

# ATAI Tutorial on CrowdSourcing Task

## Handout: Design a crowdsourcing pipeline for knowledge graph completion

*In this handout, we provide an example of a crowdsourcing pipeline to verify knowledge of facts in a knowledge graph. In this pipeline, we use one verification task that is described in Step 1. First, you will study the description and the input of the tasks. Next, you will experiment with the tasks in the MTurk environment while understanding how to work with such a platform.*

### Step 1: Get familiar with the tasks in the pipeline:

#### Task 1: Knowledge graph completion

**Goal:** The goal of this crowdsourcing task is to identify the relation between the entities and provide supporting evidence.

**Input:** Two entities from a knowledge graph and multiple choice answers.

**Task Description:** We give the crowd a short text indicating two entities, and we ask them to find a suitable relation between the entities. They can refer to the internet as many times as necessary.

Knowledge Graph Completion Instructions (Click to collapse)

**Overview:**  
We need your help to verify the knowledge or facts in a knowledge graph, which is further used by machine learning algorithms for some downstream tasks.

**Steps:**

- Read the given concepts carefully.
- Find out the best suitable relation that defines the given entities or concepts.
- You can refer to Wikipedia, which helps you to answer step 2.
- Identify the sentences or text which supports your answer.
- Please provide us the Wikipedia link and the sentences or text that support your answer in the given textbox.

Figure 1: Task Instructions (Part of it)

Knowledge Graph Completion Instructions (Click to expand)

InstructionsShortcuts

Select the relation which defines suitably the two entities in the given sentence.

Nicholas Bird is \_\_\_\_\_ of Finding Nemo

Question 1: Which answer correctly fills in the blanks?

☐ review score ☐ director ☐ voice actor ☐ box office ☐ character

Question 2: Which wikidata or wikipedia link did you use to arrive at your answer [required]?

\_\_\_\_\_

Question 3: Which sentence(s) or information from wikidata or wikipedia did you use to arrive at your answer [required]?

\_\_\_\_\_

Figure 2: Task Template

## Step 2: Try out the task (as a worker)

1. Log in to <https://workersandbox.mturk.com/> (assuming you have created your account and have the required credentials with you)
  - Username: <<ADD USER NAME>>
  - Password: <<ADD PASSWORD>>
2. Access the tasks using the following link:
  - a. <https://workersandbox.mturk.com/projects/3FDHSGPTTFDC20PD6HAW5953E799Y/tasks?>
3. This is a worker's interface, where you can preview the tasks and attempt to solve them yourself.

| Requester             | Title  | HITs | Reward | Created | Actions               |
|-----------------------|--|------|--------|---------|-----------------------|
| DeepenAI              | DeepenAI Up [3]  | 18   | \$0.30 | 1m ago  | Preview Accept & Work |
| Rosni Kottekulam Vasu | Find a suitable relation for the given entities                      | 3    | \$0.05 | 1m ago  | Preview Accept & Work |
| SmartOne              | SmartOne Upgraded [3]  | 12   | \$0.30 | 3m ago  | Preview Accept & Work |
| SmartOne              | SmartOne Upgraded [4]  | 11   | \$0.30 | 6m ago  | Preview Accept & Work |
| SmartOne              | SmartOne Upgraded [2]  | 34   | \$0.30 | 11m ago | Preview Accept & Work |
| DeepenAI              | DeepenAI Up [2]  | 16   | \$0.30 | 15m ago | Preview Accept & Work |
| DeepenAI              | DeepenAI Up [1]  | 17   | \$0.30 | 15m ago | Preview Accept & Work |
| david hailey          | Beverage Counting HIT  | 247  | \$0.15 | 17m ago | Preview Quality       |
| david hailey          | Beverage Bounding Box HIT  | 150  | \$0.15 | 17m ago | Preview Quality       |
| Terry2094014          | specify-destination-test-1114  | 30   | \$0.10 | 18m ago | Preview Quality       |
| vse-mturk             | US Proactive Nov-14-2022 - Timestamp Experiment Remoderation Request | 7    | \$0.00 | 31m ago | Preview Quality       |

## Step 3: Get familiar with MTurk GUI as a Requestor

1. Choose an existing template to start with (one that is more suited to your task).

Select a template similar to your task

Provide a customizable template to start a new project

Survey Link Instructions (Click to expand)

Survey Link:

Provide the survey code here:

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

## 2. Try to understand the various components of the task and the settings that have been presented.

The following are some suggestions:

- Give a more intuitive title to the job
- Change the instructions such that they are more clear
- Change the settings of the job (e.g., payment, number of judgments per unit)

Mechanical Turk Developer Sandbox  
This site is for test and development only. [Learn more »](#)

amazonmturk Requester Create Manage Developer

New Project New Batch with an Existing Project

Your project was successfully saved.

### Edit Project

1 Enter Properties

2 Design Layout

3 Preview and Finish

Project Name: Knowledge Graph Completion This name is not displayed to Workers.

Describe your survey to Workers

Title

Find a suitable relation for the given entities

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

Identify suitable relation for the given entities

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

Keywords

Knowledge graph, link, fact

Provide keywords that will help Workers search for your tasks.

Setting up your survey

Reward per response

\$ 0.05

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

2

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

Time allotted per Worker

1 Hours

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

Survey expires in

1 Days

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

Auto-approve and pay Workers in

3 Days

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

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:

HIT Approval Rate (%)

greater than

11

Remove

(+) Add another criterion (up to 4 more)

(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.

Task Visibility (What is task visibility?)

☒ Public - All Workers can see and preview my tasks

☐ Private - All Workers can see my tasks, but only Workers that meet all Qualification requirements can preview my tasks

☐ Hidden - Only Workers that meet my Qualification requirements can see and preview my tasks

Save

Design Layout

NEW FEATURE: Amazon SageMaker Ground Truth Plus: Fully managed data labeling service. [Learn more »](#)

[Help](#) [Pricing](#) [Contact Us](#) [Policies](#) [State Licensing](#) [Contact Sales](#) [Blog](#) [Service Health Dashboard](#) [Careers](#) [MTurk.com](#) [Workers](#) [Follow Us on Twitter](#)

©2009-2022 Amazon.com, Inc. or its Affiliates. All rights reserved.

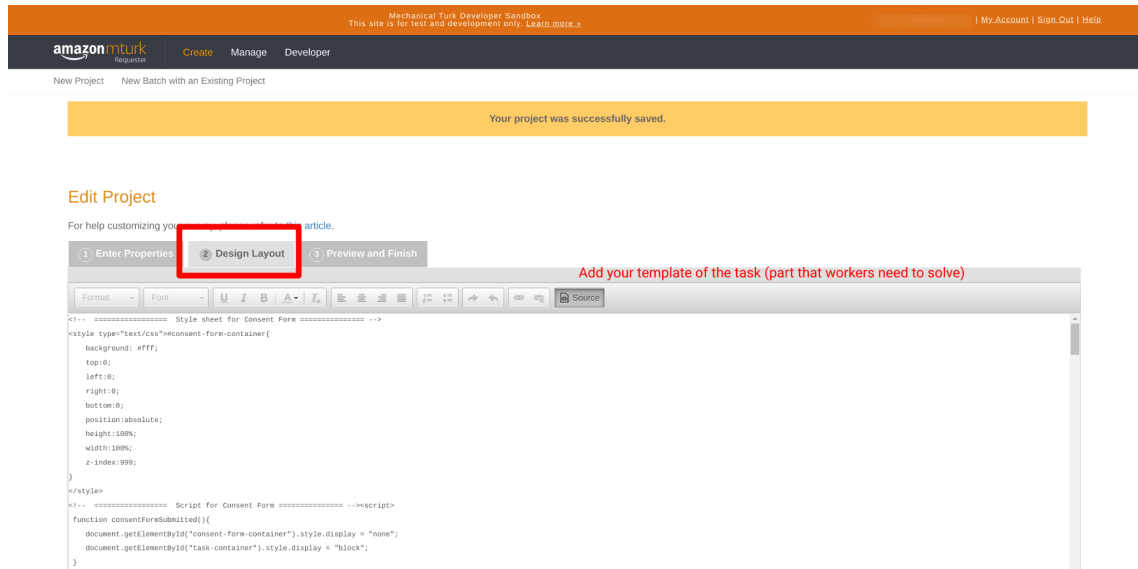
An amazon company

### 2.1. Understand the basic settings of the crowd workers

- **Master workers:**
  - Workers who demonstrated excellency across a wide range of tasks
- **Additional qualification**
  - **Location**
  - **HIT Approval Rate**
  - **Premium qualification like employment status, gender etc.**

### 3. Check the design of the crowdsourcing task

- a. You can add the source file (html css, javascript here)



### 4. Before publishing, preview the task to ensure that everything is as intended.

amazonmturk

Requester

Create Manage Developer

Mechanical Turk Developer Sandbox  
This site is for test and development only. [Learn more »](#)

My Account | Sign Out | Help

New Project New Batch with an Existing Project

Your project was successfully saved.

### Edit Project

This is how your task will look to Mechanical Turk Workers. Before you publish these tasks, any variables (eg \${variable\_name}) in the layout will be replaced with the input data that you provide when you publish your batch. You can [download a sample](#) of the CSV input file for this project or learn more about [acceptable file formats](#).

1 Enter Properties

2 Design Layout

3 Preview and Finish

Find a suitable relation for the given entities

Requester: 7Reward: \$0.05 per taskTasks available: 0Duration: 1 Hours

Qualifications Required: HIT Approval Rate (%) greater than 11

Knowledge Graph Completion Instructions (Click to expand)

InstructionsShortcuts

Select the relation which defines suitably the two entities in the given sentence.  
\$(subject) \_\_\_\_\_ \$(object)

Question 1: Which answer correctly fills in the blanks?  

☐ \$(option1)☐ \$(option2)☐ \$(option3)☐ \$(option4)☐ \$(option5)

Question 2: Which wikidata or wikipedia link did you use to arrive at your answer [required]?

Question 3: Which sentence(s) or information from wikidata or wikipedia did you use to arrive at your answer [required]?

Preview your design here  
and finish the task

Finish

NEW FEATURE: Amazon Sagemaker Ground Truth Plus: Fully managed data labeling service. [Learn more »](#)

Help

Pricing

Contact Us

Policies

State Licensing

Contact Sales

Blog

Service Health Dashboard

Careers

MTurk.com

Workers

Follow Us on Twitter

©2005-2022 Amazon.com, Inc. or its Affiliates. All rights reserved.

An amazon company

## 5. The task can be duplicated and modified at any time, or it can be created from scratch and modified later.

amazonmturk

Requester

Create Manage Developer

Mechanical Turk Developer Sandbox  
This site is for test and development only. [Learn more »](#)

My Account | Sign Out | Help

New Project New Batch with an Existing Project

Your project was successfully saved.

Start a New Batch with an Existing Project

| Project Name               | Title   | Created ▼         | Last Edited       |  |
|----------------------------|---|-------------------|-------------------|--|
| Knowledge Graph Completion | Find a suitable relation for the given entities | November 13, 2022 | November 13, 2022 | <div>Publish Batch</div> <div>Edit</div> <div>Copy</div> <div>Delete</div> |

Publish it when  
it is readyYou can edit the existing  
task templates

NEW FEATURE: Amazon Sagemaker Ground Truth Plus: Fully managed data labeling service. [Learn more »](#)

Help

Pricing

Contact Us

Policies

State Licensing

Contact Sales

Blog

Service Health Dashboard

Careers

MTurk.com

Workers

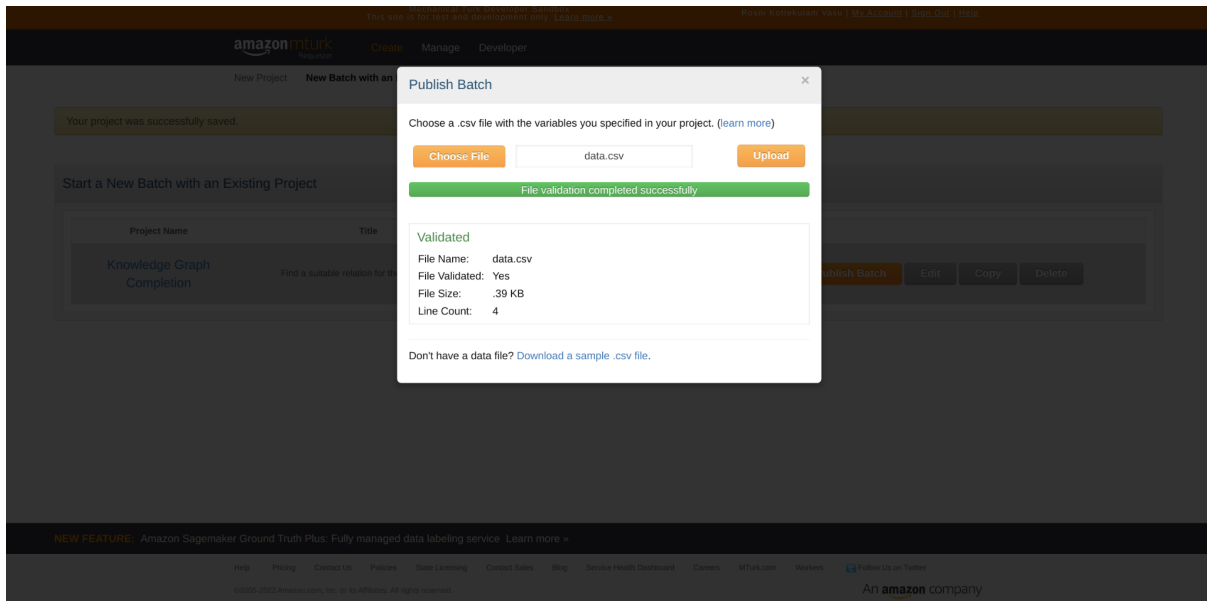
Follow Us on Twitter

©2005-2022 Amazon.com, Inc. or its Affiliates. All rights reserved.

An amazon company

## 6. Each task can be published as batches.

- a. You might have included input variables in the task template, and you can upload the data file (in CSV format) that contains the variables you included in your task template.



## 7. Launch the job and finalize the settings of the task

- a. After adding the data file, the task is ready to publish. Ensure that every task property is correct and you can also review the cost summary.

Mechanical Turk Developer Sandbox  
This site is for test and development only. [Learn more »](#)

amazon **mturk**  
Requester

CreateManageDeveloper

My Account | Sign Out | Help

New ProjectNew Batch with an Existing Project

Confirm and Publish Batch

1 Preview2 Confirm and Publish

Please review the information about the Batch, then click "Publish".

Knowledge Graph Completion

Batch Summary

Batch Name:Knowledge Graph Completion 1

Description:Identify suitable relation for the gi

Batch Properties

Title:Find a suitable relation for the given entities

Description:Identify suitable relation for the given entities

Batch expires in:1 Days

Results are auto-approved and Workers are paid after:3 Days

Workers must meet the following Qualifications to work on these tasks:

HIT Approval Rate (%) score greater than 11

Tasks

Number of tasks in this batch:3

Number of assignments per task:

x2

Total number of assignments in this batch:6

Task details and number of workers doing each task (assignments)

Cost Summary

Reward per Assignment:\$0.05

x6

(total number of assignments in this batch)

Estimated Total Reward:\$0.30

Estimated Fees to Mechanical Turk:

+ \$0.06

(fee details)

Estimated Cost:\$0.36

Applied Prepaid HITs Balance:

- \$0.36

Remaining Balance Due:\$0.00

This computes the cost for the current batch  
Remember, you can publshih you task  
with batch by batch data

Back

Publish

NEW FEATURE: Amazon SageMaker Ground Truth Plus: Fully managed data labeling service. [Learn more »](#)

HelpPricingContact UsPoliciesState LicensingContact SalesBlogService Health DashboardCareersMTurk.comWorkers

[Follow Us on Twitter](#)

©2005-2022 Amazon.com, Inc. or its Affiliates. All rights reserved.

An **amazon** company

8. Monitor the job

Mechanical Turk Developer Sandbox

This site is for test and development only. [Learn more »](#)

[My Account](#) | [Sign Out](#) | [Help](#)

amazonmturk

Create Manage Developer

Results Workers Qualification Types

Manage Batches > Batch Details

Knowledge Graph Completion 1

View the latest status of this batch, make changes, or get results.

Identify suitable relation for the given entities

Status

Status: In Progress

16% submitted

100% published

Assignments Completed: 1 / 6

Average Time per Assignment: 53 seconds

Creation Time: November 13, 2022 9:57 PM PST

Estimated Completion Time: November 13, 2022 10:16 PM PST (TODAY)

Settings

Knowledge Graph Completion

View Project

Note: If you have edited the Project after publishing this Batch, you will see the latest version.

Description:

Identify suitable relation for the given entities

Keywords:

Knowledge graph, link, fact

Qualification Requirement(s):

HIT Approval Rate (%) greater than 11

Number of Assignments per task: 2

Reward per Assignment: \$0.05

Input File: [data.csv](#)

Batch expires on: November 14, 2022 9:57 PM PST (Monday)

[Extend this batch](#)

Assignment duration: 1 hour

Auto Approval Delay: 3 days

Results

Assignments pending review: 1

Assignments approved: 0

Assignments rejected: 0

Cost Summary

Estimated Total Reward: \$0.30

Estimated Fees to Mechanical Turk: \$0.06 ([fee details](#))

Estimated Total Cost: \$0.36

These costs are only an estimate until all of the assignments have been submitted and reviewed.

Example task from this Batch

Can you find the relation between entities?

Information Sheet

We invite fluent English speakers to participate in an experiment on knowledge graph completion task.

Before you decide to participate, it is important that you understand the purpose of the research you will be participating in and what will happen to the data collected from you.

The study aims to collect knowledge triplets in the form of (subject, predicate, object). You have to identify the best relation suitable with the given entities. The resulting dataset will be used to train knowledge graph completion models.

Participants must be:

- Over 18 years of age
- Fluent English speakers: both native and non-native are welcome
- Use English in my everyday communication

How will the experiment work?

- You are asked to first accept the consent form.
- If you accept the consent form, you will be eligible to participate in the task. You will be asked to identify a best suitable relation for the given subject and object entities.
- Each HIT should not take more than 2 minutes to complete.

You will be reading content extracted from web. The nature of the phrases/terms are such that they are unlikely to cause offence, disturbance or distress to readers from a general audience. However, if you feel uncomfortable with the content of any of the task, please let us know.

NEW FEATURE: Amazon Sagemaker Ground Truth Plus: Fully managed data labeling service [Learn more »](#)

[Help](#) [Pricing](#) [Contact Us](#) [Policies](#) [State Licensing](#) [Contact Sales](#) [Blog](#) [Service Health Dashboard](#) [Careers](#) [MTurk.com](#) [Workers](#) [Follow Us on Twitter](#)

©2009-2022 Amazon.com, Inc. or its Affiliates. All rights reserved.

An amazon company

## 9. Inspect the results

- Download the complete report on the decisions, which includes input details, worker details, etc.



Mechanical Turk Developer Sandbox  
This site is for test and development only. [Learn more...](#)

My Account | Sign Out | Help

amazon **mturk**  
Requester

Create Manage Developer

Results Workers Qualification Types

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](#).

Knowledge Graph Completion 1

Customize View Filter Results

Download and upload results to evaluate it offline

Upload CSV Download CSV

2 of 2 assignments (FILTER APPLIED: only show assignments that are in 'Submitted' status)

Approve Reject

|                          | HIT ID ▲                      | Worker ID     | Lifetime Approval Rate | Input.Subject     | Input.Object  | Input.Option1 | Input.Option2        | Input.Option3        | Input.Option4 | Input.Option5 | Option1.Option 1 | Option2.Option 2 | Opt   |
|--------------------------|-------------------------------|---------------|------------------------|-------------------|---------------|---------------|----------------------|----------------------|---------------|---------------|------------------|------------------|-------|
| <input type="checkbox"/> | 3K2CEDRACB7AY1UMHGFH8UID31TMJ | A22BXIPVQWKL2 | 99% (402/405)          | Leonard Nimoy     | Star Trek     | won the       | is or was married to | Painted              | starred in    | character in  | false            | true             | false |
| <input type="checkbox"/> | 3U74KRR67MRGGJFMTHUMUNJITBTNJ | A22BXIPVQWKL2 | 99% (402/405)          | Leonardo da Vinci | The Mona Lisa | was born in   | won the              | is or was married at | Painted       | destroyed     | false            | true             | false |

Approve Reject

2 of 2 assignments (FILTER APPLIED: only show assignments that are in 'Submitted' status)

You can approve or reject the results here by selecting and click on

NEW FEATURE: Amazon Sagemaker Ground Truth Plus: Fully managed data labeling service [Learn more](#) »

Help Pricing Contact Us Policies State Licensing Contact Sales Blog Service Health Dashboard Careers MTurk.com Workers [Follow Us on Twitter](#)

©2005-2022 Amazon.com, Inc. or its Affiliates. All rights reserved.

An **amazon** company

## 10. Review the assignments by either “Approve” or “Reject”.

- If it is a “Reject”, you have to mention the reason for rejecting the assignment.

Mechanical Turk Developer Sandbox  
This site is for test and development only. [Learn more...](#)

My Account | Sign Out | Help

amazon **mturk**  
Requester

Create Manage Developer

Results Workers Qualification Types

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](#).

Knowledge Graph Completion 1

Customize View Filter Results

Upload CSV Download CSV

1 of 2 assignments (FILTER APPLIED: only show assignments that are in 'Submitted' status)

Approve Reject

|                                     | HIT ID ▲                      | Worker ID     | Lifetime Approval Rate | Input.Subject | Input.Object | Input.Option1 | Input.Option2        | Input.Option3 | Input.Option4 | Input.Option5 | Option1.Option 1 | Option2.Option 2 | Option3.Option 3 |
|-------------------------------------|-------------------------------|---------------|------------------------|---------------|--------------|---------------|----------------------|---------------|---------------|---------------|------------------|------------------|------------------|
| <input checked="" type="checkbox"/> | 3K2CEDRACB7AY1UMHGFH8UID31TMJ | A22BXIPVQWKL2 | 99% (403/406)          | Leonard Nimoy | Star Trek    | won the       | is or was married to | Painted       | starred in    | character in  | false            | true             | false            |

Approve Reject

Reject Assignments

Give Workers a reason for why you're rejecting their work

☒ Republish the rejected assignment for other Workers to complete

Cancel Reject

## 11. You can manage and view the existing workers.

Mechanical Turk Developer Sandbox  
This site is for test and development only. [Learn more »](#)

amazonmturk  
Requester

CreateManageDeveloper

ResultsWorkersQualification Types

Manage Workers

The Workers who have completed work for you are listed below. Select a Worker ID to bonus, block, unblock, assign a Qualification, or revoke a Qualification. To block, unblock, or change Qualification settings for multiple Workers, select Download CSV. Select Customize View to change which Qualification Types are displayed in the table below.

Customize View

Download CSVUpload CSV

Show my Workers by: LifetimeLast 30 daysLast 7 days

| Workers        |  |               |              |
|----------------|--|---------------|--------------|
| Worker ID ▲    | Last 7 days Approval Rate for Your tasks | Qual: test... | Block Status |
| A22BXIPVQWKLM2 | 75% (3/4)                                |               | Unblocked    |

\*NOTE: Workers are shown in the list if they have accepted and submitted an assignment in the time period selected. Workers may appear in the list but have a "0/0" approval rate displayed because you have not yet approved or rejected their work.

NEW FEATURE: Amazon Sagemaker Ground Truth Plus: Fully managed data labeling service [Learn more »](#)

HelpPricingContact UsPoliciesState LicensingContact SalesBlogService Health DashboardCareersMTurk.comWorkersFollow Us on Twitter

©2005-2022 Amazon.com, Inc. or its Affiliates. All rights reserved.

An amazon company

12. You can view and manage the existing Qualifications.

Mechanical Turk Developer Sandbox  
This site is for test and development only. [Learn more »](#)

amazonmturk  
Requester

CreateManageDeveloper

ResultsWorkersQualification Types

Manage Qualification Types

Below is a list of your Qualification Types and the corresponding number of Workers.

Create New Qualification Type

| Qualification Types |               |                                |                                     |                   |  |
|---------------------|---------------|--------------------------------|-------------------------------------|-------------------|--|
|                     | Name ▼        | ID                             | Workers who have this Qualification | Creation Date     | Description                                |
| ⓘ                   | test          | 3ZPJ66GO0GK31ET1D0AQZWU4R5FLOF | 0                                   | November 17, 2021 | test                                       |
| ⓘ                   | pre-screening | 36GZRO2Y1D7WY5RMP45HEIT2W6AF0B | 1                                   | November 15, 2021 | This is a brief knowledge test for workers |
| ⓘ                   | Pre-screening | 381E11SBE4IEZ00RMDGT2TQED843   | 0                                   | November 15, 2021 | This is a brief knowledge test for workers |
| ⓘ                   | No Repeat     | 3SIXTRR80320MHKLJMPZ1QU851UKWW | 0                                   | May 8, 2021       | Same HITs are not repeated for a worker    |

NEW FEATURE: Amazon Sagemaker Ground Truth Plus: Fully managed data labeling service [Learn more »](#)

HelpPricingContact UsPoliciesState LicensingContact SalesBlogService Health DashboardCareersMTurk.comWorkersFollow Us on Twitter

©2005-2022 Amazon.com, Inc. or its Affiliates. All rights reserved.

An amazon company