

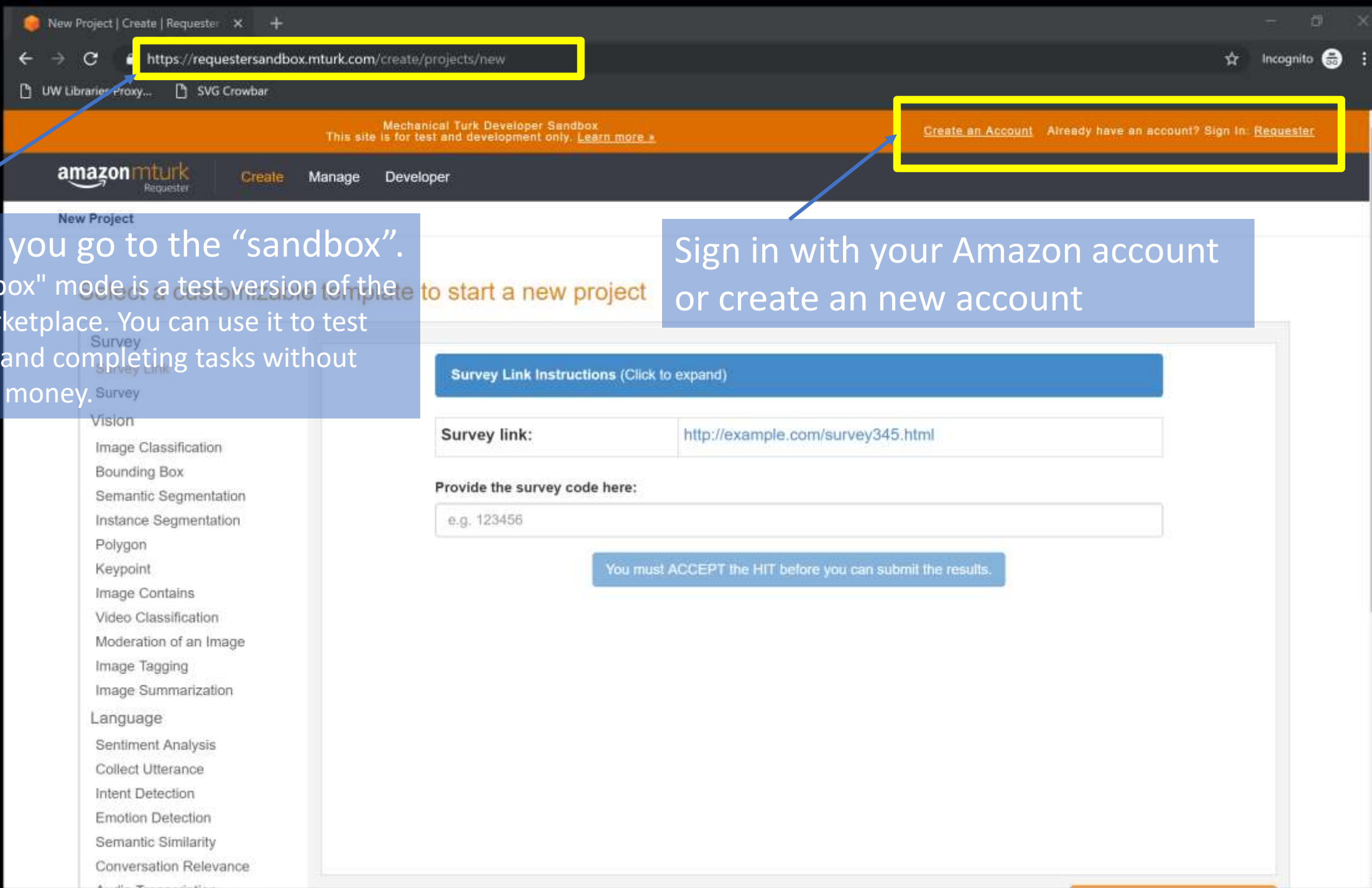
# Lab 3 Walkthrough

# Mechanical Turk

Crowdsourcing



- Mechanical Turk is a work infrastructure that allows people to perform quick tasks over the web for a small payment
- Human Intelligence Task (HIT)
  - a task that a **Requester** submits to Amazon Mechanical Turk for **Workers** to perform, e.g., transcribe an audio, tag an image
- The **requester** creates the HIT and pays the workers
- The **worker** chooses a HIT they like and are qualified to do



Make sure you go to the “sandbox”.

- The "Sandbox" mode is a test version of the MTurk marketplace. You can use it to test publishing and completing tasks without paying any money.

Sign in with your Amazon account or create a new account

# Steps

- Create a HIT
  - design the interface for annotation
  - configure the job properties
- Publish a batch
  - upload data for annotation
- Monitor and review results

New Project | Create | Requester x +

https://requestersandbox.mturk.com/create/projects/new

UW Libraries Proxy... SVG Crowbar

New Project New Batch with an Existing Project

### 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

Collect Utterance

**Intent Detection**

Emotion Detection

Semantic Similarity

Conversation Relevance

Audio Transcription

Document Classification

Translation Quality

Audio Naturalness

Other

≡

Pick the most relevant intention expressed by the below text

Hello, I'd like a large pizza

Select an option

Order	1
Cancel	2
Return	3

Submit

Create Project

Create a project using the provided template

## Edit Project

Specify the properties that are common for all of the tasks created using this project.

1 Enter Properties

2 Design Layout

3 Preview and Finish

Project Name: Intent Detection

This name is not displayed to Workers.

Project Name (not displayed to workers)

### Describe your task to Workers

Title

Select the most relevant intention expressed by the text

Describe the task 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

Select the most relevant intention expressed by the text

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

Keywords

text, classification, intent

Provide keywords that will help Workers search for your tasks.

Task information (displayed to workers)

- be informative

### Setting up your task

Reward per assignment

\$

This is how much a Worker will be paid for completing an assignment. Consider how long it will take a Worker to complete each assignment.

Number of assignments per task

3

How many unique Workers do you want to work on each task?

Time allotted per assignment

1

Hours

Keywords:

Provide keywords that will help Workers search for your tasks.

### Setting up your task

Reward per assignment

\$

This is how much a Worker will be paid for completing an assignment. Consider how long it will take a Worker to complete each assignment.

Number of assignments per task

How many unique Workers do you want to complete this task?

Time allotted per assignment

Hours

Maximum time a Worker has to work on a single task. If you set this too low, your task may be rushed.

Task expires in

Days

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

Auto-approve and pay Workers in

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:

The money you pay for each assignment.

- In the sandbox, your Current Balance is always \$10,000.00.

Number of workers (annotations) you want for each datapoint

- it is usually good to have more than one annotations on each datapoint

Within this time frame after an assignment is submitted, you can reject a worker's annotation. After this time frame, the worker will be paid for the submitted assignment.

Create | Requester | Amazon Me x +

← → ↺ https://requestersandbox.mturk.com/create/projects/83985/edit ☆ Incognito 36

UW Libraries Proxy... SVG Crowbar

How many unique Workers do you want to work on each task?

Time allotted per assignment

Maximum time a Worker has to work on a single task. Be generous so that Workers are not rushed.

Task expires in

Maximum time your task 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.

**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:

(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

You can set requirements (e.g., geographic locations) to only allow certain types of workers

- keep them to default for this lab

Save and go to the next page



## HTML Elements Reference

<https://docs.aws.amazon.com/sagemaker/latest/dg/sms-ui-template-reference.html>

### Edit Project

Use the HTML editor below to design the layout of your task, which Workers will complete.

You can use any HTML, CSS, or JavaScript to customize your layout.

Or use Crowd HTML Elements to simply design your layout.

This layout is common for all of the tasks created with this project. You can define variables (eg `$(variable_name)`) for data that will vary from task to task based on your CSV input file ([Learn more](#)).

1 Enter Properties 2 Design Layout 3 Preview and Finish

```
8 <crowd-form answer-format="flatten-objects">
9
10 <!-- The crowd-classifier element will create a tool for the Worker to select the
11 correct answer to your question -->
12 <crowd-classifier
13   categories="['Order', 'Cancel', 'Return']"
14   header="Pick the most relevant intention expressed by the below text"
15   name="intent">
16
17   <!-- The text you want classified will be substituted for the "text" variable when you
18   publish a batch with a CSV input file containing multiple text items -->
19   <classification-target> ${text} </classification-target>
20
21   <!-- Use the short-instructions section for quick instructions that the Worker
22   will see while working on the task. Including some basic examples of
23   good and bad answers here can help get good results. You can include
24   any HTML here. -->
25
26   <short-instructions> SHORT INSTRUCTIONS TBD </short-instructions>
27
28   <!-- Use the full-instructions section for more detailed instructions that the
29   Worker can open while working on the task. Including more detailed
30   instructions and additional examples of good and bad answers here can
31   help get good results. You can include any HTML here. -->
32   <full-instructions header="Intent Detection Instructions">
33     <p>Select the most relevant intention expressed by the text.</p>
34     <div>
35       <p><strong>Example: </strong>I would like to return a pair of shoes</p>
36       <p><strong>Intent: </strong>Return</p>
37     </div>
38   </full-instructions>
39
40 </crowd-classifier>
```

Categories the worker can choose from.

`${text}`: the 'text' column in the csv file you are going to upload

Instructions

Create | Requester | Amazon ME... X +

← → ↺ 🔒 https://requestersandbox.mturk.com/create/projects/83985

UW Libraries Proxy... SVG Crowbar

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

Select the most relevant intention expressed by the text

Requester: UWLib506conval Reward: \$0.02 per task Tasks available: 0 Duration: 1 Hours

Qualifications Required: None

**Previewing Answers Submitted by Workers**

This message is only visible to you and will not be shown to Workers.

You can test completing the task below and click "Submit" in order to preview the data and format of the submitted results.

**Instructions**

[View full instructions](#)

[View tool guide](#)

SHORT INSTRUCTIONS TBD

Pick the most relevant intention expressed by the below text

`${text}`

Select an option

Order	1
Cancel	2
Return	3

Finish

`<short-instructions> ...`

`<classification-target> ...`

`<crowd-classifier categories="[...]" ...`

Create | Requester | Amazon Me... X +

← → ↺ 🔒 https://requestersandbox.mturk.com/create/projects/83985

📄 UW Libraries Proxy... 📄 SVG Crowbar

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

Select the most relevant intention expressed by the text

Requester: unwe506conval Reward: \$0.02 per task Tasks available: 0 Duration: 1 Hours

Qualifications Required: None

Presenting Answers Submitted by Workers

This message is only visible to you and will not be shown to workers. You can test composing the text below and click "Submit" in order to preview the data and format of the submitted results.

Instructions

View full instructions

SHORT INSTRUCTIONS TEST

#### Intent Detection Instructions

Select the most relevant intention expressed by the text.

**Example:** I would like to return a pair of shoes

**Intent:** Return

Close

Finish

<full-instructions> ...

# Steps

- Create a HIT
  - design the interface for annotation
  - configure the job properties
- Publish a batch
  - upload data for annotation
- Monitor and review results

New Batch with an Existing Project X +

← → ↻ <https://requestersandbox.mturk.com/create/projects> 🔍 ☆ Incognito

UW Libraries Proxy... SVG Crowbar

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

uwae586convex | My Account | Sign Out | Help

amazon **mturk** Requester Create Manage Developer

New Project **New Batch with an Existing Project**

Your Project was successfully deleted.

Start a New Batch with an Existing Project

Project Name	Title	Created ▼	Last Edited	
Intent Detection	Select the most relevant intention expressed by the text	April 30, 2019	April 30, 2019	<div><div>Publish Batch</div><div>Edit</div><div>Copy</div><div>Delete</div></div>

Publish a batch to workers

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

©2019-2019 Amazon.com, Inc. or its Affiliates. All rights reserved. An amazon company



AutoSave ☐ off

File Home Insert Draw

**POSSIBLE DATA LOSS** Some features may not work as expected.

C7

	A	B	C
1	text		
2	Hit1_text_data		
3	Hit2_text_data		
4	Hit3_text_data		
5			
6			
7			
8			
9			
10			

**Publish Batch**

Choose a .csv file with the variables you specified in your project. [\(learn more\)](#)

Don't have a data file? [Download a sample .csv file.](#)

For the toy HIT, using the downloaded csv file

`${text}`

## Preview Tasks

[1 Preview](#) [2 Confirm and Publish](#)

This is how your task will look to Workers. Make sure that any variables in the task are correctly replaced by your input data, then click "Next".

### Intent Detection

Select the most relevant intention expressed by the text

Requester: uwae506convai

Reward: \$0.02 per task

Tasks available: 3

Duration: 1 Hour

Qualifications Required: None

### Instructions

[View full instructions](#)

[View tool guide](#)

SHORT INSTRUCTIONS TBD

Pick the most relevant intention expressed by the below text

Hit1\_text\_data

Select an option

Order	1
Cancel	2
Return	3

$\text{\$}\{\text{text}\}$

Submit

Showing Task 1 of 3 [Next HIT](#)

[Cancel](#) [Next](#)

## Confirm and Publish Batch

[1 Preview](#) [2 Confirm and Publish](#)

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

### Intent Detection

#### Batch Summary

Batch Name: Intent Detection 1

Description: Select the most relevant intention

#### Batch Properties

Title: Select the most relevant intention expressed by the text  
Description: Select the most relevant intention expressed by the text  
Batch expires in: 7 Days  
Results are auto-approved and Workers are paid after: 3 Days

#### Tasks

Number of tasks in this batch: 3  
Number of assignments per task: x 3  
Total number of assignments in this batch: 9

#### Cost Summary

Reward per Assignment: \$0.02  
x 9 (total number of assignments in this batch)  
Estimated Total Reward: \$0.18  
Estimated Fees to Mechanical Turk: + \$0.09 (fee details)  
Estimated Cost: \$0.27  
Applied Prepaid HITs Balance: - \$0.27  
Remaining Balance Due: \$0.00


In the sandbox, your Current Balance is always \$10,000.00.

[Back](#)[Publish](#)



Results | Manage | Requester | [An](#) X +

← → ↺ 🔒 https://requestersandbox.mturk.com/batches/236231

🔍 ☆ Incognito 

📄 UW Libraries Proxy...

📄 SVG Crowbar

Results Workers Qualification Types

Manage Batches > Batch Details

Intent Detection 1

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

Select the most relevant intention expressed by the text

Status

Cancel

Status: In Progress

0% submitted

100% published

Assignments Completed: 0 / 9

Average Time per Assignment: Not yet available

Creation Time: April 30, 2019 5:35 PM PDT

Estimated Completion Time: Not yet available

Settings

Intent Detection

View Project

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

Description: Select the most relevant intention expressed by the text.

Keywords: text, classification, intent

Qualification Requirement(s):

Number of Assignments per task: 3

Reward per Assignment: \$0.02

Input File: [input.csv](#)

Batch expires on: Not yet available

Assignment duration: 1 hour

Auto Approval Delay: 3 days

Results

Results

Assignments pending review: 0

Assignments approved: 0

Assignments rejected: 0

Cost Summary

Estimated Total Reward: \$0.18

Estimated Fees to Mechanical Turk: \$0.09 [\(fee details\)](#)

Estimated Total Cost: \$0.27

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

Example task from this Batch

☰

Pick the most relevant intention expressed by the below text

Progress of the data collection

- go to the Worker sandbox  
<https://workersandbox.mturk.com>
- search for your job name or your account name

Results | Manage | Requester | Amazon Mechanical Turk

https://workersandbox.mturk.com/projects/?end\_signin=1&openid.pape\_max\_auth\_age=43200&openid.identity=https%3A%2F%2Fwww.amazon.com%2Fap%2Fid%2Fa... Incognito

UW Libraries Proxy... SVG Crowbar

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

Worker ID: A45A98MU302QT Hello, uwee596convai | [Sign Out](#)

amazonmturk Worker **HITS** Dashboard Qualifications  [Filter](#)

**All HITS** Your HITS Queue

HIT Groups (1-20 of 1304)

[Requester](#) [Type](#)

[Rece Capture](#) Receipt Transcription

[vse-mturk](#) Algorithm tagged videos - TimelineUI May-01-2019

[Robert Atwood](#) Hummingbird Watch - Early Forest Fire Detection

[uwee596convai](#) Select the most relevant intention expressed by the text

[Ethan Perez](#) Guess the answer!

[vse-mturk](#) Customer Low Priority feedback request videos - COPSUI May-01-2019

[Ethan Perez](#) Guess the answer!

[vse-mturk-gamma](#) vsegamma testing Nudity May-01-2019

[vse-mturk-gamma](#) vsegamma testing External URL May-01-2019

[Ethan Perez](#) Guess the answer!

[vse-mturk](#) Customer High Priority feedback request videos - COPSUI May-01-2019

[Mirai](#) Identify images that contain markings (measurements, logo, text, etc.)

[Mirai](#) Identify sketches, or images of packaging or images with borders

[Show Details](#) [Hide Details](#) Items Per Page: 20

	HITS	Reward	Created	Actions	
	1,357	\$0.03	1m ago	<a href="#">Preview</a>	<a href="#">Qualify</a>
	5	\$0.00	2m ago	<a href="#">Preview</a>	<a href="#">Qualify</a>
	2	\$0.07	2m ago	<a href="#">Preview</a>	<a href="#">Accept &amp; Work</a>
	3	\$0.02	4m ago	<a href="#">Preview</a>	<a href="#">Accept &amp; Work</a>
	29	\$1.50	6m ago	<a href="#">Preview</a>	<a href="#">Accept &amp; Work</a>
	2	\$0.00	11m ago	<a href="#">Preview</a>	<a href="#">Qualify</a>
	28	\$1.12	11m ago	<a href="#">Preview</a>	<a href="#">Accept &amp; Work</a>
	6	\$0.00	12m ago	<a href="#">Preview</a>	<a href="#">Qualify</a>
	1	\$0.00	16m ago	<a href="#">Preview</a>	<a href="#">Qualify</a>
	25	\$1.12	25m ago	<a href="#">Preview</a>	<a href="#">Accept &amp; Work</a>
	2	\$0.00	31m ago	<a href="#">Preview</a>	<a href="#">Qualify</a>
	20	\$0.03	38m ago	<a href="#">Preview</a>	<a href="#">Accept &amp; Work</a>
	21	\$0.03	39m ago	<a href="#">Preview</a>	<a href="#">Accept &amp; Work</a>

Results | Manage | Requester | Amazon Mechanical Turk

https://workersandbox.mturk.com/projects/?filters%5Bsearch\_term%5D=uwee596convai&page\_size=20&page\_number=1&sort=updated\_desc&filters%5Bmin\_reward%5D=0.02

UW Libraries Proxy... SVG Crowbar

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

Worker ID: A45A98MU302QT Hello, uwee596convai | [Sign Out](#)

amazonmturk Worker **HITS** Dashboard Qualifications uwee596convai

All HITS Your HITS Queue

1-1 of 1 results containing 'uwee596convai'

### HIT Groups

Show Details Hide Details Items Per Page: 20

Requester	Title	HITS	Reward	Created	Actions
uwee596convai	Select the most relevant intention expressed by the text	3	\$0.02	4m ago	<a href="#">Preview</a> <a href="#">Accept &amp; Work</a>

Previous 1 Next

Start to work on the HIT by yourself

Help Contact Legal Service Health Feedback © 2005-2019 Amazon Mechanical Turk Inc. or its affiliates. All rights reserved. An amazon company

Results | Manage | Requester | Amazon Mechanical Turk

https://workersandbox.mturk.com/projects?filters%5Bmin\_reward%5D=0&filters%5Bsearch\_term%5D=uwee596convai&page\_number=1&page\_size=20&sort=updated\_...


UW Libraries Proxy... SVG Crowbar

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

Worker ID: A45A98MU302QT Hello, uwee596convai | [Sign Out](#)

amazonmturk Worker **HITS** Dashboard Qualifications uwee596convai

All HITS Your HITS Queue

 **There are no more of these HITS available**  
Browse [all available HITS](#).

1-1 of 1 results containing "uwee596convai"

### HIT Groups

[Show Details](#) [Hide Details](#) Items Per Page: 20

Requester	Title	HITS	Reward	Created	Actions
uwee596convai	Select the most relevant intention expressed by the text	3	\$0.02	5m ago	<a href="#">Preview</a> <a href="#">Accept &amp; Work</a>

[Previous](#) [Next](#)

The worker cannot annotate the same datapoint more than once.  
So after the worker finishes all 3 assignments, he/she cannot work on this batch any more.

[Help](#) [Contact](#) [Legal](#) [Service Health](#) [Feedback](#) © 2005-2019 Amazon Mechanical Turk Inc. or its affiliates. All rights reserved. An [amazon](#) company

Results | Manage | Requester | Amazon Mechanical Turk

https://requestersandbox.mturk.com/batches/236231

UW Libraries Proxy... SVG Crowbar

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

urwee596conva | My Account | Sign Out | Help

amazonmturk  
Requester

Create Manage Developer

Results Workers Qualification Types

Manage Batches > Batch Details

Intent Detection 1

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

Select the most relevant intention expressed by the text

Status

Status: In Progress

33% submitted

100% published

Assignments Completed: 3 / 9

Average Time per Assignment: 9 seconds

Creation Time: April 30, 2019 5:35 PM PDT

Estimated Completion Time: April 30, 2019 5:54 PM PDT (TODAY)

Settings

Intent Detection

[View Project](#)

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

Description: Select the most relevant intention expressed by the text

Keywords: text, classification, intent

Qualification Requirement(s):

Number of Assignments per task: 3

Reward per Assignment: \$0.02

Input File: [input.csv](#)

Batch expires on: May 07, 2019 5:35 PM PDT (Tuesday)

[Extend this batch](#)

Assignment duration: 1 hour

Auto Approval Delay: 3 days

Results

Assignments pending review: 3

Assignments approved: 0

Assignments rejected: 0

Cost Summary

Estimated Total Reward: \$0.18

Estimated Fees to Mechanical Turk: \$0.09 ([fee details](#))

Estimated Total Cost: \$0.27

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

Check annotations here

Progress of the data collection



Results | Manage | Requester | Amazon Mechanical Turk

https://requestersandbox.mturk.com/batches/236231/results

UW Libraries Proxy... SVG Crowbar

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

amazonmturk  
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.

#### Intent Detection 1

Customize View Filter Results

3 of 3 assignments (FILTER APPLIED: only show assignments that are in 'Submitted' status)

Approve Reject

	HIT ID	Worker ID	Lifetime Approval Rate	Input.Text	Intent.Label
<input type="checkbox"/>	338431Z1FMHH7V29YJ55DBT9UTOR1	A45A98MU302QT	0% (0/0)	Hit2_text_data	Order
<input type="checkbox"/>	36FFXPMSTAQ411062ETOR99NMLVHOX	A45A98MU302QT	0% (0/0)	Hit3_text_data	Return
<input type="checkbox"/>	3KTCJ4SCVH35ENIGN9BQ7I4BAWKMIQ	A45A98MU302QT	0% (0/0)	Hit1_text_data	Cancel

Approve Reject

3 of 3 assignments (FILTER APPLIED: only show assignments that are in 'Submitted' status)

Upload CSV Download CSV

- Input.text
- Answer.intent.label
- Approve
- Reject
- ...

You can approve or reject annotations. Annotations will be auto-approved based on the HIT configuration. (See "Auto Approve and Pay Worker in ..." in the HIT configuration page)

An amazon company

# Cohen's Kappa Coefficient

(Cohen, 1960)

- Test-Retest or Inter-rater reliability for categorical (typically dichotomous) data.
- Accounts for chance agreement

Observed		Rater 2		
		Present	Absent	Total
Rater 1	Present	20	15	35
	Absent	10	55	65
	Total	30	70	100

# Kappa Coefficient

Can you explain how to computed the expected table from the observed one?



Observed		Rater 2		
		Present	Absent	Total
Rater 1	Present	20	15	35
	Absent	10	55	65
	Total	30	70	100

Expected		Rater 2		
		Present	Absent	Total
Rater 1	Present	10.5	24.5	35
	Absent	19.5	45.5	65
	Total	30	70	100

$$\text{kappa} = \frac{P_o - P_e}{1.0 - P_e}$$

$P_o$  = observed proportion of agreements

$P_e$  = expected proportion of agreements

$$\text{kappa} = \frac{[(20+55)/100] - [(10.5+45.5)/100]}{1 - [(10.5+45.5)/100]} = 0.43$$



# Kappa Interpretation

- Interpretation:

<u>Kappa Value</u>	<u>Interpretation</u>
Below 0.00	Poor
0.00-0.20	Slight
0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Substantial
0.81-1.00	Almost perfect

(source: Landis, J. R. and Koch, G. G. 1977. *Biometrics* 33: 159-174)

- kappa could be high simply because marginal proportions are either very high or very low!!
- Best interpretation of kappa is to compare its values on other, similar scales