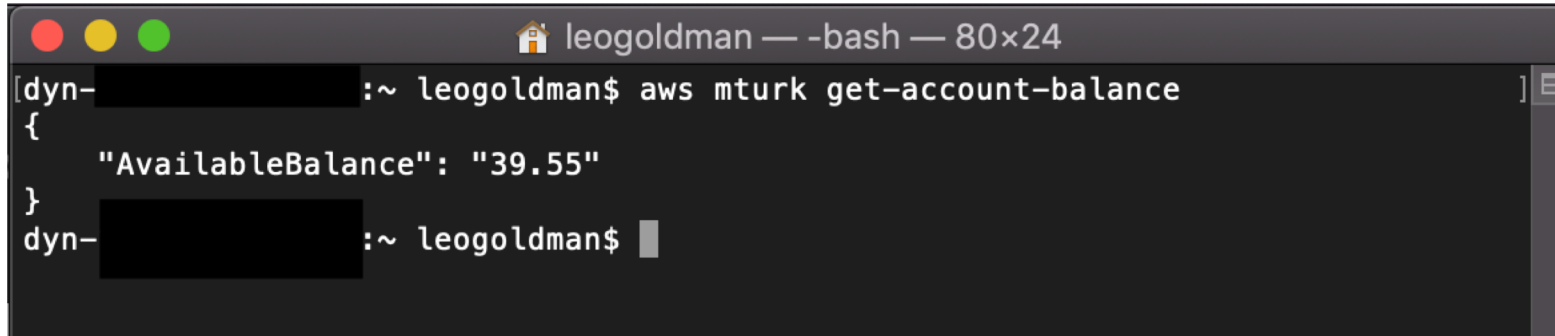
[Install »](#)
[Documentation »](#)
[Learn more »](#)

AWS Command Line

Interface
[Install »](#)
[Documentation »](#)
[Learn more »](#)

MTurk – Developer

- Once you connect your AWS account, can do most things from Developer
 - Paying bonuses in bulk
 - Assigning qualifications more easily



```
leogoldman — -bash — 80x24
[dyn- [REDACTED]:~ leogoldman$ aws mturk get-account-balance
{
  "AvailableBalance": "39.55"
}
[dyn- [REDACTED]:~ leogoldman$
```


Thank You!

Come to CELSS consulting hours!

<https://celss.iserp.columbia.edu/>

Wednesdays 2:30-5pm

Thursdays 10:30am-1pm

leo.goldman@columbia.edu